

Применение доктрины FairUse к случаям обучения генеративных нейросетей

А. В. Дюжакова

Санкт-Петербургский государственный университет,
Санкт-Петербург, Российская Федерация

В статье исследуется применение доктрины добросовестного использования (FairUse) к случаям обучения генеративных нейросетей. В связи со стремительным развитием генеративных нейросетей, которые обучаются на больших объемах данных, охраняемых авторским правом, возникает вопрос правомерности использования таких произведений для тренировки искусственного интеллекта. Вопрос имеет особую актуальность еще и потому, что за последний год неоднократно подавались коллективные иски авторов оригинальных произведений к компаниям-разработчикам генеративных нейросетей, с требованием признать такое использование незаконным.

Оценка четырех факторов, лежащих в основе доктрины FairUse, позволяет оценить, насколько в существующей парадигме использование произведений для обучения генеративных нейросетей может быть признано судом добросовестным, а также сформулировать предложение для системного изменения подхода к тренировке искусственного интеллекта в целом.

Ключевые слова: искусственный интеллект, генеративная нейросеть, интеллектуальная собственность, авторское право, доктрина добросовестного использования, fairuse.

Для цитирования: Дюжакова А. В. Применение доктрины FairUse к случаям обучения генеративных нейросетей // Ленинградский юридический журнал. – 2024. – № 1 (75). – С. 8–27. DOI: 10.35231/18136230_2024_1_8. EDN: RZDORB

Applying the Fair Use Doctrine to Generative Adversarial Network Training Cases

Anastasia V. Dyuzhakova

Saint-Petersburg State University,
Saint-Petersburg, Russian Federation

This article examines the application of the Fair Use doctrine to cases of training generative adversarial networks. In connection with the rapid extension of generative adversarial networks that are trained on large amounts of copyrighted data, the question arises of the legality of using such works for training artificial intelligence. The issue is of particular relevance also because over the past year, class action lawsuits by authors of original works have been repeatedly filed against companies developing generative adversarial networks, demanding that such use be declared illegal.

An assessment of the four factors underlying the Fair Use doctrine allows the author of the article to evaluate to what extent, in the existing paradigm, the use of works for training generative adversarial networks can be recognized by the court as fair, and also to formulate a proposal for a systematic change in the approach to training artificial intelligence generally.

Key words: artificial intelligence, generative neural network, intellectual property, copyright, fair use doctrine.

For citation: Dyuzhakova, A. V. (2024) Primenenie doktrini Fair Use k sluchayam obucheniya generativnih neirosetei [Applying the Fair Use Doctrine to Generative Adversarial Network Training Cases]. *Leningradskij yuridicheskij zhurnal – Leningrad Legal Journal*. No. 1 (75). Pp. 8–27. (In Russian). DOI: 10.35231/18136230_2024_1_8. EDN: RZDORB

Введение

В последние годы наблюдается настоящий бум по созданию и развитию инструментов искусственного интеллекта (далее – ИИ). Они стремительно набирают популярность и используются миллионами пользователей ежедневно, что делает их частью нашей реальности. На сегодняшний день наиболее прогрессивной моделью ИИ является генеративная нейросеть. Так, ее часто сравнивают с человеческим мозгом, работу которого она имитирует. Генеративные нейросети генерируют новые объекты из уже существующих данных при помощи подсказок и отличаются своей автономностью: они обрабатывают данные, находят закономерности и решают задачи (в том числе, на основе прошлого опыта) самостоятельно.

Однако для того чтобы генеративная нейросеть начала выдавать качественный результат, ее необходимо эффективно обучить. Данная модель ИИ учится на наборе различных материалов, включая научные статьи, книги, тексты на веб-сайтах и в социальных сетях, музыку, видеоматериалы, коды, изображения. Чем больше контента будет собрано для тренировки, тем точнее будут ответы, выдаваемые генеративной нейросетью. Колоссальный объем данных для обучения нейросетей берется из общего доступа – сети Интернет. Если такой контент был создан людьми, то он охраняется авторским правом. Отсюда возникает логичный вопрос – насколько правомерно используются эти произведения без согласия правообладателей в процессе обучения генеративных нейросетей? Для попытки ответа на данный вопрос предлагается проанализировать, подпадает ли такое использование под ограничения авторского права с точки зрения доктрины добросовестного использования – fair use.

Теоретический аспект данного исследования состоит в дальнейшем сравнительно-правовом анализе инсти-

тута разрешенного использования и доктрины добросовестного использования в правовом регулировании РФ и США в контексте обучения генеративных нейросетей. По сути, как эта доктрина, так и подход, используемый в континентальной системе права, касаются приблизительно одних и тех же случаев использования произведений [1]. С практической точки зрения исследование поможет потенциальным правоприменителям избежать правовых коллизий в схожих правоотношениях.

Понятие доктрины fairuse и анализ четырех факторов применительно к случаям обучения генеративных нейросетей

Добросовестное использование – это установленный законом стандарт, который уполномочивает суды определять, допустимо ли конкретное использование произведения, защищенного авторским правом, без лицензии [12]. Суть доктрины заключается в том, что третьи лица наделяются возможностью свободно использовать произведения, защищенные авторским правом, без согласия правообладателей и без ущемления их интересов при соблюдении определенных условий. Концепция fairuse нашла свое отражение в Законе США об «Авторском праве» 1976 г. (CopyrightAct)¹ (далее – Закон).

Легальное определение понятия «fairuse» в законодательстве США отсутствует, но в доктрине считается, что использование произведения будет добросовестным до тех пор, пока это способствует «прогрессу науки и полезных искусств»². О. В. Луткова отмечает: «Чтобы квалифицировать использование произведения как добросовестное, необходимо установить, как именно оно

¹ Кодекс США § 107 [Электронный ресурс]. URL: <https://www.law.cornell.edu/uscode/text/17/107> (дата обращения: 20.11.2023).

² Конституции США. П. 8, разд. 86, ст. 1. [Электронный ресурс]. URL: <https://constitutioncenter.org/the-constitution/full-text> (дата обращения: 20.11.2023).

способствует развитию знаний или прогрессу искусства, добавляя что-то новое» [3]. Является ли использование добросовестным в каждом конкретном случае, определяется судом с учетом исследования и взвешивания четырех факторов. Эти факторы не являются исключительными, но основными, а во многих случаях и единственными, которые анализируются судьями.

Применение доктрины fair use, по сути, представляет собой открытый стандарт, который наделяет судью правотворческими полномочиями и позволяет ему решать, какие факторы имеют превалирующее значение в том или ином деле, и соответственно определять наступление или ненаступление юридических последствий. Рассмотрим каждый из факторов применительно к допустимости использования произведений с целью обучения генеративных нейросетей без согласия правообладателей.

Цель и характер использования, коммерция или некоммерческие образовательные цели

Данный фактор еще называют «основным (или сердцем) четырехфакторного судебного исследования»¹. Во-первых, суд изучает вопрос, способствует ли рассматриваемое использование социально полезной деятельности, которая указана в законе: критике, комментариям, новостным репортажам, обучению, учебе или научным исследованиям. При анализе этого фактора суд также исследует вопрос, насколько использование произведения является разумной и общепринятой практикой, насколько добросовестно действовал пользователь, знал ли он о нарушении авторских прав. Хотя закон не ссылается на стандарт разумности и общепринятой практики, рассмотрение

¹ Дело On Davis. Gap, Inc., 246 F.3d 152, 2001. Компания GAP использовала в рекламе очки истца без его согласия. Суд посчитал, что такое использование не подпадает под добросовестное, так как цель не была преобразующей, использование носило коммерческий характер. [Электронный ресурс]. URL: <https://www.copyright.gov/fair-use/summaries/davis-gap-2dcir2001.pdf> (дата обращения: 13.12.2023).

этого вопроса уже давно является частью общего анализа суда на добросовестность использования [8].

Во-вторых, суд изучает вопрос, было ли преобразование исходного материала и в какой степени. Оценивая степень переработки, суд смотрит на то, добавляет ли преобразованная работа новое значение или иной характер, изменяя первоначальную новым выражением, значением или посланием [4]. При этом важное значение уделяется и цели: если цель создания и использования преобразованного произведения отличается от цели первоначального произведения, то с большей вероятностью такое использование будет оцениваться как добросовестное.

Применительно к случаям тренировки генеративного ИИ необходимо отметить его трансформационный характер. В первую очередь цель и характер использования произведений – преобразующие. Генеративная нейросеть не просто воспроизводит произведение, защищенное авторским правом, но вместо этого использует его для создания чего-то нового, т. е. преобразует исходные материалы и выдает оригинальный результат.

Обучаясь на массиве данных, нейросеть учится распознавать различные авторские техники, приемы, стилистические решения, чтобы впоследствии воплотить их в нечто новое. Если при создании работы с помощью генеративной нейросети копируется не первоначальное произведение, а лишь его стиль, то нельзя однозначно утверждать о нарушении авторских прав, ведь идеи и стиль не защищаются авторским правом. Это также благоприятствует признанию использования добросовестным.

В-третьих, суду необходимо определить, какой характер носит использование произведения: коммерческий или некоммерческий. По общему правилу, некоммерческое

использование не несет в себе цели извлечения прибыли и опять-таки с большей вероятностью будет признано добросовестным. Это не исключает того, что коммерческое использование в некоторых случаях не может быть признано добросовестным, однако требует более строгого анализа всех факторов.

Интересную позицию относительно того, что подразумевается под коммерческим использованием, высказал Верховный суд США при рассмотрении дела *Sony Corp. of America v. Universal City Studios*¹. Верховный суд США указал: «суть различия между коммерческим и некоммерческим использованием заключается не в том, является ли единственным мотивом денежная выгода, а в том, сможет ли пользователь получить прибыль от использования материала, защищенного авторским правом, без уплаты правообладателю вознаграждения». Иначе говоря, коммерческий характер использования определяется не только возможностью извлечения прибыли в момент использования произведения без согласия автора, а также возможностью его будущей монетизации.

Эта позиция Верховного суда США представляется весьма интересной в проекции использования произведений для обучения генеративных нейросетей. Ведь, вероятнее всего, первоочередные цель и характер использования произведений для тренировки ИИ – научные. Разработчики обучают генеративные нейросети для достижения новых научных результатов в области ИИ, а это, в свою очередь, социально полезная цель. Тем не менее, даже если в момент использования произведений, охраняемых авторским правом, основополагаю-

¹ Дело *Sony Corp. of America v. Universal City Studios*, 464 U.S. 417, 1984]. Sony Corp. косвенно обвинялась в том, что пользователи с помощью ее продукции осуществляли копирование телепередач. Суд посчитал, что такое копирование является добросовестным, так как пользователи продукции Sony осуществляли подобное копирование исключительно в личных целях (для домашнего просмотра). [Электронный ресурс]. URL: <https://www.copyright.gov/fair-use/summaries/sonycorp-universal-1984.pdf> (дата обращения: 29.11.2023).

щей целью была научная, это не исключает возможности монетизации созданной генеративной нейросети и ее использования в коммерческих целях в будущем.

Например, OpenAI – это некоммерческая организация, которая является «компанией, занимающейся исследованием и внедрением искусственного интеллекта», что в свою очередь является веским аргументом в пользу добросовестного использования. Тем не менее, компания OpenAI предлагает своим пользователям оформить платную подписку на пользование ChatGPT. А значит, генеративная нейросеть используется в коммерческих целях, что снижает шанс признания использования материалов, защищенных авторским правом, добросовестным. Аналогичная ситуация складывается вокруг генеративной нейросети Midjourney, которая помогает пользователям генерировать изображения по текстовому описанию. Несмотря на то что Midjourney позиционирует себя как самофинансируемый независимый проект, основной целью которого является расширение творческих способностей человечества, платформа предлагает пользователям платные подписки, более того, на платных тарифах Midjourney сгенерированные картинки можно использовать в коммерческих целях.

Такую практику еще называют «отмывание данных искусственного интеллекта», защищающую коммерческие организации от ответственности за счет использования некоммерческого исследовательского учреждения для создания набора данных и обучения механизмов ИИ, которые впоследствии могут быть использованы в коммерческих приложениях [13].

Таким образом, суд может признать использование произведений для обучения генеративных нейросетей коммерческим, если владельцы ИИ будут взимать с конечных пользователей деньги, размещать рекламу

на веб-сайте/приложении нейросети или иным образом получать прибыль от использования модели ИИ.

Сущность произведения, охраняемого авторским правом

При исследовании данного фактора суд в первую очередь оценивает творческий характер используемого произведения, его креативность. В судебной практике основным способом влияния на уровень предъявляемых требований к творческому характеру произведений является, как правило, применение определенных критериев охраноспособности: оригинальности, новизны и уникальности, с одной стороны, и создания произведения в результате самостоятельной созидательной деятельности – с другой. Считается, что в последнем случае стандарт охраноспособности ниже [2].

С точки зрения доктрины fairuse именно высокотворческие традиционные произведения (речь идет о художественных произведениях) охраняются в большей степени нежели произведения с незначительной творческой составляющей (например, фотографии, программы для ЭВМ). Использование произведений с высоким уровнем творчества снижает шансы признания использования такого объекта добросовестным.

Также суд оценивает, не ущемляет ли использование произведения его сущность и ценность. В этом аспекте интересно дело *Campbell v. Acuff-Rose Music, Inc*¹. Верховный суд США подчеркнул, что «необходимо изучить все факторы, а результаты взвесить комплексно», «судам необходимо избегать дословного и жесткого приме-

¹ Дело *Campbell v. Acuff-Rose Music, Inc.*, 510 U.S. 569, 1994. Группа 2 Live Crew использовала фрагменты песни "Oh, Pretty Woman" Роя Орбисона в своей песне "Pretty Woman". Acuff-Rose Music, владельцы авторских прав на оригинальную песню, подали иск на 2 Live Crew за нарушение авторских прав. Однако суд признал, что использование фрагментов было допустимым по доктрине fairuse. Это решение стало прецедентом для использования фрагментов в качестве пародий и комментариев в США. [Электронный ресурс]. URL: <https://www.copyright.gov/fair-use/summaries/campbell-acuff-1994.pdf> (дата обращения: 27.11.2023).

нения закона, когда это может подавить саму креативность, которую этот закон призван поощрять». Более того, «чем более преобразующей будет новая работа, тем меньше будет значимость других факторов, которые могут препятствовать установлению добросовестного использования, например, таких как коммерческий характер использования».

Суд пришел к выводу, что использование фрагментов в пародийной песне ответчика не ущемляло сущность и ценность оригинального произведения, потому что ответчик использовал эти фрагменты в качестве материала для создания новой, оригинальной и уникальной песни. Он не просто скопировал или повторил оригинал, а использовал его элементы в качестве основы для своей пародии. А коммерческое использование и определенная степень экономического вреда автоматически не опровергают аргументы в пользу добросовестного использования.

Представляется, что в контексте применения этого фактора к обучению генеративных нейросетей, он также будет иметь важное значение. Ведь генеративные нейросети трансформируют входные материалы и выдают преобразованный контент. Такое использование не влияет на сущность оригинального произведения, так как оно не теряет своей ценности, значимости, качества.

Величина и существенность использованной части по отношению ко всему произведению, защищенному авторским правом

При анализе данного фактора суд исследует, какие части произведения использовались, их объем, как они использовались, как эти части соотносятся с произведением в целом, а также, какой объем был необходим для достижения поставленной цели. Иными словами, объем и ча-

сти копируемого произведения должны быть разумными с точки зрения цели копирования. Для этого судом применяется как количественный, так и качественный подход к оценке суммы и существенности использованной части. Это значит, что суд может исследовать, какой процент оригинального произведения использовался, или же задать вопросом, насколько с качественной точки зрения использованная часть важна для произведения в целом.

При применении количественного подхода отсутствует какой-либо единообразный критерий к тому, какой процентный показатель использования чужого произведения считается добросовестным использованием, а какой нет. Например, в деле *Cambridge Univ. Press v. Becker*¹ суд первой инстанции установил, что для некоммерческого образовательного использования копирование до 10 процентов или одной главы книги будет представлять собой добросовестное использование. В ходе апелляции этот вывод суда подвергся критике и апелляционный суд постановил, что суд нижестоящей инстанции «ошибся, установив контрольный показатель в 10 процентов или одну главу», поскольку никакой фиксированный контрольный показатель вообще не должен устанавливаться, т. е. в каждом конкретном деле суд должен придерживаться индивидуального подхода, исследуя все факторы в совокупности. Именно суды определяют объем разрешенного использования материалов, защищенных авторским правом, в каждом конкретном деле.

При применении качественного подхода важен не процентный показатель, а то, насколько используемая часть важна для произведения в целом, например, представляет ли она собой самую запоминающуюся и выдающуюся

¹ Дело *Cambridge Univ. Press. Becker*, 863 F. Supp. 2d 1190, 2012. Ответчик оцифровал книги и разместил их в систему университета, предоставив студентам доступ к ним. Суд посчитал такое использование добросовестным (в отношении большей части книг). [Электронный ресурс]. URL: <https://www.copyright.gov/fair-use/summaries/cambridgeuniv-patton-11thcir2014.pdf> (дата обращения: 27.11.2023).

часть произведения. В таком случае копирование даже небольшого объема работы может быть расценено как недобросовестное использование. Например, в деле *Harper & Row Publishers, Inc. v. Nation Enterprises*¹ ответчик (журнал *The Nation*) опубликовал дословные цитаты объемом всего лишь в 300–400 слов из готовящихся к выпуску мемуаров Джеральда Форда. Верховный суд США постановил, что третий фактор препятствует добросовестному использованию, поскольку выдержки включали обсуждение Фордом вопроса о помиловании Никсона и другие центральные фрагменты книги, которые суд счел «сердцем» работы.

Иными словами, нельзя процитировать кульминационную часть романа или сделать нарезку из самых ярких частей кинофильма. При качественном подходе правило «копирование большего объема работ в большей степени влечет признание использования недобросовестным» не работает. Важно не использовать именно самые креативные и яркие части произведения (хоть и в малом объеме).

Применяя этот фактор к случаям обучения генеративных нейросетей, необходимо исходить из технических характеристик тренируемой нейросети. Логистика большей части создания и обучения генеративных нейросетей предполагает полное копирование множества произведений, защищенных авторским правом. Иными словами, набор обучающих данных должен содержать произведения в полном объеме. Такое копирование произведений в наборы обучающих данных разумно с точки зрения целей машинного обучения, которое требует анализа целых произведений для изучения целевых шаблонов.

¹ Дело *Harper & Row Publishers, Inc. v. Nation Enterprises*, 471 U.S. 539 (1985) Ответчик процитировал часть из предстоящих к публикации мемуаров. Суд посчитал, что такое использование не является добросовестным, так как: (а) мемуары еще не были опубликованы, (б) выдержки хоть и составляли незначительную часть мемуаров, однако включали в себя самые выразительные и ключевые части; (в) использование напрямую конкурировало за долю рынка и привело к реальному ущербу. [Электронный ресурс]. URL: <https://www.copyright.gov/fair-use/summaries/harperrrow-nationmagazine-1985.pdf> (дата обращения: 29.11.2023).

Говоря о копировании, необходимо сказать, что в общем доступе нет информации о том, копируется ли в процессе обучения этот контент или на него просто даются ссылки. Для ответа на вопрос, загружается ли набор обучающих данных в систему, мы спросили саму генеративную нейросеть – ChatGPT. Вот какой ответ нами был получен: «Как искусственный интеллект, я не имею доступа к информации о том, каким образом были собраны данные для обучения ChatGPT. Однако, если данные были собраны с использованием доктрины fairuse, то это не является нарушением авторских прав» (дата обращения: 25.01.2024).

В свете вышеописанного очень важно учитывать: (1) имеет ли место прямое копирование произведений (их загрузка в систему); (2) какой объем дословного цитирования генеративная нейросеть выдает конечному пользователю, то есть становится ли эта часть обучающих данных дословно доступной пользователям; (3) включают ли в себя выходные данные самые узнаваемые, творческие части элементов обучающих произведений.

Если конечный пользователь не получает доступ к оригинальному произведению, загруженному в генеративную нейросеть с целью обучения последней, и сгенерированные данные не содержат в себе самые выразительные элементы оригинального произведения, то, представляется, что такое использование можно назвать добросовестным.

Влияние использования на потенциальный рынок или стоимость произведения, защищенного авторским правом

Это фактор позволяет оценить влияние использования произведения с экономической точки зрения: на (1) потенциальный рынок или (2) стоимость произ-

ведения, защищенного авторским правом, и его производных работ. Использование, которое направлено на «замену» оригинального произведения и конкуренцию за его рынок, с меньшей вероятностью будет считаться добросовестным (например, если такое использование приведет к потере продаж оригинального произведения).

С одной стороны, представляется, что использование объектов, защищенных авторским правом, при обучении генеративных нейросетей носит минимальное влияние на потенциальный рынок или стоимость используемого произведения, поскольку ИИ не конкурирует с оригинальным произведением. При таком подходе использование произведения вряд ли повредит возможностям правообладателя извлекать прибыль из использования своего произведения. Кроме того, из-за огромного объема текста, используемого для обучения генеративной нейросети, маловероятно, что выходные данные, которые она выдает, окажут негативное экономическое воздействие на какого-либо отдельного правообладателя.

С другой стороны, на рынке наблюдается достаточно высокий спрос на работы, генерируемые нейросетями, будь то текстовые, музыкальные или художественные результаты. Генеративная нейросеть создает объекты дешевле и быстрее. При такой тенденции может произойти общий спад спроса на произведения, создаваемые людьми. Если генеративная нейросеть начнет конкурировать с людьми-творцами, такая борьба может обернуться не в пользу человека. С данной точки зрения, продукты генеративного ИИ могут отрицательно повлиять на рынок, стоимость, спрос и ценность защищенного авторским правом материала, на котором обучают генеративные нейросети.

Анализ актуальной судебной практики

Рассматривая четыре фактора доктрины через призму использования произведений с целью тренировки ИИ, нельзя не отметить, что актуальность данного вопроса обусловлена не только пробелом в нормативном регулировании, что за последний год неоднократно подавались искивые заявления против компаний-разработчиков нейросетей (например, против OpenAI – создателя генеративной нейросети ChatGPT и др.)¹. Во всех случаях истцы настаивают на том, что использование их произведений для обучения ИИ не является «добросовестным» с точки зрения доктрины fairuse.

Компания OpenAI в свою очередь уже дала официальный комментарий по защите интеллектуальной собственности на инновации в области ИИ²: «Миссия компании состоит в том, чтобы гарантировать, что искусственный интеллект принесет пользу всему человечеству», «современные системы искусственного интеллекта требуют больших объемов данных для их обучения. Только анализируя большие массивы данных, системы искусственного интеллекта могут изучать закономерности, присущие данным, созданным человеком, а затем использовать эти закономерности для синтеза аналогичных данных»,

¹ Иск газеты New York Times против OpenAI [Электронный ресурс]. URL: https://nytc0-assets.nytimes.com/2023/12/NYT_Complaint_Dec2023.pdf (дата обращения: 02.02.2024); Иск писателей Пола Трэмблэй и Моны Эвад против OpenAI [Электронный ресурс]. URL: <https://lmlitigation.com/pdf/03223/tremblay-openai-complaint.pdf> (дата обращения: 02.02.2024); Иск авторов Сары Сильверман, Криса Голдена и Ричарда Кадри против OpenAI [Электронный ресурс]. URL: <https://lmlitigation.com/pdf/03416/silverman-openai-complaint.pdf> (дата обращения: 02.02.2024); Иск авторов Джорджа Р. Р. Мартина, Майкла Коннелли, Джонатана Франзена, Джона Гришма и других против OpenAI и Microsoft. [Электронный ресурс]. URL: <https://authorsguild.org/app/uploads/2023/12/Authors-Guild-OpenAI-Microsoft-Class-Action-Complaint-Dec-2023.pdf> (дата обращения: 02.02.2024); Иск программистов против GitHub, Microsoft и Open AI [Электронный ресурс]. URL: <https://storage.courtlistener.com/recap/gov.uscourts.cand.403220/gov.uscourts.cand.403220.1.0.pdf> (дата обращения: 02.02.2024); Иск группы художников против Stability AI, Midjourney и DeviantArt [Электронный ресурс]. URL: <https://docs.justia.com/cases/federal/district-courts/california/cand-ce/3:2023cv00201/407208/67> (дата обращения: 02.02.2024). 30 октября 2023 суд отклонил все иски, кроме одного – о прямом нарушении авторских прав со стороны Stability AI.

² Официальный комментарий по защите интеллектуальной собственности на инновации в области искусственного интеллекта [Электронный ресурс]. Регистрационный № ПТО-С-2019-0038. URL: https://www.uspto.gov/sites/default/files/documents/OpenAI_RFC-84-FR-58141.pdf (дата обращения: 02.02.2024).

включение материалов, защищенных авторским правом, в наборы данных для машинного обучения является добросовестным использованием, поскольку это «невыразительное промежуточное копирование». Согласно OpenAI, цель и характер использования произведений для обучения генеративной нейросети являются преобразующими – «машинное обучение имеет целью изучение моделей, присущих созданным человеком».

Фактическое исследование суда в каждом случае будет касаться взаимосвязи между выходными данными и обучающими данными, защищенными авторским правом. На данный момент решения по исковым заявлениям еще не вынесены, однако можно с уверенностью сказать, что все эти дела служат своеобразным лакмусовым тестом жизнеспособности исков о нарушении авторских прав при обучении генеративных нейросетей.

Заключение

Обучение и использование генеративных моделей ИИ безусловно меняет наше представление о фундаментальных вопросах защиты авторских прав. Рассмотрев все четыре фактора применительно к случаям использования произведений для тренировки генеративных нейросетей, мы можем сделать следующие выводы. Во-первых, не все факторы должны склоняться в ту или иную сторону, т. е. за или против добросовестного использования произведения. Во-вторых, относительная важность каждого из четырех исследуемых судом факторов не всегда одинакова. Обучая искусственный интеллект на больших массивах данных, защищенных авторским правом, нельзя однозначно сказать, подпадает ли такое использование под добросовестное. Только результат анализа всех факторов в совокупности и значения каждого из них может привести суд к выводу о том, насколько добросовестным

является такое использование в том или ином конкретном случае.

Такой подход представляется достаточно гибким, так как позволяет быстро адаптировать закон к стремительно меняющимся технологиям и инновационным потребностям общества. Цель применения доктрины к случаям тренировки генеративной нейросети состоит в том, чтобы найти баланс между выгодой, которую получит владелец авторских прав, если копирование будет признано «недобросовестным использованием», и выгодой, которую получит общественность, если использование будет признано добросовестным.

Обратной стороной медали такого гибкого подхода является некоторая правовая неопределенность. Применение четырех факторов добросовестного использования предполагает сложный всесторонний анализ и дает судьям настолько широкие полномочия при их толковании, что может привести к непредсказуемым результатам. И, таким образом, не дает достаточных указаний на то, как доктрина будет адаптирована судами к случаям обучения генеративных моделей ИИ.

Проанализировав описанные выше факторы, полагаем, что с большей уверенностью судить о добросовестности использования можно будет при системном изменении самого подхода к обучению генеративных нейросетей. Для этого в процессе обучения предлагается соблюдать совокупность следующих правил:

1. Генеративная нейросеть должна производить анализ обучающих произведений из официальных источников, т. е. без создания их материальных копий (если технически это возможно).
2. Сохраняется недоступность оригинального произведения конечному пользователю и соблюдается его каче-

ственная охрана (в случае копирования оригинального произведения).

3. Генеративную нейросеть обучают копировать стиль, а не произведение (т. е. так, чтобы в сгенерированных данных не содержались дословные цитаты оригинальных произведений или их самые выразительные части/элементы, а при необходимости включали в себя ссылки на такие произведения).

Подводя итоги, стоит отметить, что судебные дела, оспаривающие использование контента, защищенного авторским правом, для обучения технологий ИИ, хорошо иллюстрируют конфликты, возникающие между авторским правом и ИИ, а также связанные с этим сложности. ИИ должен разрабатываться ответственно и этично, а разработчики должны уважать права художников и владельцев авторских прав.

Список литературы

1. Калятин В. О. О перспективах применения в России доктрины добросовестного использования // Закон. – 2015. – № 11. – С. 40–47.
2. Кашанин А. В. Актуальные требования к творческому характеру произведений в российской доктрине и судебной практике // Законы России: опыт, анализ, практика. – 2016. – № 7. – С. 50–59.
3. Луткова О. В. Доктрина добросовестного использования произведений в современном авторском праве США // Право. Журнал Высшей школы экономики. – 2016. – № 2. – С. 186–199.
4. Могилевский П. С. Доктрина добросовестного использования в США и свободное использование произведений в России: сравнительный анализ // Гражданское право. – 2022. – № 5. – С. 12–14.
5. Falzone A., Urban J. Demystifying Fair Use: The Gift of the Center for Social Media Statements of Best Practices // Journal of the Copyright Society of the USA. – 2010. – Vol. 57. – No. 337. – Pp. 337–350.
6. Depoorter B. Technology and Uncertainty: The Shaping Effect on Copyright Law // University of Pennsylvania Law Review. – 2009. – Vol. 157. – Pp. 1831–1868.
7. Parchomovsky G., Goldman K. A. Fair Use Harbors // Virginia Law Review. – 2007. – Vol. 93. – Pp. 1483–1532.
8. Rothman J. E. Copyright, Custom and Lessons from the Common Law // Intellectual Property and the Common Law. – 2013. – Pp. 1–19.

9. Hughes J. Fair Use and Its Politics – at Home and Abroad // Okediji RL, ed. *Copyright Law in an Age of Limitations and Exceptions*. Cambridge University Press. – 2017. – Pp. 234–274.
10. Sag M. Predicting Fair Use // *Ohio State Law Journal*. – 2012. – Vol. 73:1. – Pp. 47–91.
11. Netanel N. Making Sense of Fair Use // *Lewis & Clark Law Review*. – 2011. – Vol. 15 – Pp. 715–771.
12. Elkin-Koren N., Fischman Afori O. Rulifying Fair Use // *Arizona Law Review*. – 2017. – Vol. 59. – Pp. 1–34.
13. Radanliev P., Santos O. Ethics and Responsible AI Deployment // *Frontiers in Artificial Intelligence*. – 2023. – Vol. 7. – Pp. 1–24.
14. Leval P. Toward a Fair Use Standard // *Harvard Law Review*. – 1990. Vol. 103 (5). – Pp. 1057–1094.

References

1. Kalyatin, V. O. (2015) O perspektivah primeneniya v Rossii doctriini dobrosovestnogo ispolzovaniya. *Zakon – Law*. 2015. No. 11. Pp. 40–47. (In Russian).
2. Kashanin, A. V. (2016) Aktualnyie trebovaniya k tvorcheskomu karakteru proizvedeniy v rossiyskoy doctriine I sudebnoy praktike [Current requirements for the creative nature of works in Russian doctrine and judicial practice]. *Zakonyi Rossii: opyt, analiz, praktika – Laws of Russia: experience, analysis, practice*. No. 7. Pp. 50–59. (In Russian).
3. Lutkova, O. V. (2016) Doktrina dobrosovestnogo ispolzovaniya proizvedeniy v sovremennom avtorskom prave SShA [The Doctrine of Fair Use of Works in Modern U. S. Copyright Law]. *Pravo. Zhurnal Vyisshey shkoly iekonomiki – Right. Journal of the Higher School of Economics*. No. 2. Pp. 186–199. (In Russian).
4. Mogilevskiy, P. S. (2022) Doktrina dobrosovestnogo ispolzovaniya v SShA I svobodnoye ispolzovaniye proizvedeniy v Rossii: sravnitelnyy analiz [The Doctrine of Fair Use in the United States and the free use of works in Russia: comparative analysis]. *Grazhdanskoyepravo – Civil law*. No. 5. P. 12. Pp. 12–14. (In Russian).
5. Falzone, A., Urban, J. (2010) Demystifying Fair Use: The Gift of the Center for Social Media Statements of Best Practices. *Journal of the Copyright Society of the USA*. Vol. 57. No. 337. Pp. 337–350.
6. Depoorter, B. (2009) Technology and Uncertainty: The Shaping Effect on Copyright Law. *University of Pennsylvania Law Review*. Vol. 157. Pp. 1831–1868.
7. Parchomovsky, G., Goldman, K. A. (2007) Fair Use Harbors. *Virginia Law Review*. Vol. 93. Pp. 1483–1532.
8. Rothman, J. E. (2013) Copyright, Custom and Lessons from the Common Law. *Intellectual Property and the Common Law*. Pp. 1–19.
9. Hughes, J. (2017) Fair Use and Its Politics – at Home and Abroad. Okediji RL, ed. *Copyright Law in an Age of Limitations and Exceptions*. Cambridge University Press. Pp. 234–274.
10. Sag, M. (2012) Predicting Fair Use. *Ohio State Law Journal*. Vol. 73:1. Pp. 47–91.
11. Netanel, N. (2011) Making Sense of Fair Use. *Lewis & Clark Law Review*. Vol. 15. Pp. 715–771.
12. Elkin-Koren, N., Fischman Afori, O. (2017) Rulifying Fair Use. *Arizona Law Review*. Vol. 59. Pp. 1–34.
13. Radanliev, P., Santos, O. (2023) Ethics and Responsible AI Deployment. *Frontiers in Artificial Intelligence*. Vol. 7. Pp. 1–24.

14. Leval, P. (1990) Toward a Fair Use Standard. *Harvard Law Review*. Vol. 103 (5). Pp. 1057–1094.

Об авторе

Дюжакова Анастасия Вадимовна, аспирант, Санкт-Петербургский государственный университет, Санкт-Петербург, Российская Федерация, ORCID ID: 0009–0002–3331–2556, e-mail: anastasia.duzhakova@gmail.com

About the author

Anastasia V. Dyuzhakova, Postgraduate student, Saint-Petersburg State University, Saint Petersburg, Russian Federation, ORCID ID: 0009–0002–3331–2556, e-mail: anastasia.duzhakova@gmail.com

Поступила в редакцию: 29.01.2024
Принята к публикации: 13.02.2024
Опубликована: 30.03.2024

Received: 29 January 2024
Accepted: 13 February 2024
Published: 30 March 2024