Право. Журнал Высшей школы экономики. 2024. Том 17. № 2.
Law. Journal of the Higher School of Economics. 2024. Vol. 17, no 2.

Право в современном мире

Научная статья
УДК: 342
DOI:10.17323/2072-8166.2024.2.215.240

Зарубежный опыт правового регулирования технологии «дипфейк»

Вадим Александрович Виноградов[1], Дарья Владимировна Кузнецова[2]

[1,2] Национальный исследовательский университет «Высшая школа экономики», Россия 101000, Москва, Мясницкая ул., 20.

[1] vavinogradov@hse.ru, https://orcid.org/0000-0001-8490-2893

[2] dvkuznetsova@hse.ru, https://orcid.org/0009-0003-4059-865x

Аннотация

В последние годы технология «дипфейк» набирает стремительную популярность, развиваясь с необычайной быстротой. С помощью различных приложений каждый может самостоятельно и без особого труда создать фото- видео- или аудио-дипфейки. При этом их использование поднимает различные этические вопросы, связанные с дезинформацией и согласием, и создает риск неправомерного использования, например, в сфере политики, в различных мошеннических схемах. Все это свидетельствует о необходимости выстраивания адекватных моделей нормативного регулирования технологии «дипфейк», создания системы актов, направленных на защиту прав человека, в том числе в цифровой среде, и предупреждения ненадлежащего использования и совершения правонарушений с использованием данной технологии. Вместе с тем, технология «дипфейк» может использоваться и в полезных целях, что ставит перед законодателем довольно сложную задачу по поиску оптимального баланса. Необходимо закрепить эффективную систему правил использования технологии «дипфейк» и ответственности за их нарушение, не создавая при этом труднопреодолимые преграды развитию технологии в целом или не запрещая использование технологии «дипфейк» полностью. В настоящей статье рассматривается опыт США, Китая и Сингапура в сфере правового регулирования технологии «дипфейк» с целью поиска наиболее удачной модели. При всех различиях подходы США и Китая схожи в части принятия специального регулирования, в то время как Сингапур следует иным путем — внесения точечных изменений в законодательство и решения вопросов

с помощью расширительного правоприменения. Авторы констатируют, что законодательные меры всех стран отражают стремление адаптировать свою правовую систему к вызовам, создаваемым развивающимися цифровыми технологиями. Рассмотренный опыт (при его изучении и адаптации) может быть полезен в создании оптимальной российской модели правового регулирования технологии «дипфейк». Ключевым решением видится необходимость маркировки всех видов дипфейк-контента.

Ключевые слова

дипфейк; искусственный интеллект; фейки; Интернет; маркировка; цифровые технологии; дезинформация; синтетический контент.

Для цитирования: Виноградов В.А., Кузнецова Д.В. Зарубежный опыт правового регулирования технологии «дипфейк» // Право. Журнал Высшей школы экономики. 2024. Том 17. № 2. С. 215–240. DOI:10.17323/2072-8166.2024.2.215.240

Law in the Modern World

Research article

Foreign Experience in Legal Regulating Deepfake Technology

Vadim A. Vinogradov[1], Daria V. Kuznetsova[2]

[1,2] National Research University Higher School of Economics, 20 Myasnitskaya Str., Moscow 101000, Russia,

[1] vavinogradov@hse.ru, https://orcid.org/0000-0001-8490-2893

[2] dvkuznetsova@hse.ru, https://orcid.org/0009-0003-4059-865x

Abstract

Deepfake technology has been gaining rapid popularity in recent years, developing at an extraordinary speed. With the help of various applications, anyone can create a photo, video or audio deepfakes on their own and without much difficulty. At the same time, their use raises various ethical issues related to misinformation and consent, and creates a risk of misuse, for example, in the sphere of politics, in various fraudulent schemes. All this indicates the need to build adequate models of legal regulation of deepfake technology, to create a system of acts aimed at protecting human rights, including in the digital environment, and preventing improper use and commission of offences using this technology. At the same time, the deepfake technology can also be used for good purposes, it poses a very difficult task for the legislator to find a balance — it is necessary to fix an effective system of rules for the use of the deepfake technology and responsibility for their violation, without creating difficult to overcome obstacles to the development of technology in general or, without prohibiting the use of deepfake technology completely. The article reviews the experience of the USA,

China and Singapore in the sphere of legal regulation of the deepfake technology in order to find the most successful model. Despite all the differences, the approach of the USA and China is similar in terms of adopting specialized regulation, while Singapore is moving in a different direction — adopting point changes to legislation and addressing issues through extensive enforcement. In any case, it can be stated that the legislative measures of all countries reflect the desire to adapt their legal systems to the challenges posed by emerging digital technologies. The reviewed experience (if it will be taken into account and adapted) may be useful for the creation of an optimal Russian model of legal regulation of the deepfake technology. It is thought that labelling all types of deepfake content is the key solution.

Keywords

deepfake; artificial intelligence; fakes; Internet; labelling; digital technology; disinformation; synthetic content.

For citation: Vinogradov V.A., Kuznetsova D.V. (2024) Foreign Experience in Legal Regulating Deepfake Technology. *Law. Journal of the Higher School of Economics,* vol. 17, no. 2, pp. 215–240 (in Russ.) DOI:10.17323/2072-8166.2024.2.215.240

Введение

Стремительная цифровизация привела к тому, что технология «дипфейк» сегодня доступна каждому желающему. Пройдя довольно быстро и уверенно путь от инструмента развлечения до серьезной технологии, получившей широкое применение в сфере политтехнологий, рекламы, киноиндустрии, дипфейки заняли важную и особую нишу в среде систем искусственного интеллекта (далее — ИИ).

Сам термин «дипфейк» (deepfake) происходит от двух английских слов: deep learning — «глубокое обучение» и fake — «фальшивый». Технически дипфейк означает «реалистичную манипуляцию аудио-, фото- и видеоматериалами с помощью искусственного интеллекта для достижения максимального сходства с подлинными изображениями и звуковыми дорожками. В большинстве случаев в основе метода лежат генеративно-состязательные нейросети (GAN)»[1]. Технологию «дипфейк» определяют как «технологию, использующую искусственный интеллект для создания или редактирования содержимого видео или изображения, чтобы показать что-то, что никогда не происходило» [Young N., 2019: 8]. Концепция генеративно-состязательных сетей была выдвинута в 2014 году Яном Гудфеллоу и его соавторами в одноименной статье. Речь идет о состязательном процессе, когда обучаются две нейросети — генератор (G) и дискриминатор (D). Первая нейросеть ищет данные и генерирует подобные, вто-

[1] Available at: URL: https://rdc.grfc.ru/2020/06/research-deepfake/ (дата обращения: 12.03.2024)

рая среди сгенерированного первой сетью объема ищет несовершенные и подвергает их критике. Генеративная сеть стремится стать лучше в создании данных, в то время как дискриминативная сеть стремится стать лучше в их обнаружении [Goodfellow I. et al., 2014: 2672].

Исследования показывают, что крайне трудно отличить дипфейк от реального видео. Это наглядно иллюстрирует, например, эксперимент, описанный в одной из статей, опубликованных в научном журнале iScience. Испытуемым показали 16 видеороликов, половина из которых была сделана с помощью технологии «дипфейк». В результате бо́льшая часть таких роликов осталась неузнанной. При этом испытуемые были крайне уверены в свои силах и правоте и не сомневались, делая выбор. Это доказывает, что видео, даже сделанные не профессионалами, а в обычных приложениях, доступных к массовому использованию, настолько качественны, что человек без сомнений может воспринять их за достоверные [Köbis N. C., Doležalová B., Soraperra I., 2021: 11]. Для создания дипфейка уже не требуется длительное обучение и огромное количество данных. Уже в 2019 году эксперты отмечали, что «достаточно 10-секундного клипа» [Skibba R., 2020: 724], а с тех пор технологии продвинулись еще дальше.

Простота использования для создания базовых продуктов, появление различных приложений для широкого круга пользователей привели к увеличению рисков ненадлежащего использования, угрозам различных правонарушений с применением технологии «дипфейк». Среди тех, которые лежат на поверхности — порноместь[2], политические инсинуации, политическая дискредитация и введение в заблуждение[3], различные виды мошенничества, подлог аудио-, фото- и видеодоказательств в суде [Pfefferkorn R., 2020: 265][4], обман систем визуализации и т.п. По данным компании Deeptrace[5], дипфейки порнографического содержания составили в 2019 году 96% всего объема дипфейков. На оставшиеся 4% приходились все остальные виды возможного использования[6]. Все это вызывало

[2] Available at: https://www.wired.com/story/most-deepfakes-porn-multiplying-fast/ (дата обращения: 12.03.2024)

[3] Available at: https://edition.cnn.com/2019/05/23/politics/doctored-video-pelosi/index.html (дата обращения: 12.03.2024)

[4] Также см: https://www.telegraph.co.uk/news/2020/01/31/deepfake-audio-used-custody-battle-lawyer-reveals-doctored-evidence/ (дата обращения: 12.03.2024)

[5] Амстердамская компания, специализирующаяся на кибербезопасности, технологии глубокого обучения и компьютерного зрения для обнаружения и онлайн-мониторинга синтетических медиа.

[6] Available at: http://regmedia.co.uk/2019/10/08/deepfake_report.pdf (дата обращения: 12.03.2024)

беспокойство и стремление законодателей в различных странах отрегулировать прежде всего эту сферу использования технологии «дипфейк».

Дипфейки искажают реальность, заставляя все больше сомневаться, где правда, а где вымысел. Они подрывают доверие к информации в сети, вызывая чувства потерянности, неуверенности и тревоги. Технологии продвинулись настолько далеко, что не составляет труда придумать новый мир, полный новых придуманных событий.

Между тем технология сама по себе не является отрицательной и может использоваться и во благо, она способна помочь решить целый ряд вопросов — например, использовать в кино образы ушедших актеров, которые из-за кончины не успели закончить работу в кинопроекте[7], обеспечить органичный перевод видеороликов на разные языки[8], смоделировать внешность известных исторических личностей и т.п.

В качестве примера можно привести кейс известного американского актера Брюса Уиллиса. У актера было диагностировано заболевание, связанное с повреждением головного мозга, приводящим к расстройству речи. Поэтому Уиллис в 2022 году принял решение продать свой образ компании Deepfake для создания «цифрового двойника»; теперь двойник может участвовать в различных рекламных роликах, фильмах и т.п. вместо болеющего актера, который сам сделать это уже не в состоянии. В рекламе компании «Мегафон» как раз уже участвовал дипфейк–двойник знаменитого актера[9].

Также делаются предложения использовать дипфейки в оперативно-розыскной деятельности для дезинформирования ее объектов [Батоев В.Б., 2023: 72]. Поэтому в контексте правового регулирования было бы ошибкой ставить вопрос о запрете технологии в целом, речь должна идти о создании нормативно-правовых рамок ее правомерного использования на благо человека и общества.

В настоящей работе на примере США, Китая и Сингапура рассмотрен зарубежный опыт правового регулирования использования технологии «дипфейк» с целью поиска оптимальных конструкций и решений и выявления возможности применения их к российской действительности. Данные страны избраны ввиду из высокого технологического развития, что ставит перед законодателями серьезные вызовы при поиске оптимального регуляторного решения для сохранения баланса интересов как технологических компаний, так и рядовых граждан.

---

[7] Available at: URL: https://www.bfm.ru/news/498962 (дата обращения: 12.03.2024)

[8] Available at: URL: https://hightech.plus/2019/04/29/synthesia-kto-i-kak-snimal-dipfeik-klip-s-bekhemom-i-k-chemu-eto-privedet (дата обращения: 12.03.2024)

[9] Available at: URL: https://www.forbes.ru/forbeslife/478431-akter-brus-uillis-prodal-prava-na-svoj-obraz-kompanii-deepfake (дата обращения: 12.03.2024)

1. Правовое регулирование технологии «дипфейк» в США

Данное регулирование в США осуществляется как на федеральном уровне, так и на уровне штатов. Это объясняется устройством государства, которое сложилось исторически. «В американской модели федерализма 50 штатов, образующих Соединенные Штаты Америки, располагают большим объемом независимости и власти. По сути дела США — это страна, состоящая из 51 различного государства: 50 государственных механизмов штатов и отдельного федерального государственного механизма. В каждом из этих механизмов имеется своя правовая система» [Бернам У., 2006: 45].

Значительным шагом стало включение положений о дипфейках в Закон об оборонном бюджете на 2020 год[10] . Посвященные дипфейкам положения подчеркивают серьезность, с которой правительство США относится к угрозам, исходящим от искаженной информации, и их потенциальному влиянию на национальную безопасность и внутриполитические процессы. Раздел 256 NDAA 2020 требовал от министра обороны в течение 180 дней после вступления Закона в силу направить Конгрессу доклад о мерах, принятых для идентификации и противодействия манипуляциям с медиаконтентом. Это включает разработку технологий для обнаружения дипфейков и стратегий для борьбы с манипулятивными материалами, угрожающими национальной безопасности. В Законе подчеркивалась важность взаимодействия Министерства обороны с научным сообществом и промышленностью для разработки методов обнаружения и нейтрализации фальсифицированных медиаматериалов.

Анализ способностей потенциальных противников по созданию и распространению дипфейков назван критически важным для понимания уровня технологического превосходства и разработки соответствующих контрмер. Разделы 5724 и 5709 предусматривали конкурсы для стимулирования разработки технологий автоматического обнаружения машинно-манипулированных медиа и обязывали «разведывательное сообщество» исследовать использования дипфейков за рубежом[11]. Эти меры отражают комплексный подход к проблеме манипулирования информацией, осознание ее масштабов и воздействия на общественное

[10] National Defense Authorization Act (NDAA) — федеральный закон, принимаемый ежегодно и определяющий бюджет Министерства обороны США в соответствующем периоде.

[11] H.R.2500 — National Defense Authorization Act for Fiscal Year 2020. Available at: https://www.congress.gov/bill/116th-congress/house-bill/2500/text (дата обращения: 12.03.2024)

мнение и национальную безопасность. Помимо технологических разработок, акцент делался на сотрудничестве различных ведомств и секторов, а также на международной координации реагирования на угрозы, связанные с дипфейками.

Закон об оборонном бюджете на 2021 год (National Defense Authorization Act for Fiscal Year 2021)[12] также содержал положения, касающиеся угроз, связанных с применением технологии «дипфейк», и рекомендаций по мерам для их предотвращения. Важно, что для обозначения дипфейков используются и другие термины — «машинно-манипулированные медиа» (machine-manipulated media), «синтетические медиа» (synthetic media) или «подделки цифрового контента» (digital content forgeries). В отличие от NDAA 2020 года, который требовал от директора национальной разведки отчитываться только об использовании незаконного контента иностранными государствами или их доверенными лицами для подрыва национальной безопасности США, этот Закон поручал Министерству внутренней безопасности (Department of Homeland Security; (DHS) изучать не только как иностранные правительства используют дипфейки для нанесения ущерба национальной безопасности, но и то, как дипфейки используются для « мошенничества», нанесения вреда «уязвимым группам» или нарушения законов о гражданских правах. Министр внутренней безопасности в течение пяти лет должен выпускать ежегодный отчет о «подделках цифрового контента»[13].

В 2023 году Агентство национальной безопасности совместно с ФБР и Агентством кибербезопасности и защиты инфраструктуры (CISA) выпустили информационный листок кибербезопасности (Cybersecurity Information Sheet (CSI) «Контекстуализация угроз дипфейков для организаций». В документе подчеркиваются легкость и масштабность манипуляций киберпреступников с мультимедиа, что опасно для национальной безопасности. Указывается на настоятельную необходимость распознавания технологии «дипфейк» сотрудниками организаций и разработку стратегий реагирования для минимизации их воздействия[14]. Таким образом ведется активная внутриведомственная работа, направленная на привлечение внимания к проблеме дипфейков и предотвращение угроз, с ними связанных.

---

[12] Available at: https://www.congress.gov/116/plaws/publ283/PLAW-116publ283.pdf (дата обращения: 12.03.2024)

[13] Congress's deepening interest in deepfakes. Available at: URL: https://thehill.com/opinion/cybersecurity/531911-congresss-deepening-interest-in-deepfakes/ (дата обращения: 12.03.2024)

[14] Contextualizing Deepfake Threats to Organizations. Available at: URL: https://media.defense.gov/2023/Sep/12/2003298925/-1/-1/0/CSI-DEEPFAKE-THREATS.PDF (дата обращения: 12.03.2024)

Одним из нормативных актов федерального уровня, направленных на решение проблем, возникающих в связи с дипфейками, является Закон об идентификации результатов работы генеративных состязательных сетей (Identifying Outputs of Generative Adversarial Networks Act, IOGAN)[15]. Он признает стремительное развитие сетей GAN и потенциальные последствия их результатов. Закон предписывает Национальному научному фонду и Национальному институту стандартов и технологий (NIST) поддерживать и развивать исследования методов генеративных состязательных сетей. Это необходимо для более глубокого понимания их возможностей и того, как они могут быть использованы как в положительных, так и в отрицательных целях[16].

Одним из важных аспектов Закона является акцент на разработку методов обнаружения дипфейков. Данный Закон поощряет сотрудничество федеральных агентств, частных структур и исследователей. Решение проблем, связанных с GAN и дипфейками, требует согласованных усилий различных секторов общества.

В настоящий момент на рассмотрении Конгресса США находится Закон об ответственности за дипфейки (DEEPFAKES Accountability Act of 2023)[17]. Законопроект инициирован и внесен сенатором от штата Нью-Йорк Иветтой Кларк[18]. Предыдущий одноименный законопроект был внесен в декабре 2019 года, однако не был принят.

Законопроект направлен на «защиту национальной безопасности от угроз, исходящих от технологии «дипфейк», а также на обеспечение правовой защиты жертв вредоносных дипфейков»[19]. Текст законопроекта содержит требование об обязательной маркировке дипфейков. Дипфейки, содержащие аудио- и видео-элементы должны сопровождаться аудиопредупреждением, текстовой информацией и иметь значок, предупре-

[15] Available at: URL: https://www.govinfo.gov/content/pkg/PLAW-116publ258/pdf/PLAW-116publ258.pdf (дата обращения: 12.03.2024)

[16] Congress's deepening interest in deepfakes. Available at: https://thehill.com/opinion/cybersecurity/531911-congresss-deepening-interest-in-deepfakes/ (дата обращения: 12.03.2024)

[17] DEEPFAKES в названии закона является сокращением от Defending Each and Every Person from False Appearances by Keeping Exploitation Subject, полное наименование закона- "Defending Each and Every Person from False Appearances by Keeping Exploitation Subject to Accountability Act of 2023". Подробнее: H. R. 5586 (118th): DEEP FAKES Accountability Act. Available at: https://www.congress.gov/118/bills/hr5586/BILLS-118hr5586ih.pdf (дата обращения: 12.03.2024)

[18] Подробнее о законопроекте: Clarke leads legislation to regulate deepfakes. Available at: https://clarke.house.gov/clarke-leads-legislation-to-regulate-deepfakes/ (дата обращения: 12.03.2024)

[19] Преамбула. H. R. 5586 (118th): DEEP FAKES Accountability Act. Available at: https://www.congress.gov/118/bills/hr5586/BILLS-118hr5586ih.pdf (дата обращения: 12.03.2024)

ждающий, что аудиовизуальное произведение создано с применением технологии «дипфейк».

Для аудиодипфейков предполагается голосовое информирование перед началом самой аудиозаписи, для фото и видео дипфейков — текстовое уведомление о том, что контент создан с использование технологии «дипфейк» (в виде текста или значка). Нарушение требований об обязательной маркировке для создания злонамеренных дипфейков, включая дипфейки сексуального содержания, связанные с преступной деятельностью, используемые для подстрекательства к насилию и связанные с иностранным вмешательством в выборы признается преступлением, за которое предусмотрено наказание в виде штрафа или лишения свободы до 5 лет[20]. За все остальные виды немаркированных дипфейков будут следовать гражданско-правовые санкции, включая право на частный иск, например, на возмещение ущерба.

Если законопроект будет принят, он потребует от создателей маркировать все дипфейки, загружаемые на интернет-платформы, и делать прозрачными все изменения, внесенные в видео или другой тип контента. Согласно законопроекту, онлайн-платформы, на которых размещается генеративный ИИ-контент, также должны будут показывать происхождение этого контента[21].

Неясно, получит ли законопроект поддержку для принятия. Однако инициатива иллюстрирует желание на федеральном уровне и не секторально, а глобально подойти к регулированию технологии «дипфейк», т.е. сфокусироваться не на отдельных сферах применения дипфейков, а на их «подсвечивании». Это будет препятствовать введению в заблуждение пользователей, в том числе распространению ложной информации.

Что касается законодательства штатов, то в ряде штатов действуют законы, которые призваны ограничить вред от дипфейков в разных сферах. Гавайи, Виргиния, Техас и Вайоминг ввели уголовную ответственность за дипфейки порнографического содержания, а Калифорния и Техас допускают гражданско-правовые иски. В Техасе и Калифорнии также действуют законы, ограничивающие использование дипфейков, которые могут повлиять на политические кампании[22].

---

[20] H.R. 5586. Available at: URL: https://www.congress.gov/118/bills/hr5586/BILLS-118hr5586ih.pdf (дата обращения: 12.03.2024)

[21] Available at: URL: https://abcnews.go.com/Politics/bill-criminalize-extremely-harmful-online-deepfakes/story?id=103286802#:~:text=At%20this%20time%2C%20there%20is,to%20discuss%20possible%20government%20regulation (дата обращения: 12.03.2024)

[22] The Legal Issues Surrounding Deepfakes. Available at: https://www.honigman.com/the-matrix/the-legal-issues-surrounding-deepfakes (дата обращения: 12.03.2024)

Рассмотрим регулирование дипфейков подробнее на примере Калифорнии. Закон AB 602 (Assembly Bill No. 602)[23] направлен на борьбу с дипфейками порнографического содержания и оставляет возможность обратиться в суд, если кто-то без согласия лица создает или распространяет откровенный контент с его изображением. В Законе признается потенциальный вред от использования чьего-либо подобия в откровенном контенте без согласия, что может привести к эмоциональному, репутационному и даже экономическому ущербу. Потерпевший вправе требовать у лица, создавшего контент, возмещения ущерба. Сюда входит любая прибыль, полученная от материалов, созданных и размещенных без согласия, компенсация за эмоциональные страдания и любой другой вред, причиненный распространением таких материалов. Следует подчеркнуть, что хотя Закон AB 602 является важным шагом в решении проблемы дипфейков, он направлен именно на откровенный контент. Для решения других проблем, связанных с использованием дипфейков (дезинформация или мошенничество), необходимо использовать другие правовые средства или будущие законы.

Другой действующий в Калифорнии Закон — AB 730 (Assembly Bill No. 730)[24], признавая возможность влияния дипфейков на выборы, запрещает распространение видео-, аудио- или фотоизображений, направленных на введение избирателей в заблуждение относительно действий или слов кандидата с целью нанесения ущерба его репутации или обмана избирателя, заставив его голосовать за или против кандидата. Закон распространяется на материалы, опубликованные за 60 дней до выборов. Однако если в таких материалах содержится информация, что они подверглись манипуляциям, не являются подлинными и созданы с использованием технологии «дипфейк», это является исключением. Сообщение, призванное предупредить зрителей или слушателей о незаконном характере контента, должно быть четким и заметным. Оно должно быть размещено таким образом, чтобы зрители или слушатели были информированы, что потребляемый ими контент не является подлинным.

Согласно Закону лицо, которому был нанесен ущерб в результате нарушения этих положений, вправе подать гражданско-правовой иск с целью получения судебного постановления, запрещающего распространение дипфейк аудио- или видеоматериалов. Оно также вправе требовать возмещения ущерба от распространителя информации.

---

[23] Available at: https://leginfo.legislature.ca.gov/faces/billTextClient.xhtml?bill_id=202120220AB602 (дата обращения: 12.03.2024)

[24] Available at: https://leginfo.legislature.ca.gov/faces/billTextClient.xhtml?bill_id=201920200AB730 (дата обращения: 12.03.2024)

Основной целью AB 730 является защита честности выборов. В политической среде, где общественное мнение может быть изменено вирусным контентом, обеспечение того, чтобы избиратели не были введены в заблуждение с помощью дипфейков, имеет первостепенное значение. Закон признает проблемы, создаваемые дипфейками, и пытается дать избирателям инструменты, позволяющие отличить факт от вымысла.

Однако в США попытки ограничить использование технологии «дипфейк» могут столкнуться с препятствием в виде Первой поправки[25], запрещающей ограничение свободы слова [Joost L., 2023: 312]. Дипфейки допустимо рассматривать как форму самовыражения, что подпадает под защиту Первой поправки. В то же время существует ряд ограничений ее применения, например, клевета или диффамация. Под эти исключения подпадают дипфейки порнографического содержания, созданные и распространяемые без согласия [Blitz M. J., 2020: 167]. Такой сценарий весьма вероятен в свете позиции Верховного суда США, признающего абсолютный характер свободы слова. Это иллюстрирует решение 2013 года по иску Агентства США по международной помощи против Международного альянса за открытое общество (Alliance for Open Society International), в котором процитированы слова верховного судьи Джексона (произнесенные им в прошлом веке): «Если и есть недвижимая звезда в американском конституционном созвездии, так это правило, по которому никакое официальное лицо, высокого или низкого ранга, не может указывать, что является непреложной истиной в вопросах политики, религии или других вопросах, касающихся мнений, или принуждать граждан признаваться в этом на словах и действовать согласно этому убеждению»[26].

Также необходимо подчеркнуть, что платформы, на которых размещаются дипфейки, защищены иммунитетом от судебных исков, связанных со сторонним контентом, размещенным в их онлайн-сервисах, в соответствии с разделом 230 Закона о соблюдении пристойности в Интернете 1996 года (Communications Decency Act)[27].

Дипфейки также стали предметом обсуждения во время забастовки 2023 года, организованной Гильдией актеров (SAG) и Гильдией сценаристов Америки (WGA); обсуждалось использование ИИ и то, как он

[25] Available at: URL: https://www.archives.gov/founding-docs/bill-of-rights-transcript (дата обращения: 12.03.2024)

[26] Supreme Court Stands With NGOs in Landmark Free Speech Ruling. Available at: https://www.huffpost.com/entry/supreme-court-stands-with_b_3474671 (дата обращения: 12.03.2024)

[27] 47 U.S. Code § 230 — Protection for private blocking and screening of offensive material. Available at: https://www.law.cornell.edu/uscode/text/47/230 (дата обращения: 12.03.2024)

может влиять на кинематографическую отрасль[28]. Была поднята проблема незаконного и противозаконного использования изображений и видеозаписей голливудских актеров, созданных ИИ и используемых без их согласия. В настоящее время ведутся дискуссии о том, чтобы Гильдия продюсеров и различные другие организации занялись регулированием этой проблемы и создали ассоциацию, которая будет следить за ростом числа дипфейков и находить решения для принятия юридических мер по защите интересов актеров[29].

2. Правовое регулирование технологии «дипфейк» в Китае

В 2019 году китайские разработчики выпустили приложение для создания дипфейков — ZAO. Оно сразу стало пользоваться популярностью и завоевало высокий рейтинг в магазинах приложений. Пользователю приложения достаточно было загрузить всего одну фотографию, чтобы уже через несколько секунд стать героем одного из своих любимых фильмов. Для более качественного видео требовалось несколько больше фотографий. Выход приложения вызвал жаркие споры о возможных рисках, связанных, прежде всего, с конфиденциальностью данных, поскольку разработчик изначально указывал в условиях использования, что весь созданный контент хранится на его серверах и является его собственностью. Впоследствии, однако, была добавлена возможность удаления пользователями своего контента без возможности восстановления[30].

Еще одним китайским прорывом является Tencent Cloud AI Digital Human — платформа для создания «цифровых людей», цифровых аватаров, созданных с помощью технологии «дипфейк». Компания Tencent Cloud выпустила свой продукт в апреле 2023 года. Заявлены два сценария возможного использования — «трансляция виртуального изображения с синтезом искусственного интеллекта и голосовое взаимодействие в

[28] Гильдия актеров экрана — Американская федерация артистов телевидения и радио не смогла договориться со студиями о базовой оплате труда и остаточном вознаграждении, а также о регулировании использования этих инструментов с ИИ для защиты работы, созданной человеком, и предотвращения использования цифровых копий актеров без оплаты их труда. Available at: https://amateurphotographer.com/latest/photo-news/hollywood-strikes-and-deepfakes-how-far-will-ai-powered-tools-go/ (дата обращения: 12.03.2024)

[29] Available at: https://easternmirrornagaland.com/ai-deepfake-videos-worry-hollywood-actors-ahead-of-wage-talks/ (дата обращения: 12.03.2024)

[30] Available at: https://www.forbes.com/sites/zakdoffman/2019/09/03/chinese-viral-app-zao-forced-to-backtrack-after-wechat-ban-and-privacy-backlash/ (дата обращения: 12.03.2024)

реальном времени»[31]. Используемые технологии позволяют создать полностью реалистическую цифровую личность. Все это свидетельствует о высоком технологическом прогрессе Китая в вопросах создания искусственно созданного контента.

Опыт Китая, где широко распространена система биометрической идентификации, показывает, что риски неправомерного использования технологии «дипфейк» носят вполне реальный характер. Это, в частности, подтверждает нашумевший в 2021 году взлом группой налоговых мошенников государственной системы распознавания лиц Китая с целью подделки налоговых счетов и наживы на этом[32]. Особое беспокойство вызвало происшествие в апреле 2023 года во Внутренней Монголии — автономном районе КНР. Там некий господин Го, деятель одной из крупных компаний, стал жертвой мошенничества с использованием технологии «дипфейк». После видеочата в сети WeChat (как он думал — со своим другом) он перевел ему 4,3 млн. юаней. Впоследствии выяснилось, что и видео, и голос друга были воссозданы мошенниками с помощью технологии «дипфейк», чтобы ввести жертву в заблуждение и убедить ее перевести деньги. После этого полиция КНР обратилась к гражданам с предупреждением о новом виде мошенничества и призвала быть внимательными, а также вдумчиво относиться к защите персональных данных, прежде всего биометрических[33].

В феврале 2024 года широкий резонанс вызвало финансовое мошенничество с использованием технологии «дипфейк» в Гонконге. Оно привлекло внимание даже не суммой в 25 млн. долл., перечисленных жертвой обмана, а тем, что впервые речь шла об имитации групповой видеоконференции, т.е. жертва общалась по видеосвязи не с одним, а сразу с несколькими «знакомыми», созданными с помощью технологии «дипфейк». На проведенном после инцидента пресс-брифинге полиция Гонконга сообщила о целом ряде афер с использованием систем ИИ; в 20 случаях речь

---

[31] Tencent Cloud AI Digital Human. Available at: https://www.tencentcloud.com/products/ivh (дата обращения: 12.03.2024)

[32] «Счета-фактуры, выставленные Государственной налоговой администрацией, используются для отслеживания платежей и помогают бороться с уклонением от уплаты налогов. Прокуратура Шанхая объявила, что преступная группа обманула систему проверки личности этой платформы, используя подтасовку личной информации и фотографии высокого разрешения, которые были куплены на черном рынке в Интернете, поэтому ее зарегистрированная подставная компания может выставлять клиентам поддельные налоговые счета». Available at: https://www.scmp.com/tech/tech-trends/article/3127645/chinese-government-run-facial-recognition-system-hacked-tax?module=perpetual_scroll_0&pgtype=article&campaign=3127645 (дата обращения: 12.03.2024)

[33] Available at: https://www.chinadaily.com.cn/a/202305/22/WS646b4fd3a310b6054fad4731.html (дата обращения: 12.03.2024)

шла об использовании фотографий с украденных удостоверений личности для обмана систем распознавания лиц[34].

Успех ZAO осветил создавшуюся в Китае проблему дипфейков и рисков, связанных с их использованием.

Прежде чем рассматривать нормативное регулирование дипфейков в Китае, следует напомнить одну из особенностей китайской правовой системы — относительно небольшое количество законов по сравнению с большим количеством подзаконных актов. Административно-правовые акты являются одним из важнейших источников современного права Китая [Косихина С., Швец А., 2022: 12] и регулируют широкий круг вопросов. «Склонность современного китайского законодателя к малому числу основополагающих интегральных актов и потоку частных нормативов — особенность, корнями уходящая в далекое прошлое китайской истории» [Трощинский П.В., 2016: 45].

В 2020 году ЦК Коммунистической партии Китая издал «План реализации проекта по созданию правового общества (2020-2025 годы)»[35], 6-й раздел которого посвящен управлению киберпространством в соответствии с законом [LiY., 2023: 115]. Пункт 22 данного раздела закрепляет необходимость совершенствования правовой системы в Интернете, которое заключается в распространении законов и нормативных актов на киберпространство и их оптимизации в области информационных онлайн-услуг, а также разработке и развитии мер по регулированию и управлению использованием новых медиаформатов, в числе которых названо применение таких новых технологий, как дипфейки. Это подчеркивает обеспокоенность властей Китая быстротой развития технологий и понимание ими необходимости кибербезопасности. Власти стремятся законодательно отрегулировать деятельность в киберпространстве, подчеркивая, что управление им не должно быть хаотичным и бесконтрольным, оно должно происходить в строгом соответствии с законом.

В ноябре 2022 года Администрацией киберпространства Китая, Министерством промышленности и информационных технологий и Министерством общественной безопасности приняты «Положения об управлении глубоким синтезом информационных сервисов в Интернете», вступившие в силу 10.01.2023[36]. Указанные Положения разработаны

[34] Available at: https://www.scmp.com/news/hong-kong/law-and-crime/article/3250851/everyone-looked-real-multinational-firms-hong-kong-office-loses-hk200-million-after-scammers-stage (дата обращения: 12.03.2024); Available at: https://www.cnn.com/2024/02/04/asia/deepfake-cfo-scam-hong-kong-intl-hnk/index.html (дата обращения: 12.03.2024)

[35] Available at: https://www.gov.cn/zhengce/2020-12/07/content_5567791.htm (дата обращения: 12.03.2024)

[36] Available at: https://www.gov.cn/zhengce/zhengceku/2022-12/12/content_5731431.htm (дата обращения: 12.03.2024)

на основе законов «О кибербезопасности», «О безопасности данных», «О защите личной информации», «Об управлении информационными службами Интернета» и других законов и нормативных актов КНР. В Положениях закреплен запрет на использование дипфейков в целях угрозы национальной безопасности и интересам, для нанесения ущерба имиджу нации, посягательства на общественные интересы, нарушения экономического и социального порядка или посягательства на законные права и интересы других лиц.

Отдельно зафиксирован запрет на использование дипфейков для создания, воспроизведения, публикации или распространения ложной новостной информации (ст. 6). Положения отражают национальные особенности и традиции и указывают, что оказание услуг глубокого синтеза должно не только соответствовать законам и другим нормативным актам, но и уважать общественные нравы и этическую мораль, придерживаться правильной политической ориентации, ориентации на общественное мнение и ценностной ориентации, а также способствовать продвижению услуг глубокого синтеза во благо (ст. 4).

Важно отметить, что в отличие от американских, китайские меры регулирования дипфейков направлены на разработчиков программного обеспечения для создания дипфейков и их поставщиков, а не на платформы или пользователей таких программ [Hine E., Floridi L., 2022: 608–609]. Положения устанавливают целый ряд требований для разработчиков и поставщиков. Основным требованием является обязательная маркировка дипфейков. Данный подход нацелен на предотвращение дезинформации и манипуляции общественным мнением, а также на защиту от недобросовестного использования данной технологии в коммерческих или политических целях. Также важным требованием является обязанность поставщика систем глубокого синтеза получить согласие человека, лицо и голос которого подвергаются редактированию[37]. Этот подход отражает уважение к праву на неприкосновенность частной жизни и личных данных.

За нарушение ст. 6 «Положений об управлении глубоким синтезом информационных сервисов в сети Интернет» (устанавливает запрет на распространение ложной новостной информации) в мае 2023 года в Китае задержали человека за распространение фейковой новости о крушении поезда[38].

---

[37] Available at: https://www.oxfordmartin.ox.ac.uk/events/chinas-deepfake-regulations/#:~:text=The%20regulations%20prohibit%20the%20use,have%20wide%20latitude%20to%20interpret. (дата обращения: 12.03.2024)

[38] Согласно отчету полиции, подозреваемый использовал ChatGPT для написания самой новости, а также ботов, чтобы новость набрала больше просмотров. Available at: https://rg.ru/2023/05/10/chat-bot-dovel-do-tiurmy.html (дата обращения: 12.03.2024)

Китай стал одной из первых стран, принявших специальное регулирование, направленное на дипфейк-контент. Основу системы регулирования составляет обязательная маркировка, что является важным шагом в борьбе с цифровой дезинформацией. Обязательная маркировка дипфейк-контента будет способствовать повышению прозрачности, что позволит пользователям идентифицировать поддельный контент. Это может значительно снизить риски введения в заблуждение и манипулирования общественным мнением.

3. Правовое регулирование технологии «дипфейк» в Сингапуре

В апреле 2022 г. житель Сингапура Оуэн стал жертвой вымогательства с использованием технологии «дипфейк». Некто получил доступ к телефону и контактам Оуэна и угрожал ему распространением среди его знакомых ролика интимного характера, будто бы с участием Оуэна, созданного с помощью дипфейка. За нераспространение ролика злоумышленник требовал 8000 сингапурских долл. (около 5796 долл. США). В результате видео было распространено, Оуэн подал заявление в полицию и узнал, что некоторые из его друзей столкнулись с аналогичным шантажом[39]. Данный случай привлек к проблеме внимание широких слоев общественности страны.

Высокий уровень технического прогресса вынуждает Сингапур занимать активную позицию в поисках адекватного правового регулирования применения технологии «дипфейк» и распространения в Интернете фальсифицированной информации чтобы оградить права граждан и национальную безопасность. Пример Сингапура интересен отличным от США и Китая подходом. В этой стране пока отсутствует регулирование технологии «дипфейк», законодатель пошел путем принятия общего закона о фейках и точечных изменений в законодательстве, которые позволяют распространить его действие и на дипфейки. При этом в стране большое внимание уделяется развитию новых технологий, и ИИ, в частности, а также защите персональных данных. Основной целью является поиск баланса между стимулированием инноваций и защитой прав граждан.

В 2019 году в Сингапуре принят Закон о защите от ложных сведений и манипуляций в Интернете (Protection from Online Falsehoods and Manipulation Act, POFMA)[40]. Основной целью POFMA является предот-

[39] Available at: https://news.yahoo.com/singaporean-mans-face-ends-deepfake-171743924.html (дата обращения: 12.03.2024)

[40] Available at: https://www.pofmaoffice.gov.sg/regulations/protection-from-online-falsehoods-and-manipulation-act/ (дата обращения: 12.03.2024)

вращение распространения ложных фактов онлайн, подавление поддержки и противодействие последствиям такого распространения, а также обеспечение исправления или разъяснения ложных утверждений. Учитывая растущие проблемы, связанные с распространением фейков в Интернете, и их потенциальную опасность для общества, POFMA был введен для поддержания общественного доверия. POFMA распространяется не только на письменный контент, но и на «ложные или вводящие в заблуждение» изображения и видео, что позволяет регулировать в том числе и дипфейки. Закон препятствует влиянию ложных сведений на выборы. Таким образом, этот запрет распространим и на использование дипфейк-видеороликов для манипулирования выборами. При этом важно, что мнения, критика, сатира и пародии, как правило, не подпадают под действие POFMA.

Согласно POFMA правительство имеет право отдавать распоряжения об исправлении или удалении контента в Интернете, который оно считает ложным. Любой министр полномочен решать, что является ложью в Интернете и требовать исправления или удаления такой информации. Это реализуется путем выпуска соответствующих предписаний. Лица, получившие предписание в соответствии с POFMA, вправе обратиться к выпустившему его министру с просьбой отменить или изменить его, а в случае отказа обжаловать его в Высоком суде Сингапура. Этот механизм обжалования обеспечивает защиту от возможного превышения полномочий правительства и его членов[41].

Физические и юридические лица, уличенные в нарушении POFMA, могут быть подвергнуты серьезным наказаниям в виде крупного штрафа или тюремного заключения. В то же время отдельные юридические или физические лица, например, действующие от имени правительства Сингапура, освобождаются от выполнения некоторых положений POFMA[42].

Хотя POFMA введен с целью защиты общественных интересов, Закон подвергается критике в связи с возможностью злоупотребления или превышения полномочий со стороны правительства, а также с последствиями для свободы слова. Есть опасения, что Закон может подавить законные дискуссии и инакомыслие [Lee H., Le T., 2023: 304].

Если дипфейки используются для нанесения вреда, например, в случае мошенничества, диффамации или угроз, к ним могут быть применены соответствующие статьи Уголовного кодекса Сингапура (Penal Code

[41] Singapore Fake News Laws: Guide to POFMA (Protection from Online Falsehoods and Manipulation Act). Available at: https://singaporelegaladvice.com/law-articles/singapore-fake-news-protection-online-falsehoods-manipulation/ (дата обращения: 12.03.2024)

[42] Министр имеет право приказом, опубликованным в Бюллетене (Gazette), освободить любое лицо или группу лиц от любого положения POMFA (ст.61).

1871)[43]. Если речь идет об использовании дипфейков для порномести, то могут быть применимы статьи 377BE и 377BD Уголовного кодекса. Здесь важно отметить, что данные статьи были приняты для борьбы с порноместью в целом, а потому есть целый ряд трудностей в их применении к контенту, созданному с применением технологии «дипфейк». Например, в ст. 377BE основное внимание уделено распространению или угрозе распространения изображений или записей интимного характера, при этом Закон предполагает наличие личной связи между распространителем и жертвой, ожидая, что распространитель будет знать о возможности домогательства или унижения. Когда же распространитель не знаком с жертвой, доказать умысел на унижение или оскорбление становится проблематично.

Другой важный момент — преследованию по Закону подвергается распространитель такого контента, но не создатель, если он не занимается его распространением. Владение порнографией, созданной с помощью технологии «дипфейк», в случае, когда создатель сохраняет материал, криминализирует ст. 377BD. Таким образом, хотя ст. 377BE и 377BD Уголовного кодекса и затрагивают некоторые аспекты проблемы, они не позволяют решать многоплановые и эволюционирующие проблемы, связанные с порнографией, созданной с помощью технологии «дипфейк»[44].

Очевидно, что уголовное законодательство Сингапура в этой части требует совершенствования. Назрело четкое закрепление ответственности не только распространителей и владельцев, но и создателей таких дипфейк материалов. Также требует совершенствования регулирование положений об умысле, особенно когда распространитель не связан с жертвой. Необходимо юридическое признание того, что даже некачественные дипфейки могут нанести значительный ущерб и должны влечь за собой юридические санкции.

К дипфейкам порнографического характера может быть применим Закон о защите от домогательств (Protection From Harassment Act, POHA)[45]. Некоторые его разделы могут быть применимы к порнографическому контенту, созданному с применением технологии «дипфейк». Например, ст. 3 POHA криминализирует преднамеренное преследование, попытки вызвать тревогу или беспокойство у жертвы. Положения данного раздела могут быть применены к создателям или распространителям дипфейков порнографического содержания, если они намере-

[43] Available at: https://sso.agc.gov.sg//Act/PC1871 (дата обращения: 12.03.2024)

[44] Available at: https://lawtech.asia/fake-porn-real-harm-examining-the-laws-against-deepfake-pornography-in-singapore/ (дата обращения: 12.03.2024)

[45] Available at: https://sso.agc.gov.sg/Act/PHA2014 (дата обращения: 12.03.2024)

вались вызвать указанные чувства у жертвы. Это является ключевым моментом, поскольку в случае с дипфейками не всегда можно доказать указанный умысел. В таком случае возможно говорить о применимости ст. 4 POHA, которая аналогична ст. 3, но не содержит требования о наличии умысла [Soon J., 2022: 194]. В данном случае основное внимание уделяется тому, испытала ли жертва указанные выше чувства, что также трудно доказать применительно к дипфейкам порнографического содержания. За нарушение ст. 3 предусмотрено наказание в виде штрафа до 5000 сингапурских долл. или тюремное заключение на срок до 6 месяцев, в то время как за нарушение ст. 4 грозит только штраф в таком же размере.

Еще один возможно применимый состав содержится в ст. 7 POHA — он криминализирует причинение тревоги и беспокойства путем преследования. Однако в данном случае возникает требование о неоднократности действий. Необходимость неоднократного поведения означает, что единичные случаи создания или распространения дипфейков могут быть не охвачены. Тем не менее жертвы, подпадающие под положения указанных статей, могут обратиться с гражданским иском о возмещении ущерба, т.е. существует инструмент компенсации.

Все названные нюансы свидетельствуют о необходимости дальнейшего совершенствования законодательства Сингапура для решения уникальных проблем, встающих в связи с использованием технологии «дипфейк».

Актом, направленным на противодействие вредоносному дипфейк-контенту, также можно считать Закон о безопасности в Интернете (разные поправки) (Online Safety (Miscellaneous Amendments) Act 2022[46]). Его принятие ознаменовало значительный шаг в регулировании онлайн-платформ и обеспечении безопасности пользователей Интернета, особенно детей. Закон внес поправки в Закон о телерадиовещании (Broadcasting Act 1994)[47], возлагая на социальные сети ответственность за защиту пользователей от вреда в Интернете. Управление развития инфокоммуникационных средств массовой информации (Infocomm Media Development Authority, IMDA) получило право отдавать социальным сетям (например, YouTube, TikTok и т.п.) распоряжения об удалении вредоносного контента, например, пропагандирующего членовредительство, сексуальную эксплуатацию детей, терроризм или способного

[46] Available at: https://sso.agc.gov.sg/Acts-Supp/38-2022/Published/20221221?DocDate=20221221 (дата обращения: 12.03.2024)

[47] Available at: https://www.imda.gov.sg/-/media/imda/files/regulations-and-licensing/regulations/codes-of-practice/codes-of-practice-media/code-of-practice-for-online-safety.pdf (дата обращения: 12.03.2024)

разжечь расовую или религиозную ненависть [Manique A.E. et al., 2023: 16]. Соответственно, дипфейки подобного содержания также должны быть удалены.

Нарушителям требований Закона о безопасности в Интернете — платформам социальных сетей грозят штрафы в размере до 1 млн. сингапурских долл., а интернет-провайдеры могут быть оштрафованы на сумму до 500 тыс. сингапурских долл. за неблокировку вредоносных сервисов.

Принятие Закона о безопасности в Интернете свидетельствует о согласованных усилиях правительства Сингапура по проактивному регулированию онлайн-пространства, направленных на решение эволюционирующих проблем безопасности в Интернете при соблюдении баланса между принципами неприкосновенности частной жизни и свободы выражения мнений.

Использование дипфейков в преступных целях также подпадает под действие нового закона, принятого в июле 2023 года — Закона о преступном вреде в Интернете (Online Criminal Harms Act 2023)[48]. Этот Закон наделяет правительство Сингапура дополнительными инструментами для удаления или блокирования контента, если есть подозрение в использовании его в преступных целях. Соответствующие директивы могут даваться физическим и юридическим лицам, провайдерам онлайн-сервисов, провайдерам доступа в Интернет, а также магазинам приложений.

Закон распространяется на юридических и физических лиц независимо от их физического присутствия или ведения бизнеса в Сингапуре или за его пределами. Однако обеспечение соблюдения Закона зарубежными организациями сопряжено с трудностями, и при несоблюдении Закона правительство вправе прибегнуть к судебному преследованию или распоряжению о блокировании доступа. Закон направлен на создание более безопасного онлайн-пространства и предотвращение преступного вреда в Интернете и отражает глобальную тенденцию, когда правительства все чаще ищут способы регулирования онлайн-пространства для пресечения преступной деятельности, балансируя между необходимостью обеспечения безопасности и свободами и преимуществами, предлагаемыми Интернетом.

Законы Сингапура о кибербезопасности (Cybersecurity Act 2018[49]) и о защите персональных данных (Personal Data Protection Act 2012, PDPA)[50]

[48] Available at: https://www.parliament.gov.sg/docs/default-source/default-document-library/online-criminal-harms-bill-17-2023.pdf (дата обращения: 12.03.2024)

[49] Available at: https://sso.agc.gov.sg/Acts-Supp/9-2018/Published/20180312?DocDate=20180312 (дата обращения: 12.03.2024)

[50] Available at: https://sso.agc.gov.sg/Act/PDPA2012 (дата обращения: 12.03.2024)

также содержат положения, которые могут косвенно относиться к дипфейкам. Например, если данные человека используются без разрешения для создания дипфейка, это может быть признано нарушением правил защиты данных. Следует отметить, что хотя эти нормы направлены на предотвращение злонамеренного использования дипфейков, существует опасение, что они могут быть чрезмерными или неправомерными, а также иметь негативные последствия для свободы выражения мнений. Баланс между предотвращением вреда и защитой свободы слова — проблема, с которой сталкиваются многие страны, занимающиеся этим вопросом.

Нормативным актом, который может быть применен к использованию дипфейков в политической борьбе, является Закон о противодействии иностранному вмешательству (The Foreign Interference Countermeasures Act, FICA)[51]. Принятый в 2021 году FICA направлен на борьбу с иностранным вмешательством во внутренние дела Сингапура через информационные и телекоммуникационные технологии. Меры безопасности включают возможность блокировки вредоносного контента, запрет на использование определенных приложений и технологий, предотвращение распространения дезинформации, а также введение строгих контрольных мер для медиаплатформ и интернет-сервисов, особенно могущих быть использованы в распространении иностранного влияния[52].

Сингапур без сомнения добился успехов в безопасности в Интернете. Принятие POFMA и FICA сделали Сингапур мировым лидером борьбы с дезинформацией в Интернете [Chng K., 2023: 236]. Властям даны широкие полномочия по удалению вредоносного контента. В контексте дипфейков речь прежде всего идет о материалах порнографического содержания, распространяемых с целью причинения вреда жертве, как материального, так и морального. Однако, как признают власти Сингапура[53], предпринятых усилий может быть недостаточно, чтобы успевать за темпами технологического прогресса. Поэтому особую роль выполняют практические решения, которые позволят оперативно удалять вредоносный контент в Интернете, не накладывая при этом чрезмерного бремени на жертв. Крайне важно сосредоточить внимание на правах отдельных лиц в онлайн-пространстве, изучить возможности расширения их прав и защиты от негативных последствий вреда, наносимого в Интернете.

---

[51] Available at: https://sso.agc.gov.sg//Acts-Supp/28-2021/ (дата обращения: 12.03.2024)

[52] Available at: https://www.mha.gov.sg/docs/default-source/default-document-library/summary-factsheet-on-fica.pdf (дата обращения: 12.03.2024)

[53] Available at: https://www.straitstimes.com/singapore/further-laws-needed-to-protect-victims-of-online-harms-shanmugam (дата обращения: 12.03.2024)

Высокому уровню физической безопасности в Сингапуре должен соответствовать такой же высокий уровень безопасности в сети; страна к этому активно стремится, продолжая совершенствовать меры безопасности в цифровом пространстве. Крайне важно постоянно адаптировать и совершенствовать нормативную базу защиты граждан от этих развивающихся угроз, включая технологию «дипфейк».

Заключение

Технология «дипфейк» позволяет убедительно имитировать голоса и вносить цифровые изменения в фото-, видео- и аудиоматериалы, превращая ложь в кажущийся факт. Стремительное развитие и расширение сферы применения дипфейков, доступность и простота использования самой технологии ввиду появления в широком доступе большого количества различных инструментов для быстрого и удобного создания разнообразного дипфейк-контента привели к увеличению противоправного использования и появлению новых способов манипуляций с помощью синтетического контента. Дипфейки широко используются для создания порнографических материалов, для политической дезинформации, а также мошенничества. При этом технология может использоваться и в полезных целях. Все это подчеркивает необходимость соответствующего правового регулирования и ставит перед законодателем трудную задачу нахождения баланса — с одной стороны, закрепить систему правил использования технологии «дипфейк» и ответственности за их правонарушение, с другой стороны, не создав при этом преграды на пути развития технологии в целом или не запретив использование технологии «дипфейк» полностью.

При регулировании дипфейков, как показывает зарубежный опыт, нахождение и поддержание баланса требуется между существующими в правовой системе ценностями и границами использования новой технологии. Например, в США постоянно стоит вопрос о необходимости обеспечения свободы слова.

Три юрисдикции — США, Китай и Сингапур олицетворяют три разных подхода к решению проблемы.

В США, несмотря на растущее признание проблем, создаваемых дипфейками, нормативная база находится в процессе развития. Сочетаются инициативы штатов, которые реализуются в двух направлениях — запрета использования технологии «дипфейк» для создания контента порнографического содержания без согласия и запрета на использование дипфейков для влияния на выборы, и более широкие федеральные меры, направленные в первую очередь на защиту национальной безопас-

ности. На рассмотрении находится законопроект об ответственности за дипфейки (DEEPFAKES Accountability Act 2023). Подобные инициативы реализованы лишь в нескольких штатах, в остальных регулирование дипфейков отсутствует. Тем не менее в целом США пристально следят за развитием и применением технологии «дипфейк» с целью предотвращения возможных рисков и угроз от ее неправомерного использования.

В Китае с января 2023 года действуют «Положения об управлении глубоким синтезом информационных сервисов в Интернете». Они требуют у поставщиков услуг по созданию контента и пользователей маркировать контент, подвергнутый манипуляциям любого типа. «Положения» также содержат запрет на фейковые новости. Особое внимание уделяется рискам неправомерного использования технологии «дипфейк» и попыткам регулировать ее использование с помощью специального законодательства.

В Сингапуре нет закона, прямо направленного исключительно на регулирование дипфейков, но в Законе о защите от ложной информации и манипуляций (POMFA) закреплено положение о возможности блокировки или удаления контента в связи с освещением ложной информации. Закон о противодействии иностранному вмешательству (FICA) предусматривает меры безопасности, включающие как возможность блокировки вредоносного контента, так и запрет на использование определенных приложений и технологий, предотвращение распространения дезинформации. Целый ряд других законов также может быть применим к случаям противоправного использования технологии «дипфейк».

При всех различиях подходы США и Китая схожи в части принятия специального регулирования, а не общих норм, регулирующих широкий спектр общественных отношений, в сферу регулирования которых можно отнести и дипфейки, а именно специальных нормативных актов, направленных на регулирование дипфейков. В то же время Сингапур, будучи также технологически развитым государством и также имея потребность в решении проблемы, движется пока иным путем — точечных изменений в законодательстве и решения вопросов с помощью расширительного правоприменения. Этот подход видится спорным, поскольку пока нюансы уже существующих конструкций не позволяют охватить полностью весь спектр противоправного применения технологии «дипфейк», что требует или дальнейшего совершенствования действующих норм, или все же принятия отдельного законодательства, направленного на регулирование технологии «дипфейк».

Законодательные меры всех стран отражают стремление адаптировать правовую систему к вызовам, создаваемым развивающимися цифровыми технологиями.

Из зарубежного опыта проистекает выявление дипфейков и их «подсвечивание», которое может быть достигнуто маркировкой. Для выявления дипфейков следует широко внедрять различные программы обнаружения синтетического контента. Такие программы обязательно должны быть в распоряжении государственных органов, СМИ и у различных цифровых платформ. Ключевой мерой является обязательная маркировка всех видов дипфейк-контента. Такие подходы в правовом регулировании дадут возможность предотвратить злонамеренное использование дипфейков, сохранив при этом в целом возможность их использования.

Важна также точность регулирования использования персональных данных при создании дипфейк-контента. Внимания заслуживает опыт Китая и зафиксированная в его «Положениях об управлении глубоким синтезом информационных сервисов в Интернете» обязанность получить согласие лица, биометрические данные которого подвергнутся изменениям в результате применения технологии «дипфейк».

Рассмотренный опыт принесет пользу при создании оптимальной российской модели правового регулирования технологии «дипфейк».

Список источников

1. Батоев В.Б. Об использовании технологии «Deepfake» в оперативно-розыскной деятельности // Юридическая наука и практика: Вестник Нижегородской академии МВД России. 2023. № 1. С. 70–76.

2. Бернам У. Правовая система США. М.: Новая юстиция, 2006. 1216 с.

3. Косихина С.С., Швец А.В. Правовая система КНР. Учебное пособие. Благовещенск: Амурский государственный университет, 2022. 90 с.

4. Трощинский П.В. Правовая система Китая. М.: Наука, 2016. 472 с.

5. Blitz M.J. Deepfakes and Other Non-Testimonial Falsehoods: When Is Belief Manipulation (Not) First Amendment Speech. Yale Journal of Law & Technology, 2020, vol. 23, pp. 161-300.

6. Chng K. Falsehoods, Foreign Interference, and Compelled Speech in Singapore. Asian Journal of Comparative Law, 2023, no. 2, pp. 235–252. doi:10.1017/asjcl.2023.9

7. Goodfellow I.J. et al. Generative Adversarial Nets. In: Advances in Neural Information Processing Systems. Montreal: University Press, 2014, pp. 2672-2680. Available at: https://leimao.github.io/downloads/blog/2017-12-26-Going-into-GANs/Ian-Goodfellow-GANs.pdf (дата обращения: 15.02.2024).

8. Hine E., Floridi L. New deepfake regulations in China are a tool for social stability, but at what cost? Nature Machine Intelligence, 2022, vol. 4, pp. 608–610.

9. Joost L. The Place for Illusions: Deepfake Technology and the Challenges of Regulating Unreality. University of Florida Journal of Law and Public Policy, 2023, vol. 33, no. 2, pp. 309–332.

10. Köbis N.C., Doležalová B., Soraperra I. Fooled twice: People cannot detect deepfakes but think they can. iScience, 2021, vol. 24, issue 11, pp. 1–17.

11. Lee H., Lee T. Between two Acts: competing narratives, activism and governance in Singapore's digital sphere. Internet Histories, 2023, vol. 7, no. 4, pp. 295–312. https://doi.org/10.1080/24701475.2023.2232214

12. Li Y. Implementing Rule of Law Concept in the Digital Sphere: China's Experience. Legal Issues in the Digital Age, 2023, vol. 4, no. 4, pp. 114–133. https://doi.org/10.17323/2713-2749.2023.4.114.133

13. Manique A.E. et al. Industry approaches in handling online exploitation of children: A comparative study of the policy, guidelines and best practices in Malaysia, Singapore and Australia. Cogent Social Sciences, 2023, vol. 9(2), pp. 1–25. https://doi.org/10.1080/23311886.2023.2241713.

14.Pfefferkorn R. "Deepfakes" in the courtroom. Boston University Public Interest Law Journal, 2020, vol. 29, no. 2, pp. 245–276.

15. Skibba R. Media Enhanced by Artificial Intelligence: Can We Believe Anything Anymore? Engineering, 2020, vol. 6, issue 7, pp. 723–724. DOI: https://doi.org/10.1016/j.eng.2020.05.011.

16. Soon J. A comparative analysis of legislative protection from harassment: a view from Singapore. Oxford University Commonwealth Law Journal, 2022, vol. 22, no. 2, pp. 177–204. https://doi.org/10.1080/14729342.2022.2109272

17. Young N. Deepfake technology: complete guide to deepfakes, politics and social media. North Charleston (S.C.): Independently published, 2019, 160 p.

References

1. Batoev V.B. (2023) Using "deepfake" technology in operational-search activities. *Juridicheskaya nauka i praktika*. *Vestnik nizhegorodskoi akademii MVD*=Legal Science and Practice: Bulletin of the Nizhny Novgorod Academy of Internal Ministry, no. 1, pp. 70–76 (in Russ.)

2. Blitz M. J. (2020) Deepfakes and other non-testimonial falsehoods: when is belief manipulation (Not) First Amendment speech. *Yale Journal of Law & Technology,* vol. 23, Fall, pp. 161–300.

3. Burnham W. (2006) Legal system of the USA. Moscow: New Justice, 1216 p. (in Russ.)

4. Chng K. (2023) Falsehoods, foreign Interference, and compelled speech in Singapore. *Asian Journal of Comparative Law*, no. 2, pp. 235–252. doi:10.1017/asjcl.2023.9

5. Goodfellow I. J. et al. (2014) Generative adversarial nets. In: Advances in Neural Information Processing Systems. Montreal: University Press, pp. 2672–2680.

6. Hine E., Floridi L. (2022) New deepfake regulations in China are a tool for social stability, but at what cost? *Nature Machine Intelligence*, no. 4, pp. 608–610.

7. Joost L. (2023) Place for illusions: deepfake technology and the challenges of regulating unreality. *University of Florida Journal of Law and Public Policy,* vol. 33, no. 2, pp. 309–332.

8. Köbis N.C., Doležalová B., Soraperra I. (2021) Fooled twice: people cannot detect deepfakes but think they can. *iScience*, vol. 24, issue 11, p. 103364. DOI: https://doi.org/10.1016/j.isci.2021.103364.

9. Kosikhina S.S., Shvets A.V. (2022) Legal system of the PRC: textbook. Blagoveshchensk: Amur State University, 90 p. (in Russ.)

10. Troshchinsky P.V. (2016) Legal system of China. Moscow: Nauka, 472 p. (in Russ.)

11. Lee H., Lee T. (2023) Between two Acts: competing narratives, activism and governance in Singapore's digital sphere. *Internet Histories*, vol. 7, no. 4, pp. 295–312. DOI: https://doi.org/10.1080/24701475.2023.2232214

12. Li Y. (2023) Implementing rule of law concept in the digital sphere: China's experience. *Legal Issues in the Digital Age,* no. 4, pp. 114–133. DOI: https://doi.org/10.17323/2713-2749.2023.4.114.133

13. Manique A.E. et al. (2023) Industry approaches in handling online exploitation of children: a comparative study of policy, guidelines and best practices in Malaysia, Singapore and Australia. *Cogent Social Sciences*, no. 2. DOI: https://doi.org/10.1080/23311886.2023.2241713

14. Pfefferkorn R. (2020) "Deepfakes" in the courtroom. *Boston University Public Interest Law Journal,* vol. 29, no. 2, pp. 245–276.

15. Skibba R. (2020) Media enhanced by artificial intelligence: can we believe anything anymore? *Engineering*, vol. 6, issue 7, pp. 723–724. DOI: https://doi.org/10.1016/j.eng.2020.05.011.

16. Soon J. (2022) A comparative analysis of legislative protection from harassment: a view from Singapore. *Oxford University Commonwealth Law Journal,* no. 2, pp. 177–204. DOI: https://doi.org/10.1080/14729342.2022.2109272

17. Young N. (2019) Deepfake technology: complete guide to deepfakes, politics and social media. North Charleston (S.C.): Independently published, 160 p.

Информация об авторах:

В.А. Виноградов — доктор юридических наук, профессор.

Д.В. Кузнецова — LL.M, ведущий эксперт.

Information about the authors:

V.A. Vinogradov — Doctor of Sciences (Law), Professor.

D.V. Kuznetsova — LL.M, Senior Expert.

Статья поступила в редакцию 06.03.2024; одобрена после рецензирования 30.04.2024; принята к публикации 30.04.2024.

The article was submitted to editorial office 06.03.2024; approved after reviewing 30.04.2024; accepted for publication 30.04.2024.