

УДК 347.78

ИСПОЛЬЗОВАНИЕ ОБЪЕКТОВ АВТОРСКОГО ПРАВА В МАШИННОМ ОБУЧЕНИИ: В ПОИСКАХ БАЛАНСА ИНТЕРЕСОВ ПРАВООБЛАДАТЕЛЕЙ И ОБЩЕСТВА

А. Н. Артемова

Институт философии и права СО РАН (г. Новосибирск)

artemova-an-1991@yandex.ru

Аннотация. Внедрение искусственного интеллекта открывает большие возможности, но вместе с тем ставит новые задачи перед законодателями и правоприменителями. Одной из таких задач является поиск баланса между интересами правообладателей произведений, используемых в машинном обучении, и общества, заинтересованного в развитии технологии искусственного интеллекта. В статье автор анализирует состояние действующего законодательства в сфере авторского права и правоприменительную практику. Предметом исследования выступает судебное дело о нарушении авторских прав по иску группы писателей (Пол Трэмблей и др.) к компании-разработчику искусственного интеллекта OpenAI, которое находится в производстве Окружного суда США (Северный округ Калифорнии): автор оценивает использование произведений в машинном обучении на предмет соответствия условиям применения доктрины добросовестного использования (fair use). В заключении автор приходит к выводу о необходимости дополнения случаев свободного использования произведений посредством использования в системах машинного обучения.

Ключевые слова: авторское право; искусственный интеллект; машинное обучение; большие языковые модели; набор данных; свободное использование произведений; доктрина добросовестного использования произведений.

Для цитирования: Артемова, А. Н. (2024). Использование объектов авторского права в машинном обучении: в поисках баланса интересов правообладателей и общества. *Respublica Literaria*. Т. 5. № 3. С. 184-194. DOI: 10.47850/RL.2024.5.3.184-194

USE OF COPYRIGHT WORKS IN MACHINE LEARNING: SEARCH FOR A BALANCE OF INTERESTS OF COPYRIGHT HOLDERS AND SOCIETY

A. N. Artemova

Institute of Philosophy and Law SB RAS (Novosibirsk)

artemova-an-1991@yandex.ru

Abstract. The introduction of artificial intelligence opens up great opportunities, but at the same time poses new challenges for legislators and courts. One of these challenges is to find a balance between the interests of copyright holders of works used in machine learning and society interested in the development of artificial intelligence technology. In the article, the author analyzes the state of current copyright legislation and law enforcement practice. The subject of the study is a copyright infringement lawsuit filed by a group of writers (Paul Tremblay et al.) against the artificial intelligence developer OpenAI, which is being processed by the US District Court (Northern District of California): the author evaluates the use of works in machine learning for compliance with the conditions for applying the fair use doctrine. The author comes to the conclusion about the need to supplement cases of free use of works through use in machine learning systems.

Keywords: copyright; artificial intelligence; machine learning; large language models; dataset; free use of works; fair use doctrine.

For citation: Artemova, A. N. (2024). Use of Copyright Works in Machine Learning: Search for a Balance of Interests of Copyright Holders and Society. *Respublica Literaria*. Vol. 5. No. 3. Pp. 184-194. DOI: 10.47850/RL.2024.5.3.184-194

Искусственный интеллект (далее – ИИ) сегодня является приоритетным направлением развития экономики Российской Федерации. Большая ставка делается на генеративный искусственный интеллект, внедрение которого обещает настоящий прорыв в экономике. Как отмечается в Национальной стратегии развития искусственного интеллекта на период до 2030 г., внедрение больших генеративных моделей может привести к повышению производительности труда, увеличению валового внутреннего продукта на 1–2 % ежегодно, повышению оплаты труда специалистов во всех отраслях за счет увеличения объема выпуска продукции и улучшения ее качества¹.

Вместе с тем внедрение ИИ требует обеспечения комплексного нормативно-правового регулирования. Одним из ключевых вопросов является определение условий допустимого использования результатов интеллектуальной деятельности третьих лиц в машинном обучении². Однако до настоящего времени такие условия законодательно не определены.

Машинное обучение (Machine Learning) как процесс создания алгоритмов, самостоятельно обучающихся благодаря опыту, основано на использовании наборов данных (Dataset). Такие наборы данных могут включать в себя и произведения, охраняемые авторским правом. Так, книги выступают ключевым компонентом наборов данных для обучения больших языковых моделей (Large Language Models) – программного обеспечения на основе искусственного интеллекта, предназначенного для анализа и генерации естественного языка, поскольку служат источником качественных литературных произведений.

Представляя в 2018 г. свою разработку – большую языковую модель GPT-1 (Generated Pretrained Transformer), корпорация OpenAI сообщила, что обучала GPT-1 на BookCorpus – коллекции из более 7000 уникальных неопубликованных книг самых разных жанров. Как пояснила OpenAI, ценность книг состоит в том, что они содержат длинные отрезки непрерывного текста, что позволяет модели учиться обуславливать большие объемы генерируемого текста [Radford et al., 2018, p. 4].

Набор данных BookCorpus был создан в 2015 г. группой исследователей из Университета Торонто и Массачусетского технологического института в целях обучения языковых моделей и состоит из неопубликованных романов, размещенных в Интернете

¹ Указ Президента Российской Федерации от 10.10.2019 № 490 «О развитии искусственного интеллекта в Российской Федерации». Национальная стратегия развития искусственного интеллекта на период до 2030 года (в редакции Указа Президента Российской Федерации от 15.02.2024 № 124). П. 175. [Электронный ресурс]. *Президент России*. URL: <http://www.kremlin.ru/acts/bank/44731/page/2> (дата обращения: 18.07.2024).

² Распоряжение Правительства Российской Федерации от 19 августа 2020 г. № 2129-р «Концепция развития регулирования отношений в сфере технологий искусственного интеллекта и робототехники на период до 2024 г.». [Электронный ресурс]. *Официальный интернет-портал правовой информации*. URL: <http://government.ru/docs/all/129505/> (дата обращения: 18.07.2024).

и доступных читателям бесплатно [Zhu et al., 2015, p. 21]. Несмотря на то, что входящие в состав BookCorpus произведения охраняются авторским правом, согласие авторов на использование их произведений в обучающем наборе данных получено не было.

В 2020 г., представляя общественности большую языковую модель GPT-3, OpenAI сообщила, что 15 % обучающих данных было получено из двух интернет-наборов книг: «Книги-1» и «Книги-2». Какие именно книги входят в состав данных обучающих наборов, OpenAI не раскрыла [Brown et al., 2020, p. 8]. Однако их количество существенно превышает количество книг, входящих в BookCorpus.

Предположительно, набор данных «Книги-1» был сформирован из цифровой коллекции произведений, находящихся в общественном достоянии «Project Gutenberg», а набор данных «Книги-2» содержит данные из нелегальных «пиратских библиотек», таких как Library Genesis (также известной как LibGen), Z-Library (также известной как Vok), Sci-Hub и Bibliotik³.

Представляя GPT-4 в марте 2023 г., OpenAI уже не сообщала информацию об обучающих наборах данных, ссылаясь на конкурентную среду и соображения безопасности⁴.

Таким образом, как верно подметил Н. А. Назаров, «на данный момент объекты машинного творчества создаются по принципу “пока не поймали”, основанном на том, что незаконное использование определенного произведения в коллекции данных в машинном творчестве нельзя обнаружить» [Назаров, 2022, с. 109]. Однако появление судебной практики по незаконному использованию произведений в машинном обучении лишь вопрос времени. В США такая практика уже на этапе формирования. На данный момент на рассмотрении судов США находится несколько дел о нарушении авторских прав посредством использования произведений литературы в машинном обучении без согласия правообладателей (Tremblay et al. v. OpenAI, Inc. et al.; Silverman et al. v. OpenAI, Inc. et al.; Authors Guild et al. v. OpenAI Inc. et al. и др.) В частности, 28 июня 2023 г. группой писателей (Пол Трэмблей и др.) был подан иск к корпорации OpenAI о запрете использования их произведений, взыскании компенсации и убытков, причиненных незаконным использованием. Иск заявлен как коллективный в защиту прав всех лиц на территории США, обладающих авторскими правами на произведения, использованные OpenAI для машинного обучения без согласия правообладателей.

Как утверждают истцы, они являются авторами книг, использованных для обучения ChatGPT (программное обеспечение, основанное на языковых моделях GPT-3,5 и GPT-4, созданное, поддерживаемое и реализуемое корпорацией OpenAI). ChatGPT способен взаимодействовать с пользователем в разговорном стиле: пользователь вводит текстовый запрос, а ChatGPT генерирует ответы, имитирующие человеческий язык. Поскольку результаты ответов ChatGPT зависят от наборов данных, использованных для обучения, тот факт, что ChatGPT выдает краткое изложение содержания произведений истцов в ответ на соответствующий запрос, свидетельствует о том, что языковая модель была обучена

³ Paul Tremblay, et al. v. OpenAI, Inc., et al. Class Action Complaint. (2023). [Online]. Available at: https://storage.courtlistener.com/recap/gov.uscourts.cand.414822/gov.uscourts.cand.414822.1.0_1.pdf (Accessed: 18 July 2024).

⁴ GPT-4 Technical Report. OpenAI. (2023). [Online]. Available at: <https://cdn.openai.com/papers/gpt-4.pdf> (Accessed: 18 July 2024).

на произведениях истцов, охраняемых авторским правом. При этом OpenAI не обращалась к истцам за получением согласия на такое использование. В обоснование исковых требований истцы указали на следующие нарушения OpenAI:

1. Прямое нарушение авторских прав в результате обучения больших языковых моделей на произведениях, правообладателями которых являются истцы. В соответствии со статьей 106 Закона об авторском праве США копирование произведения, а также осуществление иных правомочий без согласия правообладателя запрещено.

2. Опосредованное нарушение авторских прав. Поскольку выходные данные основаны на информации, извлеченной из работ истцов, каждый ответ, генерируемый ChatGPT, представляет собой производное произведение, созданное в отсутствие согласия правообладателя исходного произведения, что запрещено законом (статья 106 Закона об авторском праве США).

3. Удаление информации об управлении авторскими правами. В книгах истцов содержалась информация об управлении авторскими правами (уведомление об авторских правах, информация об авторе и правообладателе, условия использования). В нарушение статьи 1202(b)(1) Закона США о защите авторских прав в цифровую эпоху OpenAI намеренно удалила информацию об авторах. Распространяя производные произведения, созданные на основе произведений авторов, без указания информации о них, OpenAI нарушила статью 1202(b)(3) указанного закона.

4. Недобросовестная конкуренция. По мнению истцов, действия OpenAI нечестны, неэтичны, аморальны, беспринципны, поскольку корпорация необоснованно извлекает прибыль и приписывает себе разработку коммерческого продукта, основанного на чужих произведениях, охраняемых авторским правом. Кроме того, удаляя информацию об авторах исходных произведений, OpenAI обманывает пользователей. Таким образом, OpenAI нарушила положения Делового и профессионального кодекса штата Калифорния.

5. Халатность. По мнению истцов, на OpenAI лежит обязанность проявлять заботливость в отношении информации, которая находится в их распоряжении, воздерживаясь от причинения вреда истцам. Эта обязанность вытекает из статьи 1714 Гражданского кодекса штата Калифорния, предписывающей действовать добросовестно по отношению к другим лицам.

6. Неосновательное обогащение. Истцы вложили значительное время и энергию в создание произведений. OpenAI извлекает выгоду из использования произведений истцов в то время, как истцы не получили денежного вознаграждения за использование их произведений компанией OpenAI⁵.

Большая часть исковых требований была отклонена судом 12 февраля 2024 г. по ходатайству ответчика. В соответствии с Федеральными правилами гражданского судопроизводства США (далее – Правила) иск должен содержать краткое и ясное изложение, демонстрирующее, что истец имеет право на судебную защиту (статья 8 (а) Правил). В случае несоблюдения данного требования ответчик вправе подать ходатайство об отклонении иска (статья 12 (b)(6) Правил).

⁵ Paul Tremblay, et al. v. OpenAI, Inc., et al. Class Action Complaint. (2023). [Online]. Available at: https://storage.courtlistener.com/recap/gov.uscourts.cand.414822/gov.uscourts.cand.414822.1.0_1.pdf (Accessed: 18 July 2024).

Так, суд отклонил исковые требования, основанные на опосредованном нарушении авторских прав, удалении информации об управлении авторскими правами, халатности и неосновательном обогащении. Как указал в определении суд, утверждения о том, что любой результат деятельности GhatGPT представляет собой производное произведение, нарушающее авторское право, недостаточно для установления нарушения авторских прав, необходимо показать существенное сходство между оригиналом и производным произведением, чего истцы не продемонстрировали.

Что касается предполагаемого удаления OpenAI информации об управлении авторскими правами, то в соответствии со статьей 1202(b)(1) Закона США о защите авторских прав в цифровую эпоху для такого заявления требуются факты, демонстрирующие, что предполагаемый нарушитель знал или имел разумные основания знать о том, что удаление информации об авторских правах спровоцирует, сделает возможным, облегчит или скроет нарушение авторских прав. Утверждения истцов голословны, не подкреплены доказательствами. Напротив, тот факт, что в приведенных истцами выдержках из ответов ChatGPT упоминаются имена истцов, говорит о том, что информация об авторах не была удалена из обучающих наборов данных. Что касается ответственности за распространение произведений или их копий без информации об авторских правах, то в силу статьи 1202(b)(3) указанного закона она наступает за распространение оригинальных произведений. Однако истцы не утверждали, что OpenAI распространяла их книги или копии их книг. Вместо этого они заявили, что каждый результат ChatGPT является производной работой, нарушающей авторские права, не представив доказательств того, что такие результаты являются копиями их книг.

Также, по мнению суда, истцам не удалось доказать наличия между ними и ответчиком особых фидуциарных правоотношений и обязанности ответчика охранять информацию, содержащуюся в произведениях истцов, а также получения OpenAI неосновательного обогащения от использования произведений истцов.

Наконец, оценивая действия ответчиков на предмет недобросовестной конкуренции, суд указал, что для признания определенной деловой практики недобросовестной конкуренцией необходимо, чтобы она была «незаконной», «несправедливой» или «мошеннической». Отклоняя требования, основанные на нарушении Закона об авторском праве в цифровую эпоху и мошенничестве, суд тем не менее посчитал заслуживающим внимания аргумент истцов о несправедливых действиях ответчика, выразившихся в том, что OpenAI использовала произведения истцов, охраняемые авторским правом, для тренировки своих языковых моделей с целью получения коммерческой выгоды. Суд пришел к выводу, что поведение OpenAI может представлять собой недобросовестную практику и оставил ходатайство ответчиков об отклонении иска в этой части без удовлетворения⁶.

⁶Paul Tremblay, et al. v. OpenAI, Inc., et al. United States District Court. Northern District Of California. Order Granting In Part And Denying in Part The Motions To Dismiss. (2024). [Online]. Available at: https://storage.courtlistener.com/recap/gov.uscourts.cand.414822/gov.uscourts.cand.414822.104.0_1.pdf (Accessed: 18 July 2024).

Истцам была предоставлена возможность внести изменения в срок до 13 марта 2024 г. В уточненном исковом заявлении истцы, ссылаясь на прямое нарушение авторских прав со стороны OpenAI в результате использования их произведений для обучения больших языковых моделей, а также на недобросовестную конкуренцию со стороны OpenAI, выразившуюся в использовании произведений истцов для разработки коммерческого продукта и получения собственной коммерческой выгоды, просили суд признать факт умышленного нарушения исключительных прав истцов со стороны ответчиков, взыскать убытки и компенсацию, предусмотренную статьей 504 Свода федеральных законов США, а также неустойку в виде процентов и судебные расходы. Кроме того истцы просили рассмотреть заявленные требования судом присяжных в соответствии со статьей 38(b) Федеральных правил гражданского судопроизводства. На данный момент рассмотрение дела не завершено.

Представляется, что позиция ответчиков относительно прямого нарушения авторских прав будет основываться на доктрине добросовестного использования произведений (fair use), в соответствии с которой допускается использование произведений без согласия правообладателя и без выплаты ему вознаграждения, если это способствует прогрессу науки и искусства. Сформировавшись в рамках системы общего права (Common Law) доктрина впоследствии нашла закрепление в статье 107 Закона об авторском праве США 1976 г.

Для оценки добросовестности использования произведения суды анализируют четыре фактора: 1) цель и характер использования; 2) характер используемого произведения; 3) объем и существенность использованной части произведения; 4) влияние на рынок.

Все факторы оцениваются в совокупности, но ключевое значение имеют первый и четвертый факторы. Чтобы использование произведения было признано добросовестным, оно должно носить преобразующий, творческий характер, быть нацелено на создание новых объектов авторского права. Новое произведение должно обогащать общество, стимулировать развитие инноваций, вносить свой вклад в культуру, а не подменять собой оригинал, не конкурировать с ним, нанося ущерб его коммерциализации.

Так, в знаменитом деле *The Authors Guild v. Google Inc.* оцифровка книг без согласия правообладателей для создания сервиса по поиску книг Google Books была признана добросовестным использованием.

Фабула дела была такова: в 2004 г. корпорация Google запустила новый сервис по поиску книг – Google Books, заключив соглашения с рядом крупнейших научных библиотек мира на оцифровку книг. С 2004 г. Google отсканировала и сделала машиночитаемыми свыше 20 млн. книг, включая книги как находящиеся в общественном достоянии, так и охраняемые авторским правом. Пользователям сервиса была предоставлена возможность осуществлять поиск книг по фразам и ключевым словам. В информацию о книге, предоставляемую сервисом, включалось краткое описание, ключевые слова, ссылки для покупки книги в Интернете и библиотеки, где можно ознакомиться с книгой. В ряде случаев предоставлялась возможность ограниченного просмотра текста (ознакомительный фрагмент).

Ряд авторов оцифрованных работ, однако, посчитали, что проект Google Books нарушает их права, поскольку не было испрашено их согласие на такое использование, и в 2005 г. обратились в суд с иском о взыскании компенсации за нарушение авторских прав.

В обоснование своей позиции Google ссылалась на доктрину добросовестного использования.

Оценив четыре условия применения доктрины, суд пришел к выводу, что использование произведений компанией Google носит преобразующий характер. Во-первых, благодаря функции поиска книг по ключевым словам / фразам Google Books является уникальным инструментом для интеллектуального анализа текста и данных. Во-вторых, оцифровка книг позволила расширить доступ к книгам для лиц с ограниченными возможностями. В-третьих, оцифровка дала жизнь старым обветшалым изданиям. Наконец, Google Books позволяет ознакомиться с информацией о книге и способами ее приобретения, тем самым Google привлекает новую аудиторию авторам и увеличивает продажи книг в интересах правообладателей. Важным моментом является то, что отображение материалов, охраняемых авторским правом, надлежащим образом ограничено, пользователям доступен лишь ознакомительный фрагмент. Таким образом, суд признал, что Google Books не заменяет и не вытесняет книги. Вместо этого проект приносит значительную общественную пользу, учитывает права авторов и увеличивает продажи книг в интересах правообладателей. В результате Окружной суд США (Южный округ Нью-Йорка) отказал в удовлетворении иска⁷. В 2015 г. Апелляционный суд США оставил в силе данное решение⁸.

Как отмечает О. В. Луткова, ни один фактор не предписан императивно, поэтому учитывая уникальность и специфику каждого конкретного спора, суды могут принимать во внимание и другие факторы, не указанные в законе, в частности, недобросовестное поведение стороны, общественный интерес, «право автора контролировать первое появление своего произведения на публике» [Луткова, 2016, с. 193, 196].

Таким образом, поскольку наличие у судей дискреционных полномочий по анализу данных факторов в конкретном деле делает применение доктрины судами противоречивым и непредсказуемым, прогнозирование исхода дела *Tremblay v. OpenAI* является крайне сложной задачей. Представляется, что использование произведений в машинном обучении удовлетворяет первому условию, поскольку является преобразующим. Большая языковая модель, несомненно, способствует развитию науки и искусства, не подменяя собой исходные произведения, а предоставляя возможность создавать новые на основе запроса пользователя.

Тот факт, что для обучения были использованы значительные по объему фрагменты художественных произведений, творческий вклад авторов которых традиционно оценивается высоко, играет не в пользу OpenAI. Однако второй и третий факторы сами по себе не являются определяющими, и оцениваются судами в совокупности с другими.

⁷*Authors Guild v. Google, Inc.* United States District Court. Southern District Of New York. Opinion Granting Summary Judgment of Fair Use. (2013). [Online]. Available at: <https://www.eff.org/document/opinion-granting-summary-judgment-fair-use> (Accessed: 30 July 2024).

⁸*Authors Guild v. Google, Inc.* United States Court of Appeals For The Second Circuit. Second Circuit Opinion (2015). [Online]. Available at: <https://www.eff.org/document/ruling-appeals-court> (Accessed: 30 July 2024).

Скрупулезного изучения требует четвертый фактор, который, по нашему мнению, является решающим. С одной стороны, формулирование с помощью ChatGPT текста по запросу пользователя не может служить заменой книг, т. е. продукт OpenAI не является конкурентом произведениям истцов. Книги и большие языковые модели служат разным целям и удовлетворяют разные потребности потребителей.

С другой стороны, даже если характер использования произведения в высшей степени преобразующий, такое использование может нанести ущерб ценности использованного произведения, охраняемого авторским правом, если приведет к широкому раскрытию значительной части данного произведения и тем самым сделает доступным конкурирующий аналог. Как утверждают истцы, ChatGPT выдает на основе запроса обобщение их книг. Такое краткое содержание уже может выступать прямым конкурентом оригинальному произведению.

Также не в пользу OpenAI играет и тот факт, что компания не раскрывает произведения, использованные для обучения, и их авторов, что может быть расценено как недобросовестное поведение.

Таким образом, суду предстоит исследовать представленные истцами доказательства (генерируемый ChatGPT текст, содержащий краткое содержание произведений истцов), чтобы сделать вывод о влиянии, которое продукт, созданный OpenAI, способен оказывать на рынок и ценность книг истцов. От того, какое решение будет в конечном итоге принято по делу, будет зависеть правоприменительная практика по данной категории дел, формирование доктрины добросовестного использования и дальнейшее развитие индустрии ИИ.

На данный момент является очевидным тот факт, что машинное обучение – это та сфера, где необходимо найти баланс интересов правообладателей и общества, поскольку от развития ИИ зависит, без преувеличения, благосостояние всего общества, и это тот случай, когда ограничение исключительных прав может быть оправдано необходимостью развития технологии ИИ, которое невозможно без использования произведений науки, литературы и искусства в машинном обучении. Установление понятных правил необходимо как для разработчиков ИИ, так и для правообладателей произведений. В доктрине справедливо указывается на то, что отсутствие правового регулирования машинного обучения будет препятствовать развитию ИИ, поскольку легализация контента, используемого для машинного обучения, будет сопряжена с дополнительными расходами, а ограничение контента доступными источниками будет вести к снижению качества создаваемого ИИ [Калятин, 2021, с. 24].

Если в странах общего права баланс интересов правообладателей и общества достигается посредством применения доктрины добросовестного использования произведений, то в странах романо-германской правовой системы данная цель достигается путем перечисления в законе случаев свободного использования произведений (информационные, научные, учебные, культурные цели и др.) Такой подход именуется перечневым: законодатель устанавливает в законе точный перечень случаев свободного использования, использование произведения в случаях, не поименованных в законе, не допускается. Отечественным законодательством также установлен перечень случаев

свободного использования (ст. 1273-1280 ГК РФ), однако ни один из случаев свободного использования произведений, установленных ГК РФ, не предусматривает возможности использования произведений в машинном обучении.

С точки зрения действующего российского законодательства к способам законного использования произведений в машинном обучении можно отнести: 1) заключение договора с правообладателем; 2) использование произведений с открытой лицензией (с учетом условий, на которых правообладатель разрешил использование своего произведения); 3) использование произведений, перешедших в общественное достояние.

К сожалению, ни один из указанных способов не является адекватным для целей машинного обучения. Первый способ требует значительных временных и финансовых затрат. Второй и третий предполагают доступ к ограниченному количеству произведений, что существенно снижает эффективность ИИ, ведь чем больший набор актуальных данных был использован для обучения, тем точнее и ценнее будет результат машинного обучения.

В этой связи представляет интерес подход к решению проблемы, реализованный в ЕС в принятом в 2024 г. Регламенте Европейского Союза «Об искусственном интеллекте» (EU Artificial Intelligence Act). Так, поставщики моделей ИИ общего назначения обязаны составить и сделать общедоступной информацию о контенте, используемом для обучения модели ИИ общего назначения (п. 107 Регламента), а также соблюдать законодательство ЕС об авторском праве, в частности, положения статьи 4 (3) Директивы (ЕС) 2019/790 (п. 105 Регламента)⁹. Данной Директивой предусмотрена возможность правообладателя посредством прямо выраженного запрета предотвратить использование его произведения для интеллектуального анализа текста и данных за исключением случаев, когда это делается в целях научных исследований. При этом интеллектуальный анализ текста и данных (Text and Data Mining) служит инструментом формирования набора данных, необходимого для машинного обучения [Афанасьева, Фурман, 2023, с. 1645-1646]. Следовательно, если такой запрет был прямо выражен, разработчикам моделей ИИ общего назначения необходимо получить разрешение от правообладателя, либо воздержаться от использования произведения. Данное решение (возможность правообладателя зарезервировать право на произведение, использование которого законодателем разрешено) критикуется учеными за создание конструкций «неопределенной правовой природы» (введение ограничений на ограничения) [Назаров, 2022, с. 108].

Более разумным выглядит предложение дополнить случаи свободного использования охраняемых объектов авторских прав посредством использования в машинном обучении. Согласно Дорожной карте развития нормативно-правового регулирования в сфере интеллектуальной собственности соответствующие изменения в Гражданский кодекс РФ планируется внести в 2025 г.¹⁰

⁹ *Artificial Intelligence Act. Regulation (EU) 2024/1689*. [Online]. Available at: <https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX:32024R1689> (Accessed: 18 July 2024).

¹⁰ Распоряжение Правительства РФ от 03.08.2020 № 2027-р (ред. от 29.06.2023) «Об утверждении плана мероприятий (“дорожной карты”) по реализации механизма управления системными изменениями нормативно-правового регулирования предпринимательской деятельности “Трансформация делового климата” “Интеллектуальная собственность”». [Электронный ресурс]. *Консультант плюс*. URL: https://www.consultant.ru/document/cons_doc_LAW_359540/a784b71d972346637c2ccceec51b63163548b642/ (дата обращения: 18.07.2024).

Баланс интересов правообладателей и общества, будучи важным принципом авторского права, нуждается в постоянной корректировке в соответствии с текущим уровнем развития технологий и интересами общества. В этой связи дополнение случаев свободного использования объектов авторского права использованием в системах машинного обучения является логически обоснованным и послужит формированию правовой основы для развития и внедрения технологии искусственного интеллекта в Российской Федерации.

Список литературы/ References

Афанасьева, Е. Н., Фурман, Д. Д. (2023). Правовые аспекты обработки искусственными интеллектуальными системами объектов авторского права в процессе машинного обучения. *Baikal Research Journal*. Т. 14. № 4. С. 1643-1656.

Afanasyeva, E. N., Furman, D. D. (2023). Legal aspects of processing copyright objects by artificial intelligent systems in the process of machine learning. *Baikal Research Journal*. Vol. 14. No. 4. Pp. 1643-1656. (In Russ.)

Калятин, В. О. (2021). *Проблема машинного творчества в системе права: регулирование создания и использования результатов интеллектуальной деятельности с применением искусственного интеллекта, зарубежный опыт и российские перспективы. Доклад НИУ ВШЭ*. М.: Издательский дом Высшей школы экономики. [Электронный ресурс]. URL: <https://www.hse.ru/mirror/pubs/share/482492820.pdf> (дата обращения: 18.07.2024).

Kalyatin, V. O. (2021). *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. Report from the National Research University Higher School of Economics*. Moscow. [Online]. Available at: <https://arxiv.org/pdf/2005.14165> (Accessed: 07 August 2024). (In Russ.)

Луткова, О. В. (2016). Доктрина добросовестного использования произведений в современном авторском праве США. *Право. Журнал Высшей школы экономики*. № 2. С. 186-199.

Lutkova, O. V. (2016). The fair use doctrine in modern US copyright law. *Right. Journal of the Higher School of Economics*. No. 2. Pp. 186-199. (In Russ.)

Назаров, Н. А. (2022). Машинное творчество и право: две части одного целого. *Труды по интеллектуальной собственности*. Т. 43. № 4. С. 101-110.

Nazarov, N. A. (2022). Machine creativity and law: two parts of one whole. *Papers on intellectual property*. Vol. 43. No. 4. Pp. 101-110. (In Russ.)

Brown, T. B., Mann, B., Ryder, N., Subbiah, M., Kaplan, J., Dhariwal, P., Neelakantan, A., Shyam, P., Sastry, G., Askell, A., Agarwal, S., Herbert-Voss, A., Krueger, G., Henighan, T., Child, R., Ramesh, A., Ziegler, D. M., Wu, J., Winter, C., Hesse, C., Chen, M., Sigler, E., Litwin, M., Gray, S.,

Chess, B., Clark, J., Berner, C., McCandlish, S., Radford, A., Sutskever, I., Amodei, D. (2020). *Language Models are Few-Shot Learners*. [Online]. Available at: <https://arxiv.org/pdf/2005.14165> (Accessed: 18 July 2024).

Radford, A., Narasimhan, K., Salimans, T., Sutskever, I. (2018). *Improving Language Understanding by Generative Pre-Training*. [Online]. Available at: https://cdn.openai.com/research-covers/language-unsupervised/language_understanding_paper.pdf (Accessed: 18 July 2024).

Zhu, Y., Kiros, R., Zemel, R., Salakhutdinov, R., Urtasun, R., Torralba, A., Fidler, S. (2015). *Aligning Books and Movies: Towards Story-like Visual Explanations by Watching Movies and Reading Books*. [Online]. Available at: https://www.cv-foundation.org/openaccess/content_iccv_2015/papers/Zhu_Aligning_Books_and_ICCV_2015_paper.pdf (Accessed: 18 July 2024).

Сведения об авторе / Information about the author

Артемова Анастасия Николаевна – кандидат юридических наук, научный сотрудник Института философии и права Сибирского отделения Российской академии наук, г. Новосибирск, ул. Николаева, 8, e-mail: artemova-an-1991@yandex.ru.

Статья поступила в редакцию: 10.08.2024

После доработки: 20.09.2024

Принята к публикации: 04.10.2024

Artemova Anastasiia – Candidate of Juridical Sciences, Researcher of the Institute of Philosophy and Law of the Siberian Branch of the Russian Academy of Sciences, Novosibirsk, Nikolaeva str., 8, e-mail: artemova-an-1991@yandex.ru.

The paper was submitted: 10.08.2024

Received after reworking: 20.09.2024

Accepted for publication: 04.10.2024