

-ИННОВАЦИИ В МИРЕ -

Дата поступления рукописи в редакцию: 17.02.2026 г.

DOI: 10.24412/2782-3830-2026-2-368-372

Дата принятия рукописи в печать: 18.03.2026 г.

КАСИХИН Лев Владимирович,
аспирант, Российская Государственная Академия
Интеллектуальной Собственности,
e-mail: lev.kasihin@mail.ru

ИСПОЛЬЗОВАНИЕ ОХРАНЯЕМЫХ ПРОИЗВЕДЕНИЙ ПРИ МАШИННОМ ОБУЧЕНИИ ГЕНЕРАТИВНОГО ИСКУССТВЕННОГО ИНТЕЛЛЕКТА

Аннотация. В статье рассматривается использование защищённых авторским правом произведений в процессе машинного обучения генеративного искусственного интеллекта. С учетом быстрого развития генеративных нейросетей, которые обучаются на больших объемах защищённых данных, возникает вопрос о законности такого использования произведений для тренировки ИИ. Этот вопрос становится особенно важным, поскольку в прошлом году имели место множество коллективных исков авторов оригинальных работ к компаниям-разработчикам генеративных нейросетей с требованием признать подобное использование незаконным. Оценка четырёх факторов, являющихся основой доктрины Fair Use, позволяет определить, может ли существующее использование произведений для обучения генеративных нейросетей быть признано судом добросовестным, а также выработать рекомендации для изменения подхода к обучению искусственного интеллекта в целом.

Ключевые слова: искусственный интеллект, комплексный подход, цифровые технологии, законопроект, генеративный искусственный интеллект, итеративный характер, право, риск-ориентированный подход, отраслевой подход.

KASIKHIN Lev Vladimirovich,
postgraduate student,
Russian State Academy of Intellectual Property

THE USE OF WORKS OF ART PROTECTED BY LAW IN MACHINE LEARNING OF GENERATIVE ARTIFICIAL INTELLIGENCE

Annotation. The article explores the use of works of art protected by law in machine learning of generative artificial intelligence. Due to the rapid development of generative neural networks, which are trained on large amounts of copyrighted data, the question arises of the legality of using such works to train artificial intelligence. The issue is also of particular relevance because over the past year, collective lawsuits have been repeatedly filed by the authors of original works against companies developing generative neural networks, demanding that such use be recognized as illegal. An assessment of the four factors underlying the Fair Use doctrine makes it possible to assess to what extent, in the existing paradigm, the use of works for training generative neural networks can be recognized by a court as fair, as well as to formulate a proposal for a systemic change in the approach to training artificial intelligence in general.

Key words: artificial intelligence, integrated approach, digital technology, bill, generative artificial intelligence, iterative nature, law, risk-based approach, industry approach.

Искусственный интеллект (ИИ) использует специально разработанные алгоритмы и базы данных, что в совокупности позволяет ему решать различные задачи, предлагать решения и строить прогнозы. Более того, ИИ способен к обучению и самообучению во время

решения поставленных вопросов. Проще говоря, благодаря соответствующим алгоритмам ИИ может самостоятельно улучшаться, анализируя входные данные. Эти функции обеспечиваются технологией машинного обучения (Machine Learning, МО), которая позволяет ИИ быстро обу-

чатся на больших объемах данных в автоматическом режиме, используя заложенные в систему алгоритмы, а не полагаясь на указания человека. Анализ юридической литературы показывает, что в настоящее время термин «машинное обучение» используется почти так же часто, как и термин «искусственный интеллект». Причем в одних случаях эти термины используются как тождественные, в других – как различающиеся [1], что нуждается в анализе в целях настоящего исследования.

В традиционных компьютерных программах разработчик создает набор правил и алгоритмов, строго следуя которым система выполняет действия. В отличие от этого, в машинном обучении (МО) модель получает входные данные и, анализируя их, формирует собственные алгоритмы для обработки последующей информации. Таким образом, в МО речь идет о обучении ИИ, а не о перепрограммировании алгоритмов принятия решений, которые при этом остаются заранее определенными. В литературе подчеркивается, что методы машинного обучения обучают машины достигать результатов, предоставляя множество примеров правильных ответов, а также могут позволять машине учиться на основе проб и ошибок, устанавливая набор правил [13]. Машинное обучение обычно используется для построения или корректировки модели поведения, но также может использоваться для интерпретации результатов модели» [2].

Между ИИ и МО не следует ставить знак равенства, так как МО является одной из технологий ИИ и представляет собой практический способ его применения. Это альтернативный подход по сравнению с ИИ, который функционирует исключительно на основе инструкций, заданных человеком.

Одной из серьезных проблем является возможность включения в обучающие данные конфиденциальной информации, такой как персональные данные, данные о геолокации и различные тайны: медицинские, банковские, коммерческие и другие. В связи с этим в Концепции развития регулирования технологий искусственного интеллекта и робототехники до 2024 года, утвержденной распоряжением Правительства РФ от 19.08.2020 № 2129-р, отмечена необходимость соблюдения баланса между требованиями защиты персональных данных и их использованием для обучения систем ИИ. Исследуя проблему соблюдения баланса между защитой данных и их использованием, В. Б. Наумов и Е. В. Тытюк отмечают, что данные можно защищать с помощью так называемых «маскировочных технологий» (safe harbour), получения согласия от субъектов данных и минимизации объема использованных данных. Авторы также указывают на законодательные усилия на разрешение данной

проблемы через экспериментальное регулирование, реализованное в Законе о внедрении ИИ, который допускает использование только анонимизированных персональных данных в машинном обучении. Однако в конечном итоге они считают это решение небесспорным, поскольку «обезличивание данных не всегда обеспечивает их полную защиту» [3].

В рамках данного исследования необходимо тщательно рассмотреть проблему использования авторских произведений в качестве обучающих данных. На практике это привело к искам авторов к разработчикам ИИ за нарушение авторских прав, так как произведения использовались без согласия их создателей. Первым значимым делом, связанным с защитой интеллектуальных прав художников, стало обращение Сары Андерсен, Келли МакКернан и Карлы Ортиз в суд против крупных компаний, занимающихся генеративным ИИ, таких как Stability AI, Midjourney и DeviantArt (Andersen v. Stability AI Ltd. (3:23-cv-00201)). В иске утверждалось, что эти компании нарушили права истцов, используя их художественные произведения для обучения без согласия. Ответчики, признавая факт использования, утверждали, что такая практика подпадает под категорию добросовестного использования.

Доктрина добросовестного использования (fair use) основана на параграфе 107 Закона США об авторском праве (US Copyright Act of 1976), который устанавливает четыре критерия для оценки возможности использования объектов авторского права без получения согласия владельца. Е. В. Балобанова отмечает, что к этим критериям относятся: 1) цель и характер использования (коммерческое или некоммерческое), 2) характер произведения (особенности его создания и т. д.), 3) количество и суть заимствования, 4) возможное воздействие на потенциальный рынок. Широкая судебная практика применения этих критериев не только уточнила содержание доктрины, но и значительно расширила её границы, что может привести к размытию ориентиров и усложнению использования этих критериев.

В контексте первого критерия важно определить, имеет ли несогласованное использование коммерческий или некоммерческий (например, образовательный) характер. В общем, считается, что приоритет в признании добросовестным отдается некоммерческому использованию, однако А. С. Ворожевич отмечает, что «американские суды также выработали правило: некоммерческое использование само по себе не гарантирует добросовестность, хотя и увеличивает вероятность её признания». О. В. Луткова добавляет, что «в американской практике устоялась позиция, что добросовестным считается преобразую-

щее, творческое использование произведения, независимо от его коммерческого или некоммерческого контекста, так как многие вторичные работы создаются с целью извлечения хоть какой-то коммерческой выгоды». В итоге, Л. В. Сагдеева заключает, что использование произведения без согласия правообладателя считается добросовестным, если оно «способствует развитию науки и полезных ремесел», что отличается от получения официальной лицензии на использование авторских прав [8].

Второй критерий касается характера первичного авторского произведения, и, как отмечает П. С. Могилевский, суд будет учитывать, имело ли это произведение творческий характер.

Третий критерий связан с объемом и значимостью заимствований, которые суд оценивает в каждом конкретном случае. А. С. Ворожевич подчеркивает, что при оценке этого критерия учитываются количественные показатели использования, такие как количество копий и объем заимствованного материала.

Четвертый критерий рассматривает, как анализируемое использование влияет на рыночную стоимость первичного произведения, что подразумевает оценку экономического ущерба, понесенного правообладателем.

А. С. Ворожевич отмечает, что американские суды оценивают критерии добросовестного использования с учетом специфики каждого отдельного случая, что приводит к непредсказуемости судебных решений в таких делах. В связи с этим предсказать итоговое решение по делу *Andersen v. Stability AI Ltd.* было бы сложно. В конце 2023 года стало известно о частичном отклонении иска, что объяснялось тем, что некоторые спорные изображения не были зарегистрированы в Бюро авторских прав США, поэтому их нельзя рассматривать как объекты авторского права.

Важно также упомянуть, что в Европейском Союзе действует Директива 2001/29/ЕС, принятой 22 мая 2001 года, которая гармонизирует определенные аспекты авторского права и смежных прав. Эта директива предоставляет государствам-членам свободу выбора и адаптации ограничений в соответствии с их традициями. В статье 5(5) Директивы закреплён трехступенчатый тест для определения границ свободного использования произведений, который устанавливает, что исключения и ограничения могут применяться только в определенных случаях и не должны вредить нормальному использованию произведений или ущемлять законные интересы авторов.

В российском праве существуют законные способы использования объектов интеллектуальных прав без разрешения правообладателя, кото-

рые закреплены в четвертой части Гражданского кодекса РФ. Для обозначения такого использования применяются термины «свободное использование результатов интеллектуальной деятельности» (ст. 1273-1279, 1306) и «действия, не являющиеся нарушением исключительного права» (ст. 1335.1, 1359 и др.).

Это свидетельствует о том, что российское законодательство придерживается иного подхода к допустимости использования авторских произведений без согласия авторов по сравнению с концепцией *fair use*. В статье 1270 Гражданского кодекса РФ прописано, что произведение может использоваться любыми способами, с приведением примерного перечня этих способов. Однако следует отметить, что использование сосредоточено на форме произведения, а не на его содержании. Например, использование произведения ИИ может рассматриваться как воспроизведение (копирование на другой носитель), публичное исполнение, переработка или предоставление доступа к произведению через искусственный интеллект.

В таких обстоятельствах трудно считать, что изображение использовалось, если применялось только его текстовое описание. Как справедливо замечает А. А. Шутыков, понятие «использование» имеет уникальный и специфический характер для каждого охраняемого объекта [10].

А. С. Васильева предлагает второй подход к решению проблемы: дополнить нормы Гражданского кодекса РФ положением о свободном использовании авторских произведений для автоматизированного анализа с целью разработки и обучения искусственного интеллекта. Это подразумевает воспроизведение и переработку объектов исключительно для указанных целей, с обязательством удалить копии, когда такая необходимость отпадает. Она подчеркивает, что реализация этого предложения требует закрепления в российском законодательстве, в отличие от зарубежных стран, таких как США, где подобные нормы не обязательны.

Третий подход заключается в предложении заключать договоры о передаче исключительного права для машинного обучения. А. С. Васильева отмечает преимущества этого метода, такие как прозрачность использования объектов интеллектуальной собственности и возможность для правообладателей получить материальную выгоду. Однако есть и недостатки, включая значительные финансовые затраты на заключение таких договоров.

В качестве четвертого варианта решения проблемы А. С. Васильева предлагает заключение лицензионного договора с правообладате-

лем. Основное преимущество данного подхода заключается в создании баланса интересов между правообладателем и разработчиком. По ее словам, цель такого договора — использование охраняемого объекта для тренировки искусственного интеллекта, что включает воспроизведение и переработку материалов в процессе генерации контента.

На основе изложенного можно сделать несколько выводов:

1. Использование обучающих данных для машинного обучения вызывает множество правовых вопросов, среди которых особенно остро стоит проблема использования авторских произведений без согласия их авторов или правообладателей.
2. Для соблюдения баланса общественных интересов и интересов авторов целесообразно предусмотреть две модели использования таких произведений в контексте машинного обучения.

Первый вариант предполагает разрешение свободного использования авторских произведений для обучения искусственного интеллекта на условиях, схожих с использованием в образовательных целях. Это значит, что такое использование возможно безвозмездно и без согласия правообладателя, при условии, что пользователи могут использовать соответствующие информационные системы без платы.

Второй вариант заключается в том, что при планируемом коммерческом использовании ИИ необходимо получать согласие правообладателя на использование его произведений в качестве обучающих данных для машинного обучения, основываясь на лицензионном или ином договоре.

Список литературы:

- [1] Kühl N., Goutier M., Hirt R., Satzger G. Machine Learning in Artificial Intelligence: Towards a Common Understanding. // Cornell University. 27.03.2020. // URL: <https://arxiv.org/ftp/arxiv/papers/2004/2004.04686.pdf>; Wakefield K. A guide to the types of machine learning algorithms and their applications. – SAS UK // URL: https://www.sas.com/en_gb/insights/articles/analytics/machine-learning-algorithms.html
- [2] Artificial Intelligence in Society, OECD Publishing, Paris, 2019. // URL: <https://www.oecd-ilibrary.org/docserver/eedfee77-en.pdf>.
- [3] Наумов В. Б., Тютюк Е. В. Правовые проблемы машинного обучения // Образование и право. – 2020. – № 6
- [4] Andersen v. Stability AI Ltd. (3:23-cv-00201) URL: <https://www.courtlistener.com/docket/66732129/andersen-v-stability-ai-ltd/>
- [5] Балобанова Е. В. К вопросу о применимости доктрины добросовестного использования в континентальном праве // Журнал Суда по интеллектуальным правам. – 2017. – № 18
- [6] Ворожевич А. С. Fair use и доктрина «расширяющего полезность использования» // URL: <https://legal-support.ru/information/blog/ip-daidzhest/fair-use-i-doktrina-rasshiraushhego-poleznost-ispolzovaniya/>
- [7] Луткова О. В. Доктрина добросовестного использования произведений в современном авторском праве США // Право. Журнал Высшей школы экономики. – 2016. – № 2
- [8] Сагдеева Л. В. Свободное использование как ограничение исключительных прав // Юридические исследования. – 2017. – № 9
- [9] Могилевский П. С. Доктрина добросовестного использования в США и свободное использование произведений в России: сравнительный анализ // Гражданское право. – 2022. – № 5
- [10] Шутьков А. А., Анищенко А. Н. Будущее искусственного интеллекта, нейросетей и цифровых технологий в АПК // Экономика и социум: современные модели развития. – 2019. – Т. 9. – № 4 (26). – С. 508–522
- [11] Доклад НИУ ВШЭ «Проблема машинного творчества в системе права: регулирование создания и использования результатов интеллектуальной деятельности с применением искусственного интеллекта, зарубежный опыт и российские перспективы» / Рук. авт. колл. В. О. Калятин. – М.: Издательский дом Высшей школы экономики. 2021.
- [12] Васильева А. С. Использование объектов авторских прав для обучения искусственного интеллекта // Журнал Суда по интеллектуальным правам. – 2024. – № 4 (46). Декабрь.
- [13] На сегодняшний день различают три основных метода МО, которые подробно описаны в технической литературе (см., например, материалы IBM: What is machine learning (ML)? // URL: <https://www.ibm.com/topics/machine-learning>). Во-первых, контролируемое МО, которое использует маркированные наборы данных для обучения алгоритмов классификации данных или точного прогнозирования результатов: по мере того, как входные данные подаются в модель, эта модель корректирует их поступление, пока они не будут соответствующим образом подобраны, что позволяет избежать как переобучения, так и недообучения модели. Во-вторых, неконтролируемое МО, которое использует алгоритмы машинного обучения для анализа и кластеризации немаркированных наборов данных – эти алгоритмы обнаруживают скрытые закономерности или группировки данных без вмешательства

человека. В-третьих, полуконтролируемое МО, которое иногда именуют как «обучение с подкреплением» и оценивают как золотую середину между контролируемым и неконтролируемым МО: такое обучение предполагает использование меньшего маркированного набора данных для классификации и извлечения признаков из большего, немаркированного набора данных (Подробнее см. Julianna Delua. Supervised versus unsupervised learning: What's the difference? // URL: <https://www.ibm.com/think/topics/supervised-vs-unsupervised-learning>).

References:

- [1] Kühl N., Goutier M., Hirt R., Satzger G. Machine Learning in Artificial Intelligence: Towards a Common Understanding. // Cornell University. 27.03.2020. // URL: <https://arxiv.org/ftp/arxiv/papers/2004/2004.04686.pdf>; Wakefield K. A guide to the types of machine learning algorithms and their applications. – SAS UK // URL: https://www.sas.com/en_gb/insights/articles/analytics/machine-learning-algorithms.html
- [2] Artificial Intelligence in Society, OECD Publishing, Paris, 2019. // URL: <https://www.oecd-ilibrary.org/docserver/eedfee77-en.pdf>.
- [3] Naumov V. B., Tytyuk E. V. Legal problems of machine learning//Education and law. – 2020. – № 6
- [4] Andersen v. Stability AI Ltd. (3:23-cv-00201) URL: <https://www.courtlistener.com/docket/66732129/andersen-v-stability-ai-ltd/>
- [5] Balobanova E.V. On the applicability of the doctrine of fair use in continental law//Journal of the Intellectual Property Court. – 2017. – № 18
- [6] Ворожевич А. С. Fair use и доктрина «расширяющего полезность использования» // URL: <https://legal-support.ru/information/blog/ip-daizhest/fair-use-i-doktrina-rasshiraushhego-poleznost-ispolzovaniya/>
- [7] Lutkova O. V. Doctrine of fair use of works in modern US copyright//Law. Journal of the Higher School of Economics. – 2016. – № 2
- [8] Sagdeeva L.V. Free use as a restriction of exclusive rights//Legal research. – 2017. – № 9
- [9] Mogilevsky P.S. The doctrine of fair use in the USA and the free use of works in Russia: comparative analysis//Civil law. – 2022. – № 5
- [10] Shutkov A. A., Anishchenko A. N. The future of artificial intelligence, neural networks and digital technologies in the agro-industrial complex//Economics and society: modern development models. – 2019. - VOL. 9. – № 4 (26). - S. 508-522
- [11] HSE report “The problem of machine creativity in the legal system: regulation of the creation and use of the results of intellectual activity using artificial intelligence, foreign experience and Russian prospects “/Ruk. auto call V.O. Kalyatin. - M.: Publishing House of the Higher School of Economics. 2021.
- [12] Vasilyeva A. S. Use of copyright objects for training artificial intelligence//Journal of the Court for Intellectual Rights. – 2024. – № 4 (46). Dec.
- [13] To date, there are three main MO methods that are described in detail in the technical literature (see, for example, IBM materials: What is machine learning (ML) ?//URL: <https://www.ibm.com/topics/machine-learning>). First, a controlled MO that uses labeled datasets to train data classification algorithms or accurately predict outcomes: as input data is fed into the model, this model adjusts its arrival until it is matched appropriately, thus avoiding both over- and under-training of the model. Second, uncontrolled MO, which uses machine learning algorithms to analyze and cluster unmarked datasets - these algorithms detect hidden patterns or groupings of data without human intervention. Third, semi-controlled MO, sometimes referred to as “reinforcement learning” and assessed as the sweet spot between controlled and uncontrolled MO: such learning involves using a smaller labeled dataset to classify and extract features from a larger, unmarked dataset (For details, see Julianna Delua. Supervised versus unsupervised learning: What's the difference? // URL: <https://www.ibm.com/think/topics/supervised-vs-unsupervised-learning>).

