права на произведения, созданные с помощью нейросетей, и допустимо ли вообще говорить об «авторском» праве на такие «произведения».

Дополнительно обращаем внимание, что в итоге диплом бакалавра студенту А. Жадану был выдан [7], а значит образовательная организация высшего образования признала у него наличие соответствующей квалификации. С учетом приведенной аргументации, полагаем, данное решение является правильным.

Список литературы

1. Витко В. Анализ научных представлений об авторе и правах на результаты деятельности искусственного интеллекта // ИС. Авторское право и смежные права. 2019. № 3. С. 5–22.

2. Искусственный интеллект не заменит работу человека. URL: https://www.kommersant.ru/doc/5798187

3. Как я написал диплом с помощью ChatGPT и оказался в центре спора о нейросетях в образовании. URL: https://journal.tinkoff.ru/neuro-diploma

4. Коданева С. И. Трансформация авторского права под влиянием развития цифровых технологий // Право и цифровая экономика. 2021. № 4. С. 31–38.

5. Положение о проведении государственной итоговой аттестации по образовательным программам высшего образования – программам бакалавриата, программам специалитета и программам магистратуры (новая редакция). URL: https://clck.ru/36oJJB

6. Синельникова В. Н., Ревинский О. В. Права на результаты искусственного интеллекта // Копирайт. 2017. № 4. С. 24–27.

7. Студенту РГГУ, который написал выпускную работу с нейросетью ChatGPT, вручили диплом. URL: https://msk1.ru/text/world/2023/03/15/72134255

8. Kasap A. Copyright and Creative Artificial Intelligence (AI) Systems: A Twenty-First Century Approach to Authorship of AIGenerated Works in the United States // Wake Forest Journal of Business & Intellectual Property Law. 2019. Vol. 19, № 4. Pp. 335–380.

Г. А. Грищенко,
кандидат юридических наук,
Московский государственный юридический
университет имени О. Е. Кутафина

ПРОБЛЕМЫ ПРАВОВОГО РЕГУЛИРОВАНИЯ ДИПФЕЙКОВЫХ ТЕХНОЛОГИЙ

Аннотация. Целью статьи является проведение анализа законодательства, правоприменительной практики, научной литературы по вопросам использования дипфейковых технологий, и выработка рекомендаций по совершенствованию правового регулирования соответствующих отношений. В условиях цифровизации вопрос о правовых аспектах применения дипфейковых технологий напрямую

связан с институтом персональных данных, что требует выработки новых подходов к установлению допустимых пределов использования личной информации. В статье предпринята попытка обобщить подходы к регулированию дипфейков не только в России, но и в других странах, акцентируя внимание, что подобные технологии могут нести как отрицательный, так и положительный эффект. Делается вывод, что совершенствование законодательства в указанном направлении должно носить комплексный характер и включать как правовые, так организационные и технические меры, связанные, в частности, с необходимостью выработки понятийного аппарата, введения обязательной маркировки дипфейкового контента, разработки свода практик и этического кодекса по применению данных технологий.

Ключевые слова: дипфейковые технологии, дипфейк, персональные данные, дискредитация, честь, достоинство, мошенничество, социальные сети, этический кодекс, правовые проблемы, совершенствование законодательства

PROBLEMS OF LEGAL REGULATION OF DEEPFAKE TECHNOLOGIES

Abstract. The purpose of the article is to analyze legislation, law enforcement practice, scientific literature on the use of deepfake technologies, and develop recommendations for improving the legal regulation of relevant relations. In the context of digitalization, the issue of legal aspects of the use of deepfake technologies is directly related to the institution of personal data, which requires the development of new approaches to establishing acceptable limits for the use of personal information. In this article, the author has attempted to summarize approaches to regulating deepfakes not only in Russia, but also in other countries, emphasizing that such technologies can have both negative and positive effects. It is concluded that improving legislation in this direction should be comprehensive and include both legal, organizational and technical measures related, in particular, to the need to develop a conceptual framework, introduce mandatory labeling of deepfake content, develop a set of practices and a code of ethics for the use of data technologies.

Keywords: deepfake technologies, deepfake, personal data, discredit, honor, dignity, fraud, social networks, code of ethics, legal problems, improvement of legislation

Развитие цифровых технологий, в частности, искусственного интеллекта, нейротехнологий, больших данных, отражается на всех сферах общественной жизни. Возможности современных цифровых технологий используются в различных процессах принятия управленческих решений, оказания государственных (муниципальных) услуг, коммуникации в сети Интернет, что зачастую связано с обработкой персональных данных, например, для идентификации и поиска физических лиц, распознавания фото-, аудио- и видеоизображений и т. д.

Особо следует отметить достижения нейросетей, которые способны копировать тембр, мимику, черты лица человека и обучаться на оригинальных интервью, фильмах, клипах, создавая новый контент, где в фокусе внимания оказываются личные данные. Уже не требуется обладание какими-то специальными навыками, чтобы заменить лицо и (или) голос абсолютно любого человека на видео или

фото и отграничить достоверную информацию от фейковой порой сложно даже специалистам в этой сфере. Многие социальные сети и иные Интернет-ресурсы предоставляют возможность путем несложных манипуляций заменить изображение (или голос) человека на лицо любой знаменитости.

В последние несколько лет одну из актуальных тем для обсуждений представляет вопрос о правовых аспектах применения дипфейковых технологий и допустимых пределах использования персональных данных.

В настоящее время в российском законодательстве определение дипфейков отсутствует. Можно отметить единственный документ, упоминающий о дипфейках применительно к возможности использования таких технологий при совершении преступлений (Приказ Генпрокуратуры России от 9 декабря 2022 г. № 746 «О государственном едином статистическом учете данных о состоянии преступности, а также о сообщениях о преступлениях, следственной работе, дознании, прокурорском надзоре»). При этом ни признаки дипфейков, ни требования и условия их применения в данном документе не раскрываются.

Конечно, определенные аспекты, связанные с обработкой персональных данных, распространением фейковой информации, в том числе в сети Интернет, и способами восстановления нарушенных прав, в настоящее время регулируются Уголовным кодексом Российской Федерации, Гражданским кодексом Российской Федерации (далее – ГК РФ), Кодексом Российской Федерации об административных правонарушениях, Федеральным законом от 27 июля 2006 г. № 149-ФЗ «Об информации, информационных технологиях и о защите информации» (далее – ФЗ «Об информации»), Федеральным законом от 27 июля 2006 г. № 152-ФЗ «О персональных данных», Законом Российской Федерации от 27 декабря 1991 г. № 2124-1 «О средствах массовой информации» и др. При этом очевидно, что в существующих реалиях указанные документы не способны в текущей редакции охватить всевозможные аспекты применения именно дипфейковых технологий, а неоднозначная правоприменительная практика указывает на необходимость выработки новых механизмов по защите прав граждан в цифровой среде, особенно в части использования их персональных данных.

Осенью 2023 г. в Государственной Думе Федерального Собрания Российской Федерации планируется рассмотрение соответствующей законотворческой инициативы, связанной с введением в нормативные правовые акты положений, направленных на регулирование применения дипфейковых технологий.

Сразу хотелось бы отметить, что дипфейки могут применяться как в мошеннических целях (и именно на борьбу с мошенничеством направлена государственная политика по законодательному регулированию данного явления), так и вполне в законных и благоприятных, например, в сфере кинематографа, искусства, рекламы. Известны случаи «сдачи» своего лица в аренду для создания новых персонажей, озвученных искусственным интеллектом, которые используются в рекламных или образовательных целях; многие режиссеры рассматривают возможность применения дипфейков для создания цифровых образов актеров, что особенно актуально для воспроизводства исторических фильмов, «оживления» умерших людей (разнообразные онлайн-выставки и виртуальные концерты известных художников, музыкантов).

Но все же дипфейки как воплощение новых технологических возможностей несут в себе серьезные риски и спектр этих рисков весьма широк и продолжает увеличиваться. Дипфейки могут представлять опасность как для отдельного человека или группы людей, так и создавать более серьезные угрозы в информационном пространстве, затрагивая государственные интересы и создавая угрозы национальной безопасности [5. С. 54–64; 8].

Повсеместный сбор биометрических персональных данных создает дополнительные риски: фейковое изображение можно будет использовать вместе с фейковыми отпечатками пальцев или следами ДНК. Биометрические атаки с технологией дипфейк выходят на новый уровень и используют уязвимость, в частности, государственных информационных систем, когда программный лайвнесс (определение, что перед камерой находится живой человек) не срабатывает. Такая практика вынуждает предъявлять новые требования к системам безопасности, обрабатывающим биометрию.

Сгенерированные изображения на основе чужих персональных данных могут быть использованы в развлекательных и шуточных целях (разнообразные пранки и розыгрыши), но зачастую они применяются в рамках дискредитации известных личностей (политиков, актеров, медийных лиц). Помимо этого, новые возможности нейросетей позволяют сгенерировать портрет несуществующего человека (URL: https://thispersondoesnotexist.com/) или создать аватар по описанию (URL: https://images.ai/), что порождает новые правовые проблемы распознавания фальшивых изображений (создание изображений, схожих до степени смешения с реальными лицами) [8].

Условно угрозы, связанные с применением дипфейковых технологий, можно разделить на следующие группы:

1. Причинение вреда репутации лицам вследствие искажения персональных данных, включая изображение и голос человека (дискредитация личности);

2. Искажение фактов (фейковые новости);

3. Мошенничество (целевой фишинг) [8].

В большинстве стран мира не существует четкого правового регулирования дипфейков. Наиболее активно противодействуют распространению вредоносного ложного контента Индия, КНР, США, Сингапур, Великобритания, Южная Корея, Япония, Австрия и Евросоюз. При этом правовая регламентация может охватывать как общие вопросы распространения дипфейков независимо от сферы регулирования соответствующих отношений, так и специальные отраслевые направления (например, избирательную сферу) [8].

В Китае, например, приняли закон, запрещающий с 1 января 2020 г. публиковать фейковые новости и вводящие в заблуждение «дипфейковые» видео, созданные с помощью искусственного интеллекта [6]. Дипфейки должны быть соответствующим образом маркированы, но при этом не уточняется, каким образом будет осуществляться дифференциация дипфейков от реальной, соответствующей действительности информации. В отношении нарушителей предусматривается уголовная ответственность, которая может быть применена как к производителям дипфейков, так и к интернет-сервисам, где соответствующий контент будет размещен [8].

В конце 2022 г. Администрация киберпространства Китая (CAC) опубликовала правила распространения фейковой информации в Интернет-пространстве (вступили в силу 10 января 2023 г.):

– фейки можно публиковать только с согласия изображаемого на них лица;

– фейки не должны наносить вред национальной безопасности;

– фейки не должны использоваться для обмана или распространения клеветы.

В США (штат Калифорния) приняли закон, запрещающий разрабатывать и распространять дипфейковый контент в пределах 60 дней до выборов; во Франции введены санкции за монтаж речи или изображения человека без его согласия; в Великобритании законодательство в основном нацелено на защиту чести и достоинства человека за распространение дипфейк-контента [8].

В настоящее время наблюдается курс на внедрение национальных законодательных мер, нацеленных именно на возложение большей ответственности за создание и распространение ложного контента на онлайн-платформы [3. С. 38].

Многие социальные сети уже разрабатывают и применяют политику по защите пользователей от фальшивого контента, и планируется, что в ближайшее время они научатся распознавать лица, сгенерированные с помощью дипфейковых технологий. Представляется, что самоцензура социальных сетей также может принести определенный положительный эффект, но очевидно, что обеспечение подлинности контента в цифровом пространстве должно сопровождаться правовыми методами со стороны государства.

В литературе высказывается мнение о целесообразности разработки и принятия унифицированного этического кодекса по применению дипфейковых технологий в различных сферах деятельности [7].

Разработка подобного этического кодекса будет способствовать установлению правил использования дипфейков, что позволит минимизировать репутационные риски и просчитывать способы, с помощью которых такие технологии могут быть использованы злоумышленниками.

Многие этические вопросы, связанные с технологиями (например, использования образа умершего человека), еще только предстоит оценить в полной мере. Но важно, чтобы принимаемое законодательство, не создав препятствий для развития искусственного интеллекта в творческих индустриях, обеспечило защиту человека, чей образ теперь так легко повторить и использовать [4. С. 87–103].

Распространение дипфейков затрагивает напрямую институт персональных данных, поскольку при создании соответствующих видео- и аудироликов используются изображения людей и (или) голос, что может причинить вред репутации абсолютно любого человека [8].

Участилась практика дискредитации публичных личностей (особенно в период избирательных кампаний), многие случаи распространения в сети Интернет дипфейкового контента сопровождаются судебными разбирательствами.

Все это свидетельствует о необходимости не только правового реагирования на уже существующие общественные отношения, но и разработки соответствующих технических решений, которые позволили бы устанавливать факт использования дипфейковых технологий. Представляется, что это может быть автоматизированный

сервис, который с помощью искусственного интеллекта мог бы выявлять факты генерации дипфейков и соответствующие нарушения законодательства в изображениях и видеоматериалах (по аналогии с разработанной информационной системой «Окулус», с января 2023 г. поэтапно внедряемой в мониторинговые программы Роскомнадзора и выявляющей в сети Интернет запрещенный контент). По сути, речь идет о механизме фактчекинга, когда любой пользователь Интернета может проверить сомнительный контент на факт применения дипфейк-технологий [8].

Подобные сервисы разрабатываются и в зарубежных странах. Например, в США внедрена программа экспертизы содержания (семантического анализа) мультимедийных материалов SemaFor, которая способна также определять манипуляции с фото и видео, созданные человеком, которые могут казаться семантически согласованными, но передавать ложную информацию [8].

Как отмечалось выше, распознать фейк зачастую сложно даже специалистам, хотя в целом можно выделить определенные моменты, на которые стоит обратить внимание при критической оценке контента (разница в качестве отрисовки элементов лица и остальных частей тела, явные пиксели в изображении, частота моргания, плохая синхронизация движений и т. д.).

Несмотря на правовое регулирование отношений, связанных с распространением «фейков» в России (ФЗ «Об информации» (ст. 15.3), УК РФ (ст. 207.1 – 207.3), КоАП РФ (ч. 9, 10, 10.1, 10.2 ст. 13.15), своевременный и хорошо реализованный дипфейк может пошатнуть всю политическую систему страны (например, в случае публикации речи главы государства с призывом к насилию или началу военных действий) [8].

Отдельно стоит отметить случаи так называемого целевого фишинга, когда дипфейк направлен на обман сотрудников определенной компании, чтобы заставить их выполнить какую-нибудь операцию (рассылка сообщений с корпоративной почты, дача поручений голосом руководителя) [8]. Известны случаи, когда дипфейки используются для подачи заявок на удаленную работу от имени несуществующих людей и «исчезают», получив аванс, или пытаются завладеть коммерческой информацией о компании изнутри.

Одной из серьезных проблем, способных нанести угрозу правам человека, является способность дипфейк-технологий с невероятной точностью имитировать интонацию, характерные паузы между словами, акцент, корректировать эмоции при произношении слов (голосовые дипфейки). И здесь возникают сложности, связанные с возможностью распространения авторского права на голос. Представляется, что дипфейки сто́ит рассматривать через призму производного произведения, при котором использование исходного произведения без согласия его правообладателя будет незаконно [2. С. 116; 8].

Очевидно, что в дальнейшем дипфейки будут разрабатываться на более высоком уровне, учитывая быстрое и качественное развитие возможностей нейросетей, в связи с чем могут увеличиться случаи фальсификации и дискредитации, однако преимуществ у данной технологии намного больше [8].

Способы защиты прав субъектов персональных данных при использовании дипфейковых технологий сводятся ко всем возможным формам защиты прав, но преимущественно касаются вопросов защиты чести, достоинства и деловой

репутации, защиты от недостоверной информации, включая возможность возмещения морального вреда в рамках гражданского судопроизводства (ст. 152 ГК РФ). При этом нельзя забывать и о возможности привлечения виновных к административной (например, при рассмотрении дел об оскорблении по ст. 5.61 КоАП РФ) и даже уголовной ответственности (например, в рамках рассмотрения дел о клевете по ст. 128.1 УК РФ).

Что касается использования, например, образов актеров или иных известных лиц для создания пародийного контента (популярных в настоящее время пранков), то нужно понимать, что само по себе использование дипфейковых технологий действующее российское законодательство не нарушает. Вместе с тем копирование образов реально существующих людей для создания подобных видео незаконно. Обнародование и использование изображения гражданина допускается только с его согласия (ст. 152.1 ГК РФ). В противном случае актер может обратиться в суд с требованием изъятия и удаления сгенерированной информации, но при этом потребуется доказать сходство внешности истца с изображенным на видео лицом, что можно сделать с помощью соответствующей экспертизы.

В целях совершенствования правового регулирования дипфейковых технологий в России весьма целесообразным является проработка вопросов о необходимости:

– введения в законодательство определения дипфейков (представляется, что подобная категория может быть отражена в понятийном аппарате ФЗ «Об информации»);

– разработки новых механизмов защиты, например, высших должностных лиц государства (по аналогии с государственной символикой и государственным флагом России) [1];

– разработки унифицированного этического кодекса (правил) по применению дипфейковых технологий в различных сферах деятельности;

– классификации дипфейков по видам угроз и последствиям с целью выработки эффективных механизмов по защите прав субъектов персональных данных (представляется, что подобная классификация должна отражать методы создания дипфейков (синтез изображения и (или) голоса), субъектов их создания (политические оппоненты, средства массовой информации), их цели (полезные или преступные), наносимый ущерб (от психологического воздействия до угроз национальной безопасности);

– создания экспертной комиссии, в которую следует включить специалистов различных направлений (от технических до юридических), в компетенцию которой входило бы изучение зарубежного опыта правового регулирования дипфейк-технологий, ведение свода практик применения таких технологий, оценка рисков и внедрение успешного опыта в национальные нормативные правовые акты;

– обязательной маркировки контента, созданного с помощью дипфейковых технологий, и привлечения к ответственности в случае ее отсутствия;

– разработки и внедрения технических решений (программ/сервисов), способных выявлять контент, созданный на базе дипфейков;

– популяризации применения дипфейковых технологий в целях положительного воздействия на сферы развлечения, образования, искусства, рекламы;

– повышения цифровой грамотности населения и разработки алгоритма действий в случае выявления подозрительного (сомнительного) контента [8].

При этом совершенствование законодательства следует осуществлять не только в отношении дипфейков как таковых, а в отношении использования технологий искусственного интеллекта в целом. Принципы использования искусственного интеллекта, указанные в Стратегии развития искусственного интеллекта в Российской Федерации, утвержденной Указом Президента Российской Федерации от 10 октября 2019 г. № 490, во многом должны определять и принципы применения дипфейков.

Одним из основных принципов правового регулирования допустимости применения дипфейковых технологий должно стать обеспечение безопасности, основанной на недопустимости использования данных технологий в целях умышленного причинения вреда гражданам и юридическим лицам, а также предупреждение и минимизация рисков возникновения негативных последствий их использования, прежде всего, в сфере обеспечения защиты персональных данных. Представляется, что вышерассмотренные предложения могут быть учтены при внесении соответствующих изменений в российское законодательство, прежде всего, регулирующее информационные правоотношения и вопросы привлечения к уголовной и административной ответственности за правонарушения в цифровой среде [8].

Список литературы

1. Алискеров М. Р. Угрозы и риски применения технологии «Deepfake» в противоправных целях // Информационная безопасность. 2022 № 2(72). С. 38–41.

2. Добробаба М. Б. Дипфейки как угроза правам человека // Lex russica. – 2022. Т. 75, № 11. С. 112–119.

3. Игнатьев А. Г., Курбатова Т. А. Аналитический обзор «Дипфейки в цифровом пространстве: основные международные подходы к исследованию и регулированию». Москва, 2023.

4. Калятин В. О. Дипфейк как правовая проблема: новые угрозы или новые возможности? // Закон. 2022. № 7. С. 87–103.

5. Киселев А. С. О необходимости правового регулирования в сфере искусственного интеллекта: дипфейк как угроза национальной безопасности // Вестник Московского государственного областного университета. Серия «Юриспруденция». 2021. № 3. С. 54–64.

6. Мун Д. В., Попета В. В. «From fake to deepfake»: угрозы и риски развития и распространения технологий искажения реальности в глобальном информационном пространстве. URL: https://cyberleninka.ru/article/n/from-fake-to-deepfake-ugrozy-i-riski-razvitiya-i-rasprostraneniya-tehnologiy-iskazheniya-realnosti-v-globalnom-informatsionnom

7. Bart van der Sloot, Yvette Wagensveld. Deepfakes: regulatory challenges for the synthetic society // Computer Law & Security Review. 2022. № 46. Art. 105716.

8. Грищенко Г. А. Обеспечение защиты персональных данных в условиях применения дипфейковых технологий // Устойчивое развитие России: правовое измерение: сборник докладов X Московского юридического форума. В 3-х частях. М., 2023. С. 197–202.