

К ВОПРОСУ НАРУШЕНИЯ АВТОРСКИХ ПРАВ ПРИ ИСПОЛЬЗОВАНИИ ТЕХНОЛОГИЙ ИСКУССТВЕННОГО ИНТЕЛЛЕКТА

О.О. Ядревский

кандидат юридических наук, доцент,
доцент кафедры права интеллектуальной собственности юридического факультета
Белорусского государственного университета
yadreus@bsu.by

Аннотация. В статье рассматриваются случаи нарушения авторского права при обучении систем искусственного интеллекта. Анализируется зарубежное регулирование указанного вопроса в рамках Европейского Союза, США, Японии. Рассматриваются положения белорусского законодательства, даются рекомендации по разрешению возможных проблемных ситуаций в национальной юрисдикции.

Ключевые слова: интеллектуальная собственность; авторское право; искусственный интеллект; правообладатели; добросовестное использование.

Abstract. The article discusses cases of copyright violation when training artificial intelligence systems. Foreign regulation of this issue within the European Union, USA, and Japan is analyzed. The provisions of Belarusian legislation are considered and recommendations are given for resolving possible problematic situations in national jurisdiction.

Keywords: intellectual property, copyright, artificial intelligence, copyright holder fair use.

Развитие технологий искусственного интеллекта неизбежно влечет возникновение не только этических, социальных и иных проблем, но и вопросов юридического характера, требующих разрешения. Так, с середины 2023 г. против компании OpenAI – разработчика системы Chat GPT – и ее основного инвестора Microsoft было возбуждено более десятка громких судебных процессов и правительственных расследований в разных странах. Несмотря на некоторые различия в предметах указанных исков, основу их составляет то, что компания нарушает закон об авторском праве, обучая алгоритмы искусственного интеллекта [1].

Важность указанных процессов и их результатов заключается в том, что искусственный интеллект обучается на данных не только крупных компаний-правообладателей, но и миллионов простых пользователей интернета. Они тоже вправе потребовать справедливой компенсации за использование результатов своей интеллектуальной деятельности.

Разработчики машинного обучения обосновывают использование объектов авторского права тем, что такое обучение носит преобразующий и ограниченный характер, то есть является добросовестным. Так, ситуацию сравнивают со сканированием книг, защищенных авторским правом, компанией Google Books в деле Гильдии авторов против компании Google, которое в конечном итоге было признано добросовестным использованием, поскольку скопированный контент не был открыт для публичного доступа и его использование носило локальный характер.

Тимоти Б. Ли указывает, что, если истцы добьются успеха, это может сместить баланс сил в пользу крупных корпораций, таких как Google, Microsoft и Meta, которые могут позволить себе лицензировать большие объёмы обучающих данных у правообладателей и использовать свои собственные наборы данных, генерируемых пользователями [2].

Как утверждают исследователи интеллектуальной собственности Брайан Кейси и Марк Лемли, наборы данных настолько велики, что «не существует разумного варианта просто лицензировать все (данные) разрешение (любого генеративного обучения) требовать авторских прав равносильно утверждению: дело не в том, что владельцам авторских прав будут платить, а в том, что использование вообще не будет разрешено» [3].

Следует отметить, что указанные иски подаются в связи со следующими причинами. Системы искусственного интеллекта часто используются в коммерческих целях и обладают высокой эффективностью производства, что увеличивает объем и частоту получаемой прибыли. При этом субъект, изначально владеющий данными, не получает экономической выгоды. Например, об этом говорится в коллективном иске к Stability AI Ltd. и Stability AI, Inc., Midjourney, Inc., DeviantArt, Inc. [4]

Кроме того, серьезным является и вопрос репутации, поскольку качество получаемого искусственным интеллектом продукта не всегда является высоким, при этом в продукте могут использоваться водяные знаки и иные способы фиксации авторства используемых продуктов. В качестве примера можно привести иск Getty Images (US), Inc. v. Stability AI, Inc. [5].

Вместе с тем, противоположной крайностью является ограничение доступа к данным, фактически монополизация обладания данными и установления кабальных договорных условий.

В 2019 году была принята Директива (ЕС) 2019/790 Европейского парламента и Совета от 17 апреля 2019 г. об авторском праве и смежных пра-

вах в рамках единого цифрового рынка, дополняющая Директивы 96/9/ЕС и 2001/29/ЕС

Статья 3 Директивы (ЕС) N 2019/790 возлагает на государства-члены ЕС установить в национальном законодательстве обязательное исключение в отношении воспроизведения и извлечения для целей интеллектуального анализа текстов и данных в целях проведения научных исследований. Данным исключением могут пользоваться только научно-исследовательские организации и учреждения культурного наследия и только в отношении произведений или иных охраняемых объектов, к которым они имеют законный доступ. При этом в п. 14 преамбулы Директивы разъяснено понятие законного доступа. Доступ признается законным, если исследуемый контент имеется в открытом доступе (свободный доступ в сети Интернет) либо на основе договорных обязательств между правообладателями и научно-исследовательскими организациями или учреждениями культурного наследия. Норма п. 2 ст. 3 Директивы, разрешающая сохранение копий произведений или иных охраняемых объектов для целей научных исследований, в том числе для проверки результатов исследований, также является весьма важной.

Статья 4 Директивы разрешает воспроизведение и извлечение для целей интеллектуального анализа текстов и данных более широким кругом субъектов, включая и коммерческие организации. 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 в сети Интернет, и, если правообладатели не сохранили за собой надлежащим образом права на воспроизведение и извлечение для целей интеллектуального анализа текстов и данных (ст. 4(3) и абз. 2 п. 18 преамбулы). При этом в случае если охраняемый контент доведен до всеобщего сведения в сети Интернет, надлежащим признается использование машинночитаемых средств, включая метаданные и условия использования веб-сайта или услуги.

В случае, если правообладатель не желает использования своих произведений, он может запретить его использование для указанных целей путем односторонней декларации либо вступления в договорные соглашения.

Таким образом, отличия статей 3 и 4 Директивы заключаются в том, что в первом случае допускается безусловное использование произведений

и иных охраняемых объектов для целей интеллектуального анализа данных и текстов научно-исследовательскими организациями и учреждениями культурного наследия, во втором случае такое свободное использование другими субъектами может быть запрещено правообладателями охраняемого контента.

Данные положения имеют важное значение. С одной стороны, они гарантируют право доступа к охраняемым материалам, в том числе к тем, которые хранятся в чужих базах данных. С другой стороны, они создают условия для формирования научных баз данных и их последующего использования.

В качестве примера других юрисдикций, где введены ограничения исключительного права в целях использования объектов авторского права для обучения систем искусственного интеллекта является Япония.

Так, ст. 30-4 Закона Японии об авторском праве разрешает анализ и использование охраняемых произведений в целях машинного обучения и анализа данных; ст. 47-4 разрешает создание случайных копий произведений, создаваемых в процессе работы компьютера; ст. 47-5 вводит возможность осуществления незначительного использования в рамках компьютеризированной обработки данных и предоставления результатов такой обработки, что важно для исследований, проверки и сохранения результатов, основанных на работе с базами данных.

В Великобритании предложено изменить положение о запрете интеллектуального анализа данных в коммерческих целях, разрешив такой анализ для поддержки развития искусственного интеллекта: Правообладатели по-прежнему будут иметь гарантии защиты своего контента, включая требование законного доступа [6].

В литературе по этому вопросу высказывается мнение, что «только инструменты интеллектуального анализа данных, предполагающие минимальное копирование нескольких слов или сканирование данных и обработку каждого элемента в отдельности, могут использоваться без потенциальной ответственности за нарушение авторских прав» [7].

Для определения возможности использования объектов интеллектуальной собственности в целях майнинга данных в США применяется концепция *fair use* – добросовестного использования, которая позволяет свободно использовать объекты авторского права, если такое использование является «разумным». Это дает возможность не указывать в законодатель-

стве исчерпывающий перечень ограничений, а также делать изъятия даже в отношении обычно допускаемых действий (ограничивать сферу свободного использования произведений) в зависимости от потребностей бизнеса. Так, согласно статье 17 U.S.C.S. § 107 устанавливаются четыре фактора, которые необходимо учитывать при определении того, является ли использование в конкретном случае добросовестным: – цель и характер использования (коммерческое оно или некоммерческое (образовательное)); – характер произведения, защищенного авторским правом (какого оно рода); – объем и существенность используемой части (какая часть произведения используется, насколько она важна для произведения в целом); – влияние использования на потенциальный рынок или ценность произведения (как использование влияет на спрос на произведение или его стоимость).

Каждый из указанных факторов должен взвешиваться судом в свете целей закона об авторском праве, а не рассматриваться как исчерпывающие или определяющие факторы.

Следует отметить, что в белорусском законодательстве об авторском праве отсутствует изъятие, подобное ранее рассмотренным, для интеллектуального анализа данных. Из наиболее похожих можно указать ст. 36 Закона Республики Беларусь от 17 мая 2011 г. № 262-З «Об авторском праве и смежных правах», предполагающую свободное использование произведений в образовательных и исследовательских целях, п. 2 которой гласит следующее: «Статьи и иные малообъемные произведения, правомерно опубликованные в сборниках, а также газетах, журналах и других печатных средствах массовой информации, отрывки из правомерно опубликованных литературных и иных произведений могут быть воспроизведены посредством репродуцирования и иного воспроизведения в образовательных и исследовательских целях».

В то же время указанная статья носит ограничительный характер, поскольку допускает лишь воспроизведение с определенными целями.

Представляется целесообразным на первом этапе применять в судебной практике доктрину добросовестного использования, при этом извлечение данных должно производиться для подготовки интеллектуального анализа данных или в процессе такого анализа, в некоммерческих целях – научных, образовательных, личных. Это способствовало бы развитию технологий искусственного интеллекта и обеспечивало бы баланс интересов разработчиков таких технологий и правообладателей.

Кроме того, полагаем необходимым обязать создателей моделей искусственного интеллекта предоставлять подробную информацию о контенте, используемом для обучения модели искусственного интеллекта, а также о сущности технологии, целях ее разработки и сферах использования, чтобы обеспечить доверие пользователей.

Необходимо также ввести возможность для правообладателя контента ввести запрет на использование его произведений для целей обучения искусственного интеллекта путем публичного оповещения либо конкретного письменного предупреждения, адресованного разработчикам систем искусственного интеллекта.

Список использованных источников

1. Open AI заплатит за нарушение авторских прав и «галлюцинации» Chat GPT? // Хабр. – URL: <https://habr.com/ru/articles/815575/> (дата обращения: 22.05.2024).
2. Stable Diffusion copyright lawsuits could be a legal earthquake for AI // Arstechnica. – URL: <https://arstechnica.com/tech-policy/2023/04/stable-diffusion-copyright-lawsuits-could-be-a-legal-earthquake-for-ai/> (дата обращения: 22.05.2024).
3. The scary truth about AI copyright is nobody knows what will happen next // The verge. – URL: <https://www.theverge.com/23444685/generative-ai-copyright-infringement-legal-fair-use-training-data> (дата обращения: 22.05.2024).
4. Andersen, V. Stability AI Ltd. (3:23-cv-00201) / V. Andersen // CourtListener. – URL: <https://www.courtlistener.com/docket/66732129/andersen-v-stability-ai-ltd/> (дата обращения: 22.05.2024).
5. Getty Images (US), Inc. v. Stability AI, Inc. (1:23-cv-00135) // Copyrightlately. – URL: https://copyrightlately.com/pdfviewer/getty-images-v-stability-ai-complaint/?auto_viewer=true#page=1&zoom=auto,-245,800 (дата обращения: 22.05.2024).
6. Artificial Intelligence and Intellectual Property: copyright and patents: Government response to consultation // UK Government (2023). – URL: <https://www.gov.uk/government/consultations/artificial-intelligence-and-ip-copyright-and-patents/outcome/artificial-intelligence-and-intellectual-property-copyright-and-patents-government-response-to-consultation> (дата обращения: 22.05.2024).
7. Text and Data Mining in the Proposed Copyright Reform: Making the EU Ready for an Age of Big Data. – URL: <https://www.ip.mpg.de/en/publications/details/text-and-data-mining-in-the-proposed-copyright-reform-making-the-eu-ready-for-an-age-of-big-data-legal-analysis-and-policy-recommendations.html> (дата обращения: 22.05.2024).