

ИСПОЛЬЗОВАНИЕ ОХРАНЯЕМЫХ ПРОИЗВЕДЕНИЙ ПРИ МАШИННОМ ОБУЧЕНИИ ГЕНЕРАТИВНОГО ИСКУССТВЕННОГО ИНТЕЛЛЕКТА

USE OF COPYRIGHTED WORKS IN MACHINE LEARNING FOR GENERATIVE ARTIFICIAL INTELLIGENCE

Мария Викторовна САМАРЦЕВА

Юридическая фирма «Дякин, Горцунян и Партнеры»,
Москва, Российская Федерация,
e-mail: mvsamartseva@yandex.ru,
ORCID: 0009-0007-5628-366X,
Web of Science ID: MUY-2674-2025

Информация об авторе

М.В. Самарцева — советник, руководитель Практики интеллектуальной собственности и технологий, член Комиссии по интеллектуальной собственности Международной торговой палаты (ИСС) в России, вице-президент IP CLUB — российского некоммерческого экспертно-образовательного проекта в сфере права интеллектуальной собственности, права информационных технологий, цифрового права (IP / IT / Digital Law), соискатель ученой степени кандидата юридических наук

Аннотация. Статья посвящена исследованию правовых аспектов использования авторских произведений в качестве обучающих данных в технологиях машинного обучения (МО). Автор анализирует специфику МО как одной из разновидностей технологий искусственного интеллекта (ИИ) и рассматривает юридические проблемы, возникающие при использовании авторских произведений без согласия правообладателей, а также персональных данных.

Рассмотрены различные подходы к решению проблемы несанкционированного правообладателем использования произведений, охраняемых авторским правом. Предложены такие варианты, как разрешение использования в технологиях МО только произведений, перешедших в общественное достояние; дополнение законодательства о свободном использовании про-

- изведений для целей МО; заключение договоров об отчуждении исключительного права или лицензионных договоров.
- Приведены примеры из зарубежной судебной практики, включая дела против крупных разработчиков систем ИИ, таких как Stability AI, Open AI; оценены перспективы рассмотрения подобных дел с позиций российского права. Предложено введение в российское законодательство конструкции, допускающей свободное использование охраняемых результатов интеллектуальной деятельности для целей развития ИИ с учетом общественных интересов.
- В заключение сделан вывод о необходимости правового регулирования использования обучающих данных в МО с учетом баланса интересов общества и правообладателей. Предложено закрепить в законодательстве возможность свободного использования произведений на условиях безвозмездности исключительно для целей обучения ИИ, а также предусмотреть использование авторских произведений с согласия правообладателя в случае коммерческого применения результатов, созданных с помощью ИИ.
- **Ключевые слова:** искусственный интеллект, машинное обучение, произведения, базы данных, самообучение, конфиденциальные данные, персональные данные, авторские права, интеллектуальная собственность

- **Для цитирования:** Самарцева М.В. Использование охраняемых произведений при машинном обучении генеративного искусственного интеллекта // Труды по интеллектуальной собственности (Works on Intellectual Property). 2025. Т. 53, № 2. С. 127–138;
• DOI: 10.17323/tis.2025.27374

Технологии искусственного интеллекта (ИИ) предполагают использование наборов данных и специально разработанных алгоритмов, что в совокупности позволяет ИИ решать разнообразные задачи, предлагать решения и строить прогнозы, а кроме того, обучаться и самообучаться в процессе разрешения обозначенных вопросов. Иными словами, ИИ, если это заложено соответствующими алгоритмами, может самостоятельно совершенствоваться путем анализа входных данных — эти возможности создает технология машинного обучения (*Machine Learning, ML*; далее — МО), позволяющая ИИ быстро обучаться на основе больших наборов данных в автоматизированном режиме с помощью заложенных в систему алгоритмов, а не в соответствии с указаниями человека.

Как показывает анализ юридической литературы, сегодня понятие МО употребляется едва ли не чаще понятия ИИ. Причем в одних случаях эти термины используются как тождественные, а в других — как различающиеся [1, 2], что нуждается в анализе в целях настоящего исследования.

Понятие «*Machine Learning*», которое в 1959 г. было введено ученым-компьютерщиком Артуром Самюэлем [3], изначально трактовалось как способность компьютера обучаться, не будучи явно запрограммированным на выполнение конкретной задачи, что подразумевало обучение компьютера распознаванию. Необходимость в обучении распознаванию обусловлена тем, что компьютер не может осязать, видеть, слышать в привычном для человека понимании — проведение анализа объектов становится доступно машине только в том случае, если анализируемый объект представлен в двоичном (бинарном) коде, состоящем из символов 0 и 1. Вследствие этого абсолютно все данные, в качестве которых могут выступать числа, слова, изображения, мелодии и прочее, требуют представления в двоичном коде.

С момента опубликования А. Самюэлем его работы содержание понятия МО несколько изменилось и в сфере применения технологий ИИ подразумевает способность машины на основе анализа наборов данных вырабатывать самостоятельные решения и прогнозы без явного участия (вмешательства) человека. Поясняя принципы работы МО, М.А. Рожкова пи-

шет: «Технология, предполагающая машинное обучение, требует создания человеком собственно самой модели — вычислительной (информационной) системы, в которой в качестве основных составляющих можно выделить решатель (компьютерная программа, которая на основе специальных алгоритмов будет разрешать поставленные задачи — делать прогнозы и вырабатывать решения), базы данных (создаваемые человеком для изучения компьютером конкретной предметной области знаний) и интеллектуальный интерфейс (совокупность средств, методов и правил, которые позволяют человеку вести общение с моделью). Для самого же машинного обучения необходима обучающаяся система, снабжаемая огромными объемами “обучающих данных” (*training data*), анализ и интерпретация которых и позволяет компьютеру впоследствии корректировать параметры модели, обновляя и улучшая ее с целью получения более точных результатов. Таким образом, компьютер, анализируя наборы “обучающих данных”, обобщает эти данные и выводит определенные закономерности (паттерны). Применение этих закономерностей при обработке следующих наборов “обучающих данных” позволяют компьютеру “учиться на основе предыдущих вычислений, чтобы вырабатывать достоверные, повторяющиеся решения и результаты”, а также совершенствовать модель для достижения более точных результатов» [4].

Сказанное позволяет подчеркнуть специфику МО. Так, в обычных компьютерных программах разработчик (человек) с помощью программного кода создает набор правил, алгоритм шагов, в точном соответствии с которыми система будет действовать, а в случае МО модель получает входные данные, анализируя и обрабатывая которые, она вырабатывает собственные алгоритмы действий для обработки следующей входной информации, то есть в рамках МО речь идет именно об обучении ИИ, а не о перепрограммировании алгоритма принятия решений (причем алгоритмы, по которым осуществляется МО, являются заранее определенными). В связи с этим в литературе отмечается: «Подходы машинного обучения часто обучают машины достигать результата, показывая им множество примеров правильных ре-

зультатов. Кроме того, они также могут определить набор правил и позволить машине учиться методом проб и ошибок. Машинное обучение обычно используется для построения или корректировки модели поведения, но также может использоваться для интерпретации результатов модели» [5].

Проанализировав результаты использования машиной обучающих данных, М.А. Рожкова подчеркнула, что МО находит наибольшее применение в двух направлениях: «В зависимости от поставленной задачи они могут либо делать прогнозы будущих результатов (например, прогнозировать потенциальную торговлю определенными акциями, рассчитывать более точные модели прогнозируемого ценообразования, оценивать имущественные риски при страховании, прогнозировать пожизненную ценность клиента и т.п.), либо вырабатывать решения на основании классификации входных данных. При этом второе направление, крайне впечатляющее по охвату, включает: создание и использование на веб-сайтах чат-ботов, отвечающих на часто задаваемые вопросы; выявление подозрительных (нетипичных) транзакций на основе информации о ставших известными мошеннических транзакциях в банках и финансовых учреждениях, интернет-магазинах, страховых организациях и пр.; извлечение значимой информации из цифровых изображений, видео и других визуальных данных (компьютерное зрение) с целью заданного использования; распознавание речи для выполнения голосового поиска в мобильных устройствах и т.д. В итоге, как отмечается в работе [4], «машинное обучение приобрело промышленную значимость: оно позволяет делать прогнозы и помогает принимать решения, основанные, например, на выявленном мошенничестве в торговых сетях, выстроенной компьютером логистической цепочке, предложении инновационного сервиса наподобие таргетинга и пр.».

Изложенное позволяет согласиться с выводом, что между ИИ и МО нет оснований ставить знак равенства: МО представляет собой одну из разновидностей технологии ИИ, практический метод применения ИИ, который надо рассматривать в качестве альтернативы применения разновидности ИИ, выстраивающей свою работу исключительно по тем инструкциям, которые заложены в нее человеком.

В развитие сказанного стоит отметить то обстоятельство, что применение обучающих данных в рамках МО связывается с рядом юридических проблем.

Одной из самых острых проблем признается возможность попадания в наборы обучающих данных таких, которые принято квалифицировать в качестве конфиденциальных. Это, в частности, персональные данные, информация о геолокации, а также различно-

го рода тайны: медицинские, банковские, коммерческие, связи и прочие. Вследствие этого в Концепции развития регулирования в сфере технологий искусственного интеллекта и робототехники до 2024 года, утвержденной распоряжением Правительства РФ от 19.08.2020 № 2129-р «Об утверждении Концепции развития регулирования отношений в сфере технологий искусственного интеллекта и робототехники до 2024 года» (далее — Концепция развития ИИ), в качестве одного из концептуальных проблемных направлений правового регулирования было обозначено соблюдение баланса между требованиями по защите персональных данных и необходимостью их использования для обучения систем ИИ.

Анализируя проблему соблюдения такого баланса, В.Б. Наумов и Е.В. Тютюк отмечают, что защита данных может осуществляться с помощью так называемых маскировочных технологий (*safe harbour*), обеспечения получения согласия субъекта данных, требования по минимизации данных [6]. Также авторы обращают внимание на предпринятую законодателем попытку разрешить обозначенную проблему посредством проведения эксперимента по установлению специального регулирования для внедрения ИИ, что было реализовано в Законе о проведении эксперимента для внедрения ИИ, предусматривающем использование в МО только обезличенных персональных данных. Следует отметить, что в итоге авторы признают предлагаемое решение небесспорным, поскольку «обезличивание персональных данных не всегда гарантирует их полноценную защиту» [6].

Еще одной проблемой, но уже частноправового характера, признается неопределенность в части принадлежности прав на наборы обучающих данных. В контексте российского права эта проблема имеет концептуальный характер ввиду неоднозначности вопроса о допустимости признания имущественных прав на данные (информацию). Обозначенная проблема во многом обусловлена произведенным в 2008 г. исключением из ст. 128 ГК РФ упоминания категории «информация», в результате чего преобладающим в юриспруденции стало мнение, что ни информация, ни данные (в качестве ее разновидности) не могут становиться объектами гражданских прав. В юридической литературе иногда высказываются предложения, нацеленные на преодоление данной преграды на пути разрешения практических задач (как правило, в качестве основного проблемного вопроса называют затруднения с оборотом данных и больших данных, вызванные пробелами в действующем законодательстве). Одним из таких предложений стала идея ввести в гражданское законодательство в качестве нового объекта гражданских прав «массив

данных», установив для него специальный правовой режим [7], что не было поддержано в доктрине [8].

В целях настоящего исследования должна быть тщательно проанализирована другая проблема применения обучающих данных — использование в этом качестве авторских произведений. На практике это вылилось в иски правообладателей к разработчикам ИИ о нарушении авторских прав ввиду того, что произведения без согласования с авторами использовались в качестве обучающих данных в МО.

Примечательно, что такие иски в зарубежной практике часто предъявляются не единичными авторами, а совместно несколькими правообладателями (коллективные иски), причем нередко — со множественностью лиц на стороне ответчика.

Одним из первых случаев предъявления требования о защите интеллектуальных прав художников, чьи работы были использованы в качестве обучающих данных, стало дело художниц Сары Андерсен (*Sarah Andersen*), Келли МакКернан (*Kelly McKernan*) и Карлы Ортиз (*Karla Ortiz*), которые обратились в суд в январе 2023 г. с иском к крупнейшим компаниям, занимающимся генеративным ИИ, — генераторам изображений *Stability AI*, *Midjourney* и *DeviantArt* (*Andersen v. Stability AI Ltd.*)¹.

В этом иске утверждалось, что компании-ответчики нарушили права истцов, когда без согласования использовали принадлежащие последним художественные произведения для МО. При этом ответчики², не отрицая факта такого использования, указывали, что обучение машин на уже существующих авторских изображениях подпадает под категорию добросовестного использования.

Окружной суд частично удовлетворил и частично отклонил ходатайства ответчиков, признав, что претензии истцов о косвенном нарушении авторских прав со стороны *Stability AI* имеют достаточные основания для продолжения судебного разбирательства. Однако требования по ст. 1202(a) и 1202(b) Закона о защите авторских прав в цифровую эпоху (*DMCA*)³

¹ Дело № № 3:23-cv-00201 District Court, N.D. California. — URL: https://www.courtlistener.com/docket/66732129/andersen-v-stability-ai-ltd/?filed_after=&filed_before=&entry_gte=&entry_lte=&order_by=desc

² Ответчики по этому делу: *DeviantArt, Inc.*, *Getty Images (US), Inc.*, *Midjourney, Inc.*, *Runway AI, Inc.*, *Stability AI, Inc.*, *Stability AI Ltd.*

³ «Акт о защите авторских прав в цифровом тысячелетии» (*DMCA*), принятый в США в 1998 г. Этот акт ввел ответственность (в том числе и уголовную) за обозначенные в ст. 11 и 12 Договора ВОИС составы правонарушений. Параграф 1202 *DMCA* предусматривает ответственность за размещение недостоверной информации, за подмену либо удаление информации о принадлежности авторских прав, включая им-

были отклонены без права на повторное предъявление, поскольку суд не нашел достаточных доказательств для их поддержки.

Также суд отклонил ходатайства *Midjourney* об отклонении исков о ложном одобрении и защите торгового облика бренда (*trade dress*), указав, что истцы представили достаточные доказательства возможного введения потребителей в заблуждение. Суд также отклонил иск о нарушении контракта против *DeviantArt* и иск о неосновательном обогащении, указав, что последний дублирует претензии по нарушению авторских прав. В некоторых случаях истцам была предоставлена возможность внести изменения в исковые заявления. На начало апреля 2025 г. данное дело еще находилось на рассмотрении в Окружном суде Калифорнии, США.

Доктрина добросовестного использования (*fair use*) основана на параграфе 107 Закона США об авторском праве (*US Copyright Act of 1976*), который закрепляет четыре критерия для определения допустимости использования объекта авторского права без согласия правообладателя. Описывая их, Е.В. Балобанова отмечает, что «к числу общих критериев (факторов) *fair use* относятся: 1) цель и характер использования (коммерческий или некоммерческий); 2) характер произведения (особенности создания и др.); 3) количественные и содержательные характеристики заимствования; 4) характер предполагаемого воздействия на потенциальный рынок» [9]. Довольно обширная судебная практика применения этих критериев способствовала не только уточнению содержания этой доктрины, но и серьезному расширению границ ее применения, что в конечном счете может привести к размыванию ориентиров и затруднениям в использовании указанных критериев.

Применительно к первому критерию определяющим будет то, носит ли несогласованное использование коммерческий характер или осуществляется для некоммерческих (например, образовательных) целей. В качестве общего правила нередко указывают, что приоритет в признании добросовестным использованием будет за применением, которое не преследовало коммерческой (финансовой) выгоды, однако А.С. Ворожечин пишет, что «американские суды сформулировали и обратное правило: некоммерческая цель использования сама по себе не означает, что использование является добросовестным, хотя и повышает шанс на то, что оно будет признано таковым» [10]. В свою очередь, О.В. Луткова подчеркивает:

порт, за распространение или иное введение в оборот копий произведений, в которых такая информация была удалена или изменена.

«В американской правоприменительной практике упрочилась позиция, что добросовестным следует считать преобразующее, творческое последующее использование произведения, независимо от его коммерческого или некоммерческого характера, поскольку многие, если не все, вторичные работы создаются, по крайней мере, с неосновной целью получения хотя бы некоторой коммерческой выгоды от их использования» [11]. Резюмируя, Л.В. Сагдеева указывает, что использование произведения без согласия правообладателя считается добросовестным до тех пор, «пока оно способствует “развитию науки и полезных ремесел”, что отличается от выдачи официальной лицензии (частной или публичной) на использование объекта авторских прав» [12].

Второй критерий — это характер первичного авторского произведения, применительно к которому П.С. Могилевский замечает: «При оценке данного фактора суд будет обращать внимание на то, имело ли использованное произведение творческий характер» [13].

Третий критерий — это объем и существенность заимствований из первичного авторского произведения, что определяется и оценивается судом в каждом конкретном случае. По мнению А.С. Ворожевич, при оценке поведения на соответствие данному критерию «учитываются количественные показатели использования (число копий, созданных нарушителем, объем используемой части произведения)» [10].

Четвертый критерий предполагает решение вопроса о том, насколько анализируемое использование влияет на рыночную стоимость первичного авторского произведения. То есть, по сути, оценивается экономический ущерб, причиненный правообладателю.

Как подчеркивает А.С. Ворожевич, «американские суды исходят из того, что приведенные критерии должны оцениваться с учетом особенностей каждого конкретного случая» [10], поэтому следование критериям доктрины добросовестного использования вовсе не обеспечивает предсказуемость судебных решений по такого рода делам. Следовательно, предсказать, каким должно быть итоговое решение по обозначенному делу *Andersen v. Stability AI Ltd.* было бы довольно затруднительно.

Доктрина добросовестного использования действует отнюдь не во всех юрисдикциях — в законодательстве стран континентальной системы права эта доктрина не нашла отражения. На это обратил внимание В.О. Калятин, подчеркнувший, что доктрина *fair use* является «одной из базовых концепций, используемых в странах англосаксонского права для определения допустимости использования произведения пользователем, в отличие от стран континентального

права, где установление вариантов разрешенного поведения пользователя произведения осуществляется путем прямого перечисления ограничений исключительного права» [14].

Здесь важно заметить, что на уровне Европейского союза принята Директива 2001/29/ЕС Европейского парламента и Совета ЕС от 22.05.2001 о гармонизации определенных аспектов действия авторского права и смежных прав в информационном обществе. Этой директивой «государствам-членам предоставляются свобода выбора и возможность адаптации в соответствии со своими традициями определенного набора ограничений из представленных в перечне и, следовательно, определенная свобода в формировании своего собственного ограничительного механизма» [9]. А в ст. 5(5) Директивы закреплена трехступенчатый тест для определения границ свободного использования произведений (с учетом правил национального регулирования): предусматривается, что исключения и ограничения могут быть применены только в некоторых особых случаях, и они не должны наносить ущерба нормальному использованию произведений и необоснованно ущемлять законные интересы авторов.

Подход, закрепляющий ограничения исключительного права, проистекает из положений Бернской конвенции, в п. 2 ст. 9 которой предусмотрено право стран-участниц закрепить в национальном законодательстве особые случаи воспроизведения произведений при условии, что такое воспроизведение не наносит ущерба нормальному использованию произведения и не ущемляет необоснованным образом законные интересы автора, а в ст. 10 и 10bis перечислены конкретные (возможные) случаи свободного использования произведений.

В российском праве и легальные способы использования объектов интеллектуальных прав без согласия правообладателя, и ограничения исключительного права закреплены непосредственно в действующем законодательстве — в части четвертой Гражданского кодекса РФ. Причем для обозначения разрешенного использования объектов интеллектуальных прав третьими лицами без согласия правообладателя используются такие термины, как «свободное использование результатов интеллектуальной деятельности» (ст. 1273–1279, 1306 Гражданского кодекса РФ) и «действия, не являющиеся нарушением исключительного права» (ст. 1335.1, 1359, 1422, 1456, 1487, п. 5 ст. 1519 Гражданского кодекса РФ).

Таким образом, российское законодательство, очевидно, закрепляет иной подход для решения вопроса о допустимости использования авторских произведений без согласия их авторов, нежели *fair use*.

В этих условия несомненный интерес представляет исследование В.А. Чебодаевой [15], которая приняла попытку дать оценку перспектив рассмотрения обозначенного выше дела *Andersen v. Stability AI Ltd.* с позиций российского права. Подобные авангардные изыскания представляются чрезвычайно полезными, поскольку в набор обучающих данных могут попадать авторские произведения, правообладателями которых являются граждане разных стран.

Оценив обстоятельства дела и доводы сторон, изложенные в процедурных документах, В.А. Чебодаева подчеркнула, что сам факт использования произведений истцов подтверждался указаниями на сайте, содержащем сведения о том, какие из работ художников нашли применение в качестве обучающих данных для генератора иллюстраций. На основе произведенного анализа она делает вывод, что по смыслу российского права в тех случаях, когда авторские произведения выступают в качестве обучающих данных, сначала имеет место их воспроизведение (посредством записи в памяти компьютера), а затем переработка путем применения соответствующего алгоритма. Это, по мнению В.А. Чебодаевой, в силу ст. 1270 Гражданского кодекса РФ должно рассматриваться как использование произведений, допускаемое лишь с согласия автора. Поскольку в упомянутом деле произведения художников включались в наборы обучающих данных без их согласия, В.А. Чебодаева сделала вывод, что в рассматриваемом случае есть все основания признать состоявшимися нарушения исключительных прав. Кроме того, она считает нарушенными и личные немущественные права художников, поскольку при воспроизведении в цифровой форме нарушается право на неприкосновенность произведений в значении ст. 1266 Гражданского кодекса РФ.

Примечательно, что в своем исследовании В.А. Чебодаева не остановилась на указанных выводах, а приняла попытку оценить, можно ли в контексте российского законодательства признать нарушением авторского права генерацию ИИ результатов в фирменном стиле истцов. Но в итоге она была вынуждена признать отсутствие эффективных механизмов защиты прав и законных интересов авторов произведений в ситуации, когда ИИ генерирует контент, созданный в фирменном, узнаваемом стиле художников (хотя и предложила решать проблему посредством обращения к нормам конкурентного законодательства).

Далеко не со всеми доводами, изложенными в статье В.А. Чебодаевой, можно согласиться безоговорочно. Например, делая вывод о том, что имело место воспроизведение спорных произведений в памяти компьютера, автор не учитывала, что генератор изображений *Stable Diffusion* был обучен на данных

германской компании *LAION*, которая утверждала, что не воспроизводит и не хранит сами изображения — данные этой компании представляют собой только список URL-адресов, ведущих к изображениям в интернете, а также список извлеченных по ним текстовых описаний [16]. В этих условиях встает вопрос о том, имело ли место использование ИИ самих объектов авторских прав, как оно закреплено в действующем законодательстве.

В ст. 1270 Гражданского кодекса РФ говорится об использовании произведения в любой форме и любым способом и приводится примерный перечень способов использования. Вместе с тем надо признать, что действия по использованию обращены к форме произведения, а не к его содержанию. В связи с этим говорить об использовании произведения, созданного ИИ, можно, например, в том случае, когда произведение (или его часть) воспроизводится (осуществляется его копирование на иной носитель, в том числе в память того устройства, с которым связан искусственный интеллект), публично исполняется или демонстрируется, перерабатывается или к нему предоставляется доступ третьим лицам в результате переработки и выдачи итогов работы искусственного интеллекта.

В таких условиях вряд ли можно говорить об использовании изображения, если применялось его текстовое описание. Представляется очень уместным процитировать замечание А.А. Шутыкова и А.Н. Анищенко о том, что «понятие “использование” по своему объему (сфере применения) и по содержанию практически для каждого охраняемого объекта является своим, особым, специфическим» [17].

Вместе с тем правовые исследования, подобные проанализированной работе В.А. Чебодаевой, свидетельствуют о том, что кажущаяся, на первый взгляд, довольно частной проблема использования авторских произведений в качестве обучающих данных, которая изначально рассматривалась как проблема прикладного, правоприменительного характера, нуждающаяся в анализе правомерности и обоснованности подобных требований, на самом деле имеет концептуальный характер. Иными словами, требует доктринальной проработки вопрос о допустимости квалифицировать использование авторских произведений в качестве обучающих данных как «свободное использование» или как «действия, не являющиеся нарушением исключительного права» (в терминологии Гражданского кодекса РФ).

Надо заметить, что предъявляемые иски о нарушении авторских прав посредством использования произведений, охраняемых авторским правом, в качестве обучающих данных характерны для зарубежной судебной практики. Вместе с тем обобщение и анализ

такой судебной практики в контексте российского права может сработать на опережение и позволит поставить этот вопрос на обсуждение правоведов еще до того, как такие иски начнут массово возникать в российской практике.

Изложенное обуславливает необходимость обратить внимание на выявленные А.С. Васильевой варианты разрешения проблемы несогласованного использования авторских произведений в качестве обучающих данных [18].

Первым прозвучало предложение ограничить разработчиков в возможности использовать для МО любые произведения, ограничив их только теми, которые перешли в общественное достояние. Однако этот подход, представляющийся самым простым с юридической точки зрения, не поддерживают разработчики, поскольку результаты, генерируемые ИИ, в таком случае не будут отвечать современным реалиям и современным потребностям, что, по признанию самой А.С. Васильевой, делает данный вариант, хотя и самым приемлемым для действующего законодательства, но не лучшим в условиях, когда «общество стремится к развитию новых технологий и их внедрению в повседневную практику» [18].

В качестве второго подхода к разрешению исследуемой проблемы А.С. Васильева предложила дополнить нормы Гражданского кодекса РФ о свободном использовании авторского произведения следующим способом использования: «воспроизведение и переработка объектов авторского произведений и баз данных исключительно для автоматизированного анализа в целях разработки и обучения искусственного интеллекта (первоначального или в процессе совершенствования) с исключением возможности использования этих объектов за пределами указанной цели и обязанностью удалить соответствующую копию при отпадании такой цели» [19]. При этом она обоснованно отметила, что реализация подобного предложения может быть осуществлена только путем закрепления соответствующей нормы непосредственно в российском законодательстве, тогда как за рубежом, например в США, введение в закон подобной нормы необязательно.

Третьим стало предложение заключать договоры об отчуждении исключительного права для МО. По мнению А.С. Васильевой, преимущества этого варианта состоят в прозрачности (понятно, какие объекты интеллектуальной собственности использовались для обучения), а также в возможности получения правообладателями материальной выгоды от отчуждения исключительного права. Недостатки же состоят в значительных дополнительных материальных затратах на заключение возмездных договоров.

В качестве четвертого варианта разрешения исследуемой проблемы А.С. Васильева назвала возможность заключения лицензионного договора с правообладателем объекта. Ключевое преимущество этого варианта видится ей в установлении баланса интересов между правообладателем и разработчиком: «В случае заключения данного договора целью договора выступает использование охраняемого объекта для обучения искусственного интеллекта, что включает в себя использование в виде воспроизведения и переработки в ходе создания (генерации) искусственным интеллектом чего-либо» [18].

Тщательный анализ предлагаемых А.С. Васильевой вариантов разрешения проблемы использования авторских произведений в качестве обучающих данных без согласия на это авторов позволяет утверждать следующее.

Достаточно сложно поддержать первый предлагаемый вариант, поскольку даже сама автор подспудно считает его непригодным — в условиях бурного внедрения инноваций подобный подход будет шагом назад.

Сложнореализуемыми и практически нерентабельными следует признать третий и четвертый варианты, предусматривающие заключение договоров авторов с разработчиками, поскольку крайне затратным представляются отыскание авторов всех произведений, включаемых в наборы обучающих данных, установление, являются ли эти произведения объектами авторского права по действующему законодательству, обеспечение заключения с ними договоров и выплата соответствующего вознаграждения. Той же позиции придерживается М.О. Кухно, который отметил: «Такой путь во многих случаях выглядит нереалистичным (особенно при анализе большого объема данных), поскольку это затратный (по времени и денежным средствам) процесс индивидуального взаимодействия с каждым из правообладателей» [20].

Применительно к четвертому подходу следует заметить также, что с практической точки зрения едва ли можно определить, произошло ли использование конкретного авторского произведения во время генерации ИИ нового результата, то есть существует ли проблема определения конкретного автора и выплаты ему вознаграждения за использование именно его произведения.

В этих условиях представляется правильным, опираясь на второй вариант из предложенных, предполагающий свободное использование авторских произведений для целей МО, проанализировать допустимость его закрепления в российском праве.

Как уже указывалось, действующее законодательство предусматривает две конструкции использова-

ния результата творческой деятельности без согласия автора: в качестве «свободного использования результатов интеллектуальной деятельности» и в качестве «действий, не являющихся нарушением исключительного права». Думается, применительно к МО предпочтителен первый вариант, поскольку он допускает установление границ такого использования, что несколько сложнее осуществить в рамках второй категории.

С учетом сказанного представляется верным обратиться к положениям Директивы (ЕС) 2019/790 Европейского Парламента и Совета от 17.04.2019 об авторском праве и смежных правах на едином цифровом рынке и о внесении изменений в Директивы 96/9/ЕС и 2001/29/ЕС (*Directive (EU) 2019/790 of the European Parliament and of the Council of 17 April 2019 on copyright and related rights in the Digital Single Market and amending Directives 96/9/EC and 2001/29/EC*). Директива (ЕС) 2019/790 закрепляет две разновидности ограничений и исключений в области интеллектуальных прав авторов и иных правообладателей, причем эти исключения были прописаны до появления генеративного ИИ и связанных с ним вопросов авторского права.

Первое исключение, закрепленное в ст. 3 названной Директивы, подразумевает в целях проведения научных исследований свободное воспроизведение и извлечение для интеллектуального анализа текстов и данных. Это исключение предназначено для деятельности аккредитованных исследовательских организаций и учреждений культурного наследия, причем они могут использовать только те данные, к которым они имеют законный доступ, в связи с чем Д.А. Мотовилова пишет: «Как разъясняется в п. 1 преамбулы, доступ признается законным, если между правообладателями и научно-исследовательскими организациями или учреждениями культурного наследия имеются договорные соглашения (например, подписка) или исследуемый контент свободно доступен онлайн» [21].

Второе исключение, закрепленное в ст. 4 названной Директивы, охватывает более широкий круг субъектов, включая коммерческие организации, которым разрешено воспроизведение и извлечение произведений для целей интеллектуального анализа текстов и данных. При этом Д.А. Мотовилова отмечает: «Однако такое воспроизведение и извлечение возможно только в том случае, если к произведению или иному охраняемому объекту имеется законный доступ со стороны пользователя, в том числе, когда оно было доведено до всеобщего сведения онлайн, и если правообладатели не сохранили за собой надлежащим образом права на воспроизведение и извлечение для целей интеллектуального анализа текстов и данных

(см. ст. 4(3) и абз. 2 п. 18 преамбулы). При этом... фактически правообладатели могут запретить подобное использование охраняемого контента» [21].

Обозначенные подходы представляются недостаточно продуманными и несколько половинчатыми: допуская возможность несогласованного использования произведений, Директива предусматривает, что правообладатели могут заявить требование о неправомерном использовании произведения. Предоставляя исследовательским организациям право использовать такие данные, Директива не позволяет делать это в коммерческих целях.

Более правильным в целях эффективного развития ИИ в Российской Федерации, повышения конкурентоспособности, расширения экономических, экологических и социальных преимуществ Российской Федерации во всем спектре отраслей и в социальной сфере допустить беспрепятственное и безвозмездное использование авторских произведений в качестве обучающих данных для МО (посредством введения в законодательство специального исключения в сфере авторских прав). Иными словами, в интересах общества представляется правильным создание правовой конструкции, допускающей свободное использование охраняемых результатов интеллектуальной деятельности для целей обучения и развития стратегически значимых технологий. В поддержку предлагаемого подхода можно сослаться на следующее заключение: «По своей природе создание и обучение искусственного интеллекта предполагает использование возможно большего количества примеров, и его эффективность в будущем во многом зависит от выбора и количества используемых для обучения примеров. Необходимость получения разрешения на использование каждого такого объекта сделает создание искусственного интеллекта крайне затруднительным. Нельзя забывать и о том, что правообладатель может потребовать уплаты ему вознаграждения, а это приведет к значительному возрастанию расходов на создание искусственного интеллекта» [19]. Причем иной подход, как подчеркивается в настоящем исследовании, способен затруднить развитие ИИ в нашей стране.

Вследствие сказанного в целях сохранения баланса общественных интересов и интересов авторов и иных правообладателей произведений предлагается дополнить действующее законодательство следующими положениями. Во-первых, указанием на допустимость свободного использования произведений для обучения ИИ на условиях, приравненных к свободному использованию в образовательных целях, то есть безвозмездно и без согласования с правообладателем, при условии предоставления пользователям возмож-

ности пользования информационными системами, использующими этот ИИ, на условиях безвозмездности. Во-вторых, в ситуации, когда подразумевается коммерческое применение ИИ, использование правомерно обнародованных произведений в качестве обучающих данных для МО осуществляется с согласия правообладателя на основании лицензионного договора.

Подобные положения должны найти закрепление не в Законе об информации, как полагает, в частности, Н.А. Назаров, предложивший предоставить лицам неограниченный доступ к объектам интеллектуальной собственности в контексте сбора, обработки и использования данных и сославшийся при этом на дополнение Закона об информации [22] (в ст. 1 которого указано, что его положения не распространяются на сферу интеллектуальной собственности), а в части четвертой Гражданского кодекса РФ, поскольку речь идет об одном из вариантов использования объектов интеллектуальной собственности.

В целях практической реализации выработанных научных положений и сделанных выводов в рамках участия в рабочей группе по разработке проекта Федерального закона «О регулировании систем искусственного интеллекта» были сформулированы следующие нормативные положения о дополнении ст. 1274 Гражданского кодекса пунктом 5 следующего содержания:

«5. Допускается без согласия автора или иного правообладателя и без выплаты вознаграждения использование правомерно обнародованных произведений для обучения генеративной системы искусственного интеллекта при условии предоставления пользователям возможности пользования этой системой на условиях безвозмездности.

В случае предоставления пользователям возможности пользоваться такой системой только на возмездных основаниях, использование правомерно обнародованных произведений для обучения генеративной системы искусственного интеллекта осуществляется с согласия правообладателя на основании лицензионного договора».

Изложенное в настоящем исследовании позволяет сделать ряд выводов.

Применение обучающих данных в рамках МО порождает ряд правовых проблем. Одной из самых острых признается проблема применения в качестве обучающих данных авторских произведений без согласия на это правообладателей таких произведений.

В целях сохранения баланса общественных интересов и интересов авторов и иных правообладателей правомерно обнародованных произведений пред-

ставляется правильным предусмотреть две конструкции использования таких произведений для МО.

Первый вариант — установление допустимости свободного использования авторских произведений для обучения ИИ на условиях, приравненных к свободному использованию в образовательных целях, то есть безвозмездно и без согласования с правообладателем, при условии предоставления пользователям возможности пользования информационными системами, применяющими этот ИИ, на условиях безвозмездности.

Второй вариант — закрепление для случаев, когда изначально предполагается коммерческое использование ИИ, возможности использования авторских произведений в качестве обучающих данных для МО только с согласия правообладателя на основании лицензионного договора.

СПИСОК ИСТОЧНИКОВ

1. Kühl N., Goutier M., Hirt R., Satzger G. Machine Learning in Artificial Intelligence: Towards a Common Understanding. 2019. — URL: <https://arxiv.org/ftp/arxiv/papers/2004/2004.04686.pdf>
2. Wakefield K. A guide to the types of machine learning algorithms and their applications. — URL: https://www.sas.com/en_gb/insights/articles/analytics/machine-learning-algorithms.html
3. Samuel Arthur L. Some Studies in Machine Learning Using the Game of Checkers // IBM Journal. 1959. Vol. 3, No. 3. P. 210—229. — URL: <https://people.csail.mit.edu/brooks/idocs/Samuel.pdf>
4. Рожкова М.А. Машинное обучение как метод применения искусственного интеллекта: суть технологии и обзор основных правовых проблем // Закон.ру. 2024. 18 марта. — URL: https://zakon.ru/blog/2024/03/18/mashinnoe_obuchenie_kak_metod_primeneniya_iskusstvennogo_intellekta_sut_tehnologii_i_obzor_osnovnyh/
5. Artificial Intelligence in Society, OECD Publishing, Paris, 2019. — URL: <https://www.oecd-ilibrary.org/docserver/eedfee77-en.pdf>
6. Наумов В.Б., Тютюк Е.В. Правовые проблемы машинного обучения // Образование и право. 2020. № 6. С. 219—231. — URL: <https://cyberleninka.ru/article/n/pravovye-problemy-mashinnogo-obucheniya?ysclid=ltigehrqf2339477194>
7. Савельев А.И. Гражданско-правовые аспекты регулирования оборота данных в условиях попыток формирования цифровой экономики // Вестник гражданского права. 2020. № 1. С. 67.
8. Рожкова М.А. Нематериальные объекты и имущественные права на них в системе абсолютных прав.

- М.: ГАУГН Пресс, 2024. 258 с. Серия: «IP & Digital Law».
9. Балобанова Е.В. К вопросу о применимости доктрины добросовестного использования в континентальном праве // Журнал Суда по интеллектуальным правам. 2017. № 18. С. 92–99. — URL: <https://ipcmagazine.ru/articles/1729017/>
 10. Ворожевич А.С. Fair use и доктрина «расширяющего полезность использования». — URL: <https://legal-support.ru/information/blog/ip-daidzhest/fair-use-i-doktrina-rasshiryayushhego-poleznost-ispolzovaniya/>
 11. Луткова О.В. Доктрина добросовестного использования произведений в современном авторском праве США // Право. Журнал Высшей школы экономики. 2016. № 2. С. 186–199.
 12. Сагдеева Л.В. Свободное использование как ограничение исключительных прав // Юридические исследования. 2017. № 9.
 13. Могилевский П.С. Доктрина добросовестного использования в США и свободное использование произведений в России: сравнительный анализ // Гражданское право. 2022. № 5. С. 12–14.
 14. Калятин В.О. О перспективах применения в России доктрины добросовестного использования // Закон. 2015. № 11. С. 40–47.
 15. Чебодаева В.А. Контент, сгенерированный нейросетью: кто виноват и что делать? // Журнал Суда по интеллектуальным правам. 2023. № 4 (42). С. 46–58. — URL: <https://ipcmagazine.ru/articles/1729306/>
 16. Создатели нейросетей обвиняются в нарушении авторских прав художников. 19 января 2023. — URL: <https://22century.ru/computer-it/113839?ysclid=m7wg2c00p0538107917>
 17. 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.
 18. Васильева А.С. Использование объектов авторских прав для обучения искусственного интеллекта // Журнал Суда по интеллектуальным правам. 2024. № 4 (46). С. 80–84. — URL: <https://ipcmagazine.ru/articles/1765658/>
 19. 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.: ИД ВШЭ, 2021. 28 с. С. 22. — URL: <https://www.hse.ru/mirror/pubs/share/482492820.pdf>
 20. Кухно М.О. Искусственный интеллект — новый субъект авторского права: недалекое будущее или фикция? // Журнал Суда по интеллектуальным правам. 2023. Вып. 3 (41). С. 72–86. — URL: <https://ipcmagazine.ru/articles/1729283>
 21. Мотовилова Д.А. Ключевые положения Директивы (ЕС) № 2019/790 об авторском праве и смежных правах на едином цифровом рынке // Журнал Суда по интеллектуальным правам. 2019. № 26 (декабрь). С. 106–120. — URL: <https://ipcmagazine.ru/articles/1729025/>
 22. Назаров Н.А. Машинное творчество и право: две части одного целого // Труды по интеллектуальной собственности (Works on Intellectual Property). 2022. Т. 43. № 4. С. 101–110. — URL: <https://tis.hse.ru/article/view/16358/15428>

REFERENCES

1. Kühl N., Goutier M., Hirt R., Satzger G. Machine Learning in Artificial Intelligence: Towards a Common Understanding. 2019. — URL: <https://arxiv.org/ftp/arxiv/papers/2004/2004.04686.pdf>
2. Wakefield K. A guide to the types of machine learning algorithms and their applications. — URL: https://www.sas.com/en_gb/insights/articles/analytics/machine-learning-algorithms.html
3. Samuel Arthur L. Some Studies in Machine Learning Using the Game of Checkers // IBM Journal. 1959. Vol. 3, No. 3. July, P. 210–229. — URL: <https://people.csail.mit.edu/brooks/idocs/Samuel.pdf>
4. Rozhkova M.A. Mashinnoe obuchenie kak metod primeneniya iskusstvennogo intellekta: sut' tekhnologii i obzor osnovnykh pravovykh problem // Zakon. ru. 2024. 18 marta. — URL: https://zakon.ru/blog/2024/03/18/mashinnoe_obuchenie_kak_metod_primeneniya_iskusstvennogo_intellekta_sut_tehnologii_i_obzor_osnovnyh_
5. Artificial Intelligence in Society, OECD Publishing, Paris, 2019. — URL: <https://www.oecd-ilibrary.org/docserver/eedfee77-en.pdf>
6. Naumov V.B., Tytyuk E.V. Pravovye problemy mashinnogo obucheniya // Obrazovanie i pravo. 2020. No. 6. P. 219–231. — URL: <https://cyberleninka.ru/article/n/pravovye-problemy-mashinnogo-obucheniya?ysclid=ltigehrqf2339477194>
7. Savelev A.I. Grazhdansko-pravovye aspekty regulirovaniya oborota dannykh v usloviyakh popytok formirovaniya tsifrovoy ekonomiki // Vestnik grazhdanskogo prava. 2020. No. 1. P. 67.
8. Rozhkova M.A. Nematerial'nye obekty i imushchestvennye prava na nikh v sisteme absolyutnykh prav. М.: GAUGN Press, 2024. 258 p. (seriya "IP & Digital Law").
9. Balobanova E.V. K voprosu o primenimosti doktriny dobrosovestnogo ispolzovaniya v kontinentalnom

- prave // Zhurnal Suda po intellektualnym pravam. 2017. No. 18. P. 92–99. — URL: <https://ipcmagazine.ru/articles/1729017/>
10. Vorozhevich A.S. Fair use i doktrina “rasshiryayushchego poleznost ispolzovaniya”. — URL: <https://legal-support.ru/information/blog/ip-daidzhest/fair-use-i-doktrina-rasshiraushhego-poleznost-ispolzovaniya/>.
 11. Lutkova O.V. Doktrina dobrosovestnogo ispolzovaniya proizvedeniy v sovremennom avtorskom prave SSHA // Pravo. Zhurnal Vyshey shkoly ekonomiki. 2016. No. 2. P. 186–199.
 12. Sagdeeva L.V. Svobodnoe ispolzovanie kak ogranichenie isklyuchitel'nykh prav // Yuridicheskie issledovaniya. 2017. No. 9.
 13. Mogilevskiy P.S. Doktrina dobrosovestnogo ispolzovaniya v SSHA i svobodnoe ispolzovanie proizvedeniy v Rossii: sravnitel'nyy analiz // Grazhdanskoe pravo. 2022. No. 5. P. 12–14.
 14. Kalyatin V.O. O perspektivakh primeneniya v Rossii doktriny dobrosovestnogo ispolzovaniya // Zakon. 2015. No. 11. P. 40–47.
 15. Chebodaeva V.A. Kontent, sgenirovannyy neyroset'yu: kto vinovat i chto delat'? // Zhurnal Suda po intellektualnym pravam. 2023. № 4 (42) (dekabr). P. 46–58. — URL: <https://ipcmagazine.ru/articles/1729306/>
 16. The creators of neural networks are accused of violating the copyrights of artists. 19 January 2023. — URL: <https://22century.ru/computer-it/113839?ysclid=m7wg2c00p0538107917>
 17. Shutkov A.A., Anishchenko A.N. Budushchee iskusstvennogo intellekta, neyrosetey i tsifrovyykh tekhnologiy v APK // Ekonomika i sotsium: sovremennye modeli razvitiya. 2019. T. 9. No. 4 (26). P. 508–522.
 18. Vasileva A.S. Ispolzovanie obektov avtorskiy prav dlya obucheniya iskusstvennogo intellekta // Zhurnal Suda po intellektualnym pravam. 2024. Dekabr. No. 4 (46). P. 80–84. — URL: <https://ipcmagazine.ru/articles/1765658/>.
 19. Doklad NIU VSHE “Problema mashinnogo tvorchestva v sisteme prava: regulirovanie sozdaniya i ispolzovaniya rezul'tatov intellektual'noy deyatel'nosti s primeneniem iskusstvennogo intellekta, zarubezhnyy opyt i rossiyskie perspektivy” / Ruk. avt. koll. V.O. Kalyatin. M.: ID VSE, 2021. 28 p. P. 22. — URL: <https://www.hse.ru/mirror/pubs/share/482492820.pdf>
 20. Kukhno M.O. Iskusstvennyy intellekt — novyy subekt avtorskiy prava: nedalekoe budushchee ili fiktsiya? // Zhurnal Suda po intellektualnym pravam. 2023. Vyp. 3 (41). P. 72–86. — URL: <https://ipcmagazine.ru/articles/1729283>
 21. Motovilova D.A. Klyuchevye polozheniya Direktivy (ES) No. 2019/790 ob avtorskom prave i smezhnykh pravakh na edinom tsifrovom rynke // Zhurnal Suda po intellektualnym pravam. 2019. No. 26 (dekabr). P. 106–120. — URL: <https://ipcmagazine.ru/articles/1729025/>
 22. Nazarov N.A. Mashinnoe tvorchestvo i pravo: dve chasti odnogo tselogo // Trudy po intellektual'noy sobstvennosti (Works on Intellectual Property). 2022. T. 43. No. 4. P. 101–110. — URL: <https://tis.hse.ru/article/view/16358/15428>