Ирина Ивановна КАРПЕНКО

ГЕНЕРАТИВНЫЕ ТЕХНОЛОГИИ ИСКУССТВЕННОГО ИНТЕЛЛЕКТА В СОВРЕМЕННЫХ МЕДИА: ПРОБЛЕМЫ АВТОРСКОГО ПРАВА И МАРКИРОВАНИЯ НЕЙРОСЕТЕВОГО ПРОДУКТА

Развитие технологий искусственного интеллекта открывает перед современными медиа большие перспективы. Однако сегодня, наряду с неоспоримыми положительными моментами использования традиционных и генеративных технологий ИИ, остро встает ряд проблем, связанных с соблюдением авторского права. В статье рассматриваются законодательные аспекты регулирования технологий искусственного интеллекта, проблемы использования результатов интеллектуальной деятельности человека для обучения моделей ИИ и современные практики недопущения этого, а также вопросы маркирования нейросетевых продуктов.

Ключевые слова: *генеративные технологии, искусственный интеллект, генеративные нейросети, нейросети, маркирование нейросетевого продукта, авторское право.*

Irina I. KARPENKO

CONVERGENCE OF MEDIA CONTENT DISTRIBUTION ENVIRONMENTS: QR CODES AND AUGMENTED REALITY

The development of artificial intelligence technologies opens up great prospects for modern media. However, today, along with the undeniable positive aspects of using traditional and generative AI technologies, a number of problems related to copyright compliance are acute. The article examines the legislative aspects of regulating artificial intelligence technologies, the problems of using the results of human intellectual activity to train AI models and modern practices to prevent this, as well as issues of labeling neural network products.

Key words*: generative technologies, artificial intelligence, generative neural networks, neural networks, labeling of neural network products, copyright.*

Глобальное распространение и внедрение генеративных нейросетевых технологий в деятельность современных медиа вызывает целый ряд проблем, связанных с соблюдением авторского права и этических принципов медиа. Сложности, которые возникают сегодня в части соблюдения авторского права в области применения продукта генеративных нейросетей, связаны с несколькими факторами: отсутствием сбалансированной

законодательной базы в области регулирования технологий искусственного интеллекта (ИИ), сложностью отслеживания источников для обучения генеративных нейросетей, практикой создания и использования изображений, нарушающих этику и право, а также проблемой определения авторства при генерировании нейросетевого продукта.

На сегодняшний день лишь несколько стран в мире, осознавая опасности и риски и учитывая преимущества и перспективы, которые несет человечеству ИИ, проявили инициативу в области регулирования его деятельности. Так, Великобритания, страна, находящаяся на третьем месте в мире по исследованиям и разработкам в области искусственного интеллекта, приняла позицию «мягкого» регулирования, «…чтобы новаторы могли процветать, а риски, связанные с ИИ, можно было устранить. <…> жесткий подход может подавить инновации и замедлить внедрение ИИ. <…> Если мы поспешим принять законы слишком рано, мы рискуем возложить на бизнес неоправданное бремя»[1].

Соединенные Штаты Америки также в последние годы последовательно фиксировали в различных документах принципы работы с технологиями ИИ. Еще в 2016 г. администрация Барака Обамы начала уделять особое внимание рискам и правилам, связанным с ИИ. В докладе о будущих направлениях работы с ИИ под названием «Подготовка к будущему искусственного интеллекта» (Preparing for the Future of Artificial Intelligence) идет речь о текущем на момент публикации документа состоянии ИИ, существующих и потенциальных областях его применения, а также вопросах, которые прогресс в области ИИ ставит перед обществом и государственной политикой Соединенных Штатов. В отчете также даются рекомендации относительно конкретных дальнейших действий в области государственной политики по регулированию ИИ. «В целом, подход к регулированию продуктов с поддержкой искусственного интеллекта для защиты общественной безопасности должен основываться на оценке аспектов риска, которые добавление искусственного интеллекта может снизить, наряду с аспектами риска, которые оно может увеличить. Более того, если риск подпадает под рамки существующего режима регулирования, обсуждение политики должно начинаться с рассмотрения того, адекватно ли существующие нормативные акты уже учитывают этот риск или их необходимо адаптировать с учетом добавления

[1] A pro-innovation approach to AI regulation. URL: https://www.gov.uk/government/publications/ai-regulation-a-pro-innovation-approach/white-paper (date of access: 16.09.2024).

искусственного интеллекта»[1].

В этом же году Национальным советом по науке и технологиям (National Science and Technology Council) был разработан и обнародован Национальный стратегический план исследований и разработок искусственного интеллекта (The National Artificial Intelligence Research and Development Strategic Plan), включавший в себя стратегию финансирования исследований и разработок в области ИИ [2]. Стратегический план направлен на реализацию в трех направлениях: «Во-первых, он предполагает, что технологии искусственного интеллекта будут продолжать становиться все более сложными и повсеместными благодаря инвестициям правительства и промышленности в исследования и разработки в области искусственного интеллекта. Во-вторых, этот план предполагает, что влияние ИИ на общество будет продолжать возрастать, в том числе на занятость, образование, общественную и национальную безопасность, а также влияние на экономический рост США. В-третьих, предполагается, что отраслевые инвестиции в ИИ будут продолжать расти, поскольку недавние коммерческие успехи увеличили предполагаемую отдачу от инвестиций в НИОКР»[3].

11 февраля 2019 г. президентом США был подписан приказ «Сохранение лидерства Америки в области искусственного интеллекта» – достаточно нейтральный документ, направленный на «поддержание и укрепление научного, технологического и экономического лидерства Соединенных Штатов в области исследований и разработок в области искусственного интеллекта и их внедрения» [4].

Как справедливо отмечают авторы доклада Высшей школы экономики (ВШЭ) «Проблема машинного творчества в системе права…» «Несмотря на то, что количество нормативных актов, принятых в разных странах в отношении искусственного интеллекта, достаточно велико, такие документы преимущественно <…> носят прикладной характер (например, имея целью стимулирование развития

[1] PREPARING FOR THE FUTUREOF ARTIFICIAL INTELLIGENCE. URL: https/obamawhitehouse.archives.gov/sites/default/files/whitehouse_files/microsites/ostp/NSTC/preparing_for_the_future_of_ai.pdf (date of access: 17.09.2024).

[2] The National Artificial Intelligence Research and Development Strategic Plan. URL: https/obamawhitehouse.archives.gov/sites/default/files/whitehouse_files/microsites/ostp/NSTC/national_ai_rd_strategic_plan.pdf (date of access: 01.10.2024).

[3] Там же.

[4] Maintaining American Leadership in Artificial Intelligence. URL: https://www.federalregister.gov/documents/2019/02/14/2019-02544/maintaining-american-leadership-in-artificial-intelligence (date of access: 16.09.2024).

искусственного интеллекта в стране)»[1].

Новизна и непродолжительное время использования человечеством данной технологии, а также недостаточная законодательная база в части практики применения технологий искусственного интеллекта поднимает ряд острых проблем, связанных с работой нейросетевых моделей. В современном научном дискурсе мы находим работы, которые концентрируются на проблемах, связанных с правовым регулированием как автономно машинно-сгенерированных произведений, так и подобных объектов, созданных при посредстве человека [Оморов 2020; Ястребов, 2018; Мохат, 2018 и др.]. Однако работ, рассматривающих проблему использования машинами объектов интеллектуальной собственности в качестве материала для обучения модели нами найдено не было. Это говорит о том, что современная наука пока не до конца способна критически анализировать существующую в мире практику ИИ и представлять теоретическое обоснование существующим угрозам и проблемам.

Технологии генеративного искусственного интеллекта – самообучаемые системы, которые используют в качестве баз данных триллионы единиц информации. Разработчики подобных нейросетевых моделей неосознанно или намеренно допускают использование в качестве материала для обучения нейронных сетей не только «открытых» объектов – текстов, иллюстративного, аудио- или видеоматериала, распространяемых свободно, но и охраняемых объектов чьей-то интеллектуальной собственности.

В России и в мире на сегодняшний день не существует правоприменительной практики или законодательно закрепленных положений о допущении или недопущении подобных ситуаций. Взгляды медиапрактиков на этот счет сильно поляризованы. Есть точки зрения, что «эти системы, безусловно, способны нарушать авторские права и могут столкнуться с серьезными юридическими проблемами в ближайшем будущем. …<существуют и точки зрения, что то – И.К.>, что в настоящее время происходит в области генеративного ИИ, юридически безупречно и любые судебные иски обречены на провал»[2].

[1] Проблема машинного творчества в системе права: регулирование создания и использования результатов интеллектуальной деятельности с применением искусственного интеллекта, зарубежный опыт и российские перспективы. URL: https://www.hse.ru/mirror/pubs/share/482492820.pdf (date of access: 30.08.2024).

[2] The scary truth about AI copyright is nobody knows what will happen next. URL: https://www.theverge.com/23444685/generative-ai-copyright-infringement-legal-fair-use-training-data (date of access: 13.09.2024).

Вместе с тем, параллельно с развитием и совершенствованием технологий ИИ, в мире наметилась практика проведения охранных мероприятий объектов интеллектуальной собственности от индексации и включения в базы данных для машинного обучения, которые проводят, к примеру, владельцы фотобанков (Shutterstock) и социальных сетей (DeviantArt), защищая работы своих авторов.

Другим инструментом, ограждающим графический контент от индексации и обучения ИИ, являются программные продукты типа Nightshade и Glaze, задача которых состоит в защите стиля произведения и авторских прав в целом художников и фотографов при использовании их работ для обучения различных моделей искусственного интеллекта. Задачей подобных программ является «выделение набора особенностей оригинального изображения (назовем его «А»). Далее инструмент ищет замену этим особенностям, ориентируясь на какой-либо определенный стиль, жанр рисования или живописи, например, «кубизм» или «картины Ван Гога» и делая изображение «А» похожим именно на эти жанры или стили. Пользователь, делая промт на вывод изображения «А», ожидает получить картинку в стилистике определенного автора, однако в итоге видит совершенно бесполезную для него работу. То есть нейросети при обработке изображений с наложенными фильтрами Glaze ошибочно воспринимают запрос и выдают совершенно искаженный результат» [1]. Результативность инструментов доказана: «при использовании Glaze нейросети не удалось имитировать фирменные стили художников в 92% случаев, в то время как при отсутствии «Глазури» эти показатели равнялись примерно 7%» [2]. Подобные обработки, которые совершаются на уровне пикселей, невозможно выделить ни человеческому глазу, ни машине. Не требуется в этом случае и добавления никаких водяных знаков. Nightshade в отличие от Glaze позиционируется как более «агрессивная» система, работающая «аналогично Glaze, но вместо защиты от мимикрии стиля он разработан как инструмент нападения для искажения представлений признаков внутри генеративных моделей изображений ИИ <…> Glaze – это защитный инструмент, который отдельные художники могут использовать для защиты от атак мимикрии стиля, в то время как Nightshade – это наступательный инструмент, который художники могут

[1] Nightshade и Glaze — как работает защита авторского права художников от AI? URL: https://club.dns-shop.ru/blog/t-328-prilojeniya/118167-nightshade-i-glaze-kak-rabotaet-zaschita-avtorskogo-prava-hudojn/?utm_referrer=https%3A%2F%2Fwww.yandex.ru%2F (date of access: 01.10.2024).

[2] Там же.

использовать, чтобы помешать моделям, которые копируют их изображения без согласия (таким образом защищая всех художников от этих моделей)»[1].

На фоне мощной волны негатива в адрес ИИ выделяются редкие сторонники противоположной точки зрения, декларирующие важную роль ИИ в борьбе с нарушениями прав на интеллектуальную собственность. Так, некоторые исследователи полагают, что «с целью минимизации рисков нарушения прав владельцев РИД[2] необходимо иметь возможность собирать данные» [Фролова 2023: 490] из различных источников с целью поиска несанкционированных заимствований. «Модели распознавания изображений могут сканировать миллионы списков» [Там же], включая логотипы и товарные знаки, с целью обнаружения заимствованного контента. Подобные сканирования могут помочь и с позиции анализа текста, определяя его подлинность по стилистике, использованию лексики или устойчивых конструкций, которые были, к примеру, изъяты полностью или частично из других текстов по схожей тематике.

Целесообразным в контексте разговора о защите ИИ в области использования генеративного ИИ является и вопрос об определении авторства нейросетевого продукта. В ряде концепций, определяющих гипотетического правообладателя, рассматриваются: 1) непосредственный автор промта – команды, которая дается нейросети человеком и которая включает в себя набор признаков будущего произведения; 2) автор-разработчик нейросети как программного продукта; 3) сама нейросеть.

Относительно того, кто или что является субъектом права в отношении сгенерированного нейросетевого продукта, и мнения, и практика расходятся. Китайская, американская и европейская правоприменительные практики оставляют авторство за человеком, дающим команду программе, вследствие которой устанавливаются правила создания произведения, выбираются команды, факторы, пропорции, ставятся ограничения и устанавливаются допуски, конструируется форма будущего контента. Процесс же создания произведения с процессом творческим, составляющим описательный этап, не соотносится. В России понятие творческого труда, который становится главным опосредующим фактором результата

[1] What Is Nightshade? URL: https://nightshade.cs.uchicago.edu/whatis.html (date of access: 01.10.2024).

[2] Результат интеллектуальной деятельности.

интеллектуальной собственности, «делает невозможным (в рамках действующей в России модели) охрану произведений, созданных искусственным интеллектом» [Калятин 2022: 32].

Единственным государством в мире, где искусственный интеллект наделяется авторскими правами, является Новая Зеландия. «В законе об авторских правах 1994 г. под автором понимается лицо, которым в случае создания литературного, драматического, музыкального или художественного произведения, сгенерированного компьютером, предпринимались меры, необходимые для создания произведения. Сгенерированное компьютером произведение является таковым, если автором объекта авторских прав не является человек»[1]. Вместе с тем, следует отметить, что законодательство имеет ряд пробелов, так как обязанностями и ответственностью, приобретаемыми вместе с правами на авторство, искусственный интеллект не наделяется, что в принципе невозможно, однако нарушение им чьих-то интересов и прав вполне возможно.

Еще одним вопросом, актуальным в свете разговора об авторском праве в сфере ИИ для теории и практики медиа, является необходимость маркирования продукта генеративных нейросетей. Это происходит по ряду причин. Во-первых, с позиции авторского права указание авторства на любые произведения является общепринятой нормой. Классификация же продукта генеративных нейросетей выходит за рамки нормы. Так, согласно Гражданскому кодексу РФ, «автором произведения <…> признается гражданин, творческим трудом которого оно создано»[2]. Аналогичное определение автора произведения мы находим и в Законе РФ «Об авторском праве и смежных правах»: «автор – физическое лицо, творческим трудом которого создано произведение»[3]. В случае же, если продукт произведен программным образом, случайной выборкой, формированием произведения на основе материала, входящего в базу исходных данных нейросети, речь не идет об авторстве человека. В этой ситуации человек выполняет роль своего рода модератора, который предоставляет параметры для отбора базового материала, но в самом процессе создания произведения

[1] Васильева А. К вопросу о наличии авторских прав у искусственного интеллекта. URL: https://www.garant.ru/article/1605912/ (дата обращения: 01.10.2024).

[2] Ст. 1257 ГК РФ ч.4. Автор произведения. URL: https://www.consultant.ru/document/cons_doc_LAW_64629/7dde8dbb10c5ce94297e5eb859712be091044d70/ (дата обращения: 31.09.2024).

[3] Ст. 4. Основные понятия. Закон РФ Об авторском праве и смежных правах/ URL: https://www.consultant.ru/document/cons_doc_LAW_2238/ (дата обращения: 31.09.2024).

участия не принимает. Из этого следует: маркировка нейросетевого продукта может помочь избежать путаницы и конфликтов относительно того, кто имеет право на использование и распространение произведения.

Во-вторых, другим аспектом, оправдывающим практику маркирования продукта генеративных нейронных сетей, является этика доверия, которая может возникать между создателями и потребителями медиаконтента: потребители должны знать, что тот контент, с которым они взаимодействуют в процессе медиапотребления, сгенерирован машиной, а не произведен человеком. Сегодня, в целом, активные интернет-пользователи уже способны распознавать графический продукт, сгенерированный искусственным интеллектом, в силу его стилевой выраженности. Однако другой контент – текст, реалистичные видео, виртуализированные модели человека – сложно идентифицировать без дополнительного маркирования. В этом контексте стоит упомянуть о дипфейках – проблеме, которая стала актуальной в свете развития глубоких генеративных технологий.

В-третьих, маркирование нейросетевого продукта может также помочь реализовать СМИ их медиаобразовательный потенциал, помогая аудитории, столкнувшейся с подобным контентом, обучаться его идентифицировать. Эта практика может способствовать лучшему пониманию возможностей и ограничений искусственного интеллекта в современной медиапрактике.

Несмотря на очевидные, казалось бы, доводы в пользу маркирования нейросетевого продукта, существуют и связанные с этим сложности. Во-первых, в большинстве случаев использование нейросетевого продукта, особенно при генерации текстовых материалов, в неизменном виде не имеет места. Соотношение роли человека и роли машины в данной ситуации выявить сложно.

Также стоит отметить, что автоматическое маркирование нейросетевого продукта – нанесение, к примеру, водяных знаков будет требовать дополнительных усилий и ресурсов от медиакомпаний, что будет сказываться на удорожании подобного контента.

На сегодняшний день проблема маркирования объективно обсуждается на законодательном уровне во многих странах. Так, с августа 2023 г. маркировка нейросетевого продукта является обязательной в Китае. Регулирование и надзор за подобной практиков берет на себя государство. Маркировочные знаки должны находиться на видном месте и содержать «как минимум фразу “сгенерировано

искусственным интеллектом", "сгенерировано", "сгенерировано ИИ"»[1]. Документ, наряду с явной идентификацией посредством расстановки визуальных водяных знаков, предполагает и иной тип маркирования – неявный, посредством включения в файлы метаинформации, считывание которой возможно только с помощью технических средств, что упрощает программный контроль в условиях авторитарного госнадзора за подобным контентом.

Весной 2025 г. обязательным станет маркирование сгенерированного искусственным интеллектом продукта и в Европе. Регламент об искусственном интеллекте в части, регламентирующей требования о прозрачности, включает тезис: «Поставщики систем ИИ, включая системы ИИ общего назначения, генерирующие синтетический аудио-, графическое, видео- или текстовое содержимое, должны гарантировать, что выходные данные системы ИИ маркируются в машиночитаемом формате и могут быть обнаружены как искусственно сгенерированные или обработанные»[2].

В марте 2024 г. в США также был представлен законопроект, требующий «идентификации и маркировки онлайн-изображений, видео- и аудиоматериалов, созданных с использованием искусственного интеллекта»[3]. Подобную практику законодатели обязывают взять на себя разработчиков программных продуктов, а не непосредственных создателей контента, через расстановку водяных знаков или «вживления» в графические объекты метаданных. Как следует из этого заявления, маркирование не будет предусматриваться в случае генерации текстового контента, что останется только в области саморегулирования медиакомпаний.

Российская законодательная практика в области маркирования сгенерированного нейросетевого продукта также не отстает от общемировой. С весны 2024 г. в Государственной думе РФ ведется разработка концепции закона о маркировке нейросетевого продукта или продукта, измененного технологиями ИИ. В отличие от упомянутых ранее законодательных актов Европейского союза, Китая и США, Россия обсуждает возможности маркирования и текстового контента:

[1] 网络安全标准实践指南—生成式人工智能服务内容标识方法. URL: https://www.tc260.org.cn/upload/2023-08-25/1692961404507050376.pdf (date of access: 13.08.2024).

[2] The EU Artificial Intelligence Act. URL: https://artificialintelligenceact.eu/ (date of access: 18.09.2024).

[3] New bipartisan bill would require labeling of AI-generated videos and audio. URL: https://www.pbs.org/newshour/politics/new-bipartisan-bill-would-require-labeling-of-ai-generated-videos-and-audio (date of access: 18.09.2024).

«развитие ИИ-сервисов несет опасность появления огромных массивов не только никем не проверенных текстов, но и даже полностью выдуманных фактов, цифр и данных»[1]. Этот вопрос, однако, остается пока непроясненным, так как ни визуальной, ни программной возможности маркировать текстовые данные на уровне мета-информации сегодня нет. Маркировка графических объектов предполагается априори – посредством графических или водяных знаков. Ответственность за это налагается на российские ИИ-сервисы, прежде всего «Сбер» и «Яндекс».

Предполагается, что подобное маркирование предотвратит участившиеся случаи кибермошенничества с использованием, к примеру, аудио-дипфейков, которые преступники используют в телефонной коммуникации. Стоит, однако, заметить, что трансляция в режиме реального телефонного потокового общения подобных дипфейков не даст возможности идентифицировать ее как ИИ-генерацию, так как подобный контент возможно распознать программным образом только в случае обладания данным аудиофайлом с последующим чтением метаданных. То же можно сказать и о графических и видеофайлах.

Подытоживая, отметим, что современные мировые практики регулирования различных аспектов деятельности ИИ еще не синхронны. В этом ключе важным является нахождение точек соприкосновения в разных национальных правовых системах и соблюдение баланса между интересами всех субъектов права, вовлеченных в индустрию искусственного интеллекта. Также важна расстановка акцентов в части разумных сдерживаний развития практики использования ИИ и разработка конструктивных мер, стимулирующих его дальнейшее развитие и совершенствование.

Список литературы:

Калятин В.О. Определение субъекта прав на результаты интеллектуальной деятельности, созданные с использованием искусственного интеллекта и его влияние на развитие гражданского законодательства // Право. Журнал Высшей школы экономики. 2022. Т. 15. № 4. С. 24–50/

Морхат П.М. Право на результаты интеллектуальной деятельности, произведенные юнитом искусственного интеллекта: гражданско-правовые проблемы // Право и государство. 2018. № 1–2 (76–79). С. 204–215.

[1] Метки ставить негде. URL: https://www.kommersant.ru/doc/6729928?ysclid=m1qd85zerz859126906 (дата обращения: 01.10.2024).

Оморов Р.О. Интеллектуальная собственность и искусственный интеллект // E-Management. 2020. № 1. С. 43–49.

Фролова Е. Е., Купчина Е. В. Цифровые инструменты защиты прав на интеллектуальную собственность: на примере блокчейн и искусственного интеллекта // Вестник Пермского университета. Юридические науки. 2023. Вып. 3(61). С. 479–498.

Ястребов О.А. Правосубъектность электронного лица: Теоретико-методологические подходы // Труды Института государства и права РАН. 2020. Т. 13. № 2. С. 36–55.

КАРПЕНКО Ирина Ивановна – доктор филологических наук, доцент, доцент кафедры журналистики, Белгородский государственный национальный исследовательский университет (г. Белгород), *irinakarpenko@bsu.edu.ru*