

Юридические исследования

*Правильная ссылка на статью:*

Бодров Н.Ф., Лебедева А.К. Анализ судебной практики установления обстоятельств в случаях противоправного распространения генеративного контента, созданного с помощью технологий искусственного интеллекта // Юридические исследования. 2024. № 11. DOI: 10.25136/2409-7136.2024.11.72540 EDN: TLSBYY URL: [https://nbpublish.com/library\\_read\\_article.php?id=72540](https://nbpublish.com/library_read_article.php?id=72540)

## **Анализ судебной практики установления обстоятельств в случаях противоправного распространения генеративного контента, созданного с помощью технологий искусственного интеллекта**

**Бодров Николай Филиппович**

ORCID: 0000-0002-9005-3821

кандидат юридических наук

доцент; кафедра судебных экспертиз; Московский государственный юридический университет имени О.Е. Кутафина (МГЮА)

123242, Россия, г. Москва, ул. Садовая-Кудринская, 9, каб. 729

✉ [bodrovnf@gmail.com](mailto:bodrovnf@gmail.com)



**Лебедева Антонина Константиновна**

кандидат юридических наук

доцент; кафедра судебных экспертиз; Московский государственный юридический университет имени О.Е. Кутафина

125993, Россия, г. Москва, ул. Садовая-Кудринская, 9

✉ [tonya109@yandex.ru](mailto:tonya109@yandex.ru)



[Статья из рубрики "Уголовный закон и правопорядок"](#)

**DOI:**

10.25136/2409-7136.2024.11.72540

**EDN:**

TLSBYY

**Дата направления статьи в редакцию:**

02-12-2024

**Дата публикации:**

09-12-2024

**Аннотация:** В статье авторами рассматриваются различные случаи из судебной и следственной практики, связанные с противоправным распространением и использованием генеративного контента, созданного с помощью технологий искусственного интеллекта. Особое внимание уделено вопросам доказывания факта фальсификации цифровых продуктов, созданных нейросетями, включая голосовые и графические дипфейки. Анализируются правовые, технологические аспекты использования таких доказательств в судебных разбирательствах. Авторы также подчёркивают, что определение дипфейка в действующем законодательстве отсутствует, а без этого не представляется возможным говорить о модернизации отечественного законодательства. С учётом темпов развития технологий искусственного интеллекта необходимо законодательно закрепить определение дипфейка. В контексте стремительного развития технологий искусственного интеллекта следует предложить нормативное регулирование дипфейков с учётом правовых пробелов, которые сопровождают современный уровень развития искусственного интеллекта, и угроз, которые уже являются реальными, что подтверждается рассмотренной судебной практикой. Обоснована необходимость создания набора данных для проведения экспериментальных фоноскопических исследований фонограмм с записями, клонированных с использованием нейросетей голосов. Методологическую основу исследования составляют всеобщий диалектический метод, общенаучные (описание, сравнение, обобщение, моделирование и др.) и частнонаучные методы. Новизна исследования заключается в выявлении и систематизации ключевых проблем, связанных с проведением судебных экспертиз и правовым регулированием генеративного контента. В работе представлены рекомендации по совершенствованию законодательных норм и экспертных методик (на примере судебной фоноскопической экспертизы), включая необходимость создания специализированных баз данных и научно-методических подходов для исследования генеративного контента. Выводы статьи подчеркивают значимость разработки стандартов для диагностики факта использования генеративного искусственного интеллекта при создании цифровых продуктов, а также необходимость повышения квалификации экспертов, осуществляющих производства судебных экспертиз в отношении подобных объектов. Полученные результаты могут быть использованы для формирования более эффективных механизмов правового реагирования на вызовы, связанные с технологиями искусственного интеллекта.

**Ключевые слова:**

дипфейки, генеративный контент, судебная практика, экспертная практика, нейросети, клонирование голоса, Судебная экспертология, криминалистика, фоноскопическая экспертиза, набор данных

*Указанная публикация подготовлена в рамках выполнения работ по государственному заданию на тему: «Российская правовая система в реалиях цифровой трансформации общества и государства: адаптация и перспективы реагирования на современные вызовы и угрозы (FSMW-2023-0006)». Регистрационный номер ЕГИСУ НИОКТР 124012000079-6*

В настоящее время технологии искусственного интеллекта (далее – ИИ) для создания генеративного контента развиваются крайне стремительно и задействуются в самых различных областях общественной жизни. Одновременно формируется и практика

рассмотрения судебных дел, в которых в качестве доказательств используются цифровые следы, представляющие собой сгенерированный при помощи ИИ контент.

Данный контент – это цифровые продукты в различном виде: в виде текста, графики, звука или их сочетания.

Безусловно, активное развитие нейросетевых технологий в сфере создания генеративного контента предоставляет огромнейшие возможности для его противоправного использования и распространения. Актуальность выбранной темы обусловлена тем, что всё больше случаев, когда синтезированный контент используется в целях введения в заблуждение или преодоления пользователем систем контроля и управления доступом [\[1, с. 38\]](#), в рамках дел о клевете, различных мошенничествах, при рассмотрении гражданских дел (например, бракоразводный процесс).

В подобных ситуациях перед правоприменителями встаёт задача установления обстоятельств, связанных с созданием, использованием и распространением сгенерированного при помощи нейросетевых технологий контента:

- технологические сложности выявления подобного генеративного контента, доказательства факта его создания как в целом при помощи нейросетевых технологий, так и конкретных нейросетей, использованных для его создания;
- целесообразность назначения судебной экспертизы. С учетом того, что методики большинства судебных экспертиз не адаптированы для исследования объектов – результатов нейросетевой генерации, даже категорический вывод эксперта может не иметь значения для последующего доказывания в том случае если в ходе экспертного исследования не был установлен факт того, что объект исследования действительно является дипфейком, без этого достоверность выводов эксперта в современных реалиях может быть поставлена под сомнение. В настоящее время вопросы судебно-экспертного противодействия дипфейк-контенту анализируются в научных трудах очень небольшого количества отечественных специалистов, например, в публикациях Галяшиной Е. И., Громовой А. В., Ошкукова С. С. и некоторых других;
- неоднозначность судебной практики в квалификации противоправных деяний, связанных с распространением и использованием генеративного контента, в том числе отсутствие унифицированных подходов к определению термина «дипфейк», а также вопросы установления субъекта, ответственного за распространение/использование генеративного контента. Как зарубежное, так и отечественное законодательство, таким образом, пока не адаптировано к вызовам, связанным с использованием нейросетевых технологий.

В настоящее время, как следствие, возникают принципиально новые ситуации, требующие уточнения процедуры доказывания. Проблема распространения дипфейков затрагивает всех участников судопроизводства [\[2, с. 254\]](#). Однако, как нам представляется, центральное место в доказывании по подобным делам будет занимать заключение эксперта, поскольку первоочередной проблемой становится сложность в доказывании того, является ли представленный цифровой продукт, сгенерированный при помощи нейросетевых технологий или нет.

Судебная практика в рассматриваемой сфере ещё только формируется, сталкиваясь безусловно с рядом сложностей, как правового, так и технологического характера.

В данной статье рассмотрена отечественная и зарубежная практика судебного и

следственного установления обстоятельств по делам о противоправном распространении синтезированного контента (дипфейков).

Методологическую основу исследования составляют всеобщий диалектический метод, общенаучные (описание, сравнение, обобщение, моделирование и др.) и частнонаучные методы.

Мы анализируем те, пока немногочисленные, кейсы, когда синтезированный контент использовался в качестве доказательства в гражданских, арбитражных и уголовных делах. По нашему мнению, в ближайшей перспективе объём подобных дел в судопроизводстве будет неуклонно возрастать, а первичный анализ и обобщение специфики подобных дел смогут быть использованы для предотвращения судебных ошибок.

Важно рассмотреть следственные и судебные ситуации, связанные с установлением обстоятельств распространения синтезированного контента, а также обобщить практику доказывания по делам, где дипфейки выступали в качестве орудия или предмета противоправной деятельности.

Полагаем, что рассмотрение проблематики установления обстоятельств противоправного распространения синтезированного контента следует начать с анализа весьма показательного случая раскрытия преступления за рубежом.

### **1. Уголовное дело по обвинению директора школы**

Один из кейсов, в котором в качестве доказательств был предоставлен дипфейк, произошёл в США. Директор средней школы Пайксвилла Эрик Эйсверт стал подозреваемым по уголовному делу о диффамации после того, как в социальных сетях была распространена фонограмма с записью его голоса с различными расистскими и антисемитскими комментариями [Бывший учитель физкультуры, подставивший директора школы с помощью ИИ, был задержан в аэропорту с оружием. URL: <https://www.thebaltimorebanner.com/education/k-12-schools/eric-eiswert-ai-audio-baltimore-county-YBJNJAS6OZEE5OQVF5LFOFYN6M/> (дата обращения: 28.11.2024)] Дипфейки все чаще используются при осуществлении диффамации, как указывает, например, Парыгина Н. Н. «Создать ложное представление о том, что тот или иной человек что-либо сказал или сделал, теперь беспрецедентно легко» [3, с. 178].

Распространение данной фонограммы привело к ostracism директора от работы в школе, послужило причиной общественного порицания в адрес директора в социальных сетях. В адрес директора и его семьи поступали письма с угрозами, всё это, безусловно, способствовало умалению его деловой репутации и привело к дискредитации его как личности и как директора школы.

Спорная фонограмма была размещена в аккаунте в социальной сети Инстаграм [Запрещённая в России социальная сеть]. Впоследствии экспертным путём удалось установить, что данная фонограмма являлась сгенерированной при помощи технологий ИИ. В экспертном заключении было указано, что голос на фонограмме являлся клонированным голосом директора. В ходе следствия было установлено, что спорная фонограмма распространялась с использованием недавно зарегистрированного анонимного адреса электронной почты.

Установить принадлежность адреса электронной почты удалось благодаря корректно использованной следователем тактической операции. После установления круга

подозреваемых в клонировании голоса компьютерно-технические специалисты были привлечены к изучению логов компьютеров и учётных записей, зарегистрированных в школьной сети.

В учётной записи учителя физкультуры, Дажона Дэрьена, была обнаружена информация о поисковых запросах о «больших языковых моделях» (БЯМ), нейросетях для распознавания голоса и создания клонов голоса. Даты поисковых запросов непосредственно предшествовали дате опубликования фонограммы.

Для установления принадлежности адреса электронной почты Д. Дэрьену в ходе его допроса следователь зашёл на почтовый сервер, где зарегистрирован адрес электронной почты и воспользовался функцией «восстановить пароль». После чего на телефон Д. Дэрьена пришло сообщение со ссылкой, необходимой для восстановления пароля электронной почты [Дело с использованием сгенерированного при помощи ИИ голоса директора школы. URL: <https://www.wbaltv.com/article/police-former-pikesville-high-ad-facing-charges-ai-generated-voice-hoax/60603313>(дата обращения: 28.11.2024)].

С помощью ряда экспертиз и исследований, проведённых в Калифорнийском университете (University of California, Berkeley) [Обвинение директора школы. Дело с использованием ИИ. URL: <https://www.baltimorecountymd.gov/departments/police/news/2024/04/25/athletic-director-charged-in-pikesville-high-school-ai-case> (дата обращения: 28.11.2024)] было установлено, что сгенерированная фонограмма была создана при помощи технологий ИИ и распространена учителем физкультуры, Дажоном Дэрьеном, в январе 2024 года.

На спорной фонограмме среди прочих лиц, в адрес которых были сделаны расистские заявления, было и прозвище данного учителя физкультуры «DJ». В отношении учителя физкультуры ранее было начато расследование по делу о краже в школе после заявления, составленного директором школы, а сгенерированная фонограмма, таким образом, послужила своеобразной местью учителя физкультуры директору.

Данное происшествие привело к серьёзным волнениям в школе, аудиозапись ввела в заблуждение значительный круг лиц, сильно повлияла на доверие учащихся и родителей, администрации школы. Прокурор округа Балтимор заявил, что это первое дело, когда доказательство было сфальсифицировано при помощи технологий ИИ, в результате чего делом занимался весь округ, так как сложно было найти экспертов, компетентных в вопросах исследования подобного синтезированного контента.

В рассматриваемой ситуации проблема связана с тем, что само право на реализацию судебной защиты своей чести, деловой репутации и доброго имени зависит от возможности субъекта доказать, что реалистичное убедительное цифровое доказательство является сфальсифицированным с помощью нейросетевых технологий.

Возможности доказывания в описанной ранее следственной ситуации существенно ограничены. Во-первых, Дажон Дэрьен синтезировал монологическую речь Эрика Эйсверта, что лишило последнего возможности использовать показания свидетелей для защиты от обвинений. На фонограмме не были зафиксированы реплики кого-то из собеседников Эрика Эйсверта, которые могли бы дать свидетельские показания о том, что он не произносил в разговоре с ними подобных высказываний.

Во-вторых, Дажон Дэрьен распространил информацию с использованием рассылки, а заявителями о правонарушении Эрика Эйсверта выступили получатели такой рассылки, для которых качество синтеза голоса и речи оказалось достаточно убедительным.

В-третьих, единственным источником информации о подозреваемом стала общедоступная информация о конфликте Дажона Дэрьяна и Эрика Эйсверта. Если бы конфликт не был публичным или иной злоумышленник распространил бы сгенерированную фонограмму до возникновения какого-либо конфликта, в распоряжении следствия и самого Эрика Эйсверта вряд ли бы оказалась информация, достаточная для определения круга лиц, потенциально причастных к синтезу и распространению фонограммы.

В-четвертых, для доказывания факта фальсификации необходимо проведение диагностического фоноскопического исследования. В случае досудебного установления обстоятельств на лицо, обратившееся за помощью специалиста-фоноскописта, возлагаются достаточно существенные финансовые расходы. Обращение за досудебным исследованием в такой высокотехнологичной сфере несёт за собой достаточно серьёзные финансовые издержки, а отсутствие методик исследования и малая результативность исследований результатов высокотехнологичного синтеза фонограмм обуславливает весьма значительные финансовые риски для заявителя даже в тех ситуациях, когда он абсолютно уверен в том, что фонограмма была синтезирована.

Вопрос оплаты за производство судебной экспертизы может в ряде случаев быть существенной преградой для реализации права на защиту от обвинения или возражения на иск. Так, например, в отечественной гражданско-процессуальной практике расходы на оплату судебной экспертизы возлагаются судом, как правило, на сторону, ходатайствовавшую о её производстве [\[4\]](#). Согласно положениям ст. 96 ГПК РФ, денежные суммы, подлежащие выплате экспертам, предварительно вносятся на счёт, открытый в порядке, установленном бюджетным законодательством РФ, стороной, заявившей соответствующую просьбу.

В уголовном судопроизводстве финансовые расходы часто возникают ещё на стадии внепроцессуальных исследований, когда для обоснования заявления требуется заключение специалиста, оплата подготовки которого происходит за счёт собственных средств заявителя.

Производство судебной фоноскопической экспертизы также связано с некоторыми рисками. Во-первых, в ситуации, когда результатом экспертного исследования станет вывод в форме НПВ (не представляется возможным ответить на поставленный судом или следователем вопрос), заявитель может быть фактически лишён возможностей последующего доказывания.

Во-вторых, в ситуации, когда судебно-экспертные методики не содержат соответствующих разделов, посвящённых выявлению признаков синтеза, шанс выявить сам факт синтеза весьма мал, так как у эксперта при производстве исследования не будет методического подспорья для выявления и описания каких-либо типовых признаков генерации или обстоятельств, подлежащих установлению при проверке версии о синтезе.

В-третьих, в ситуации, когда экспертная практика по такого рода делам ещё не сформировалась, логично предположить, что выявление факта генерации окажется под силу только наиболее высококвалифицированным экспертам, которых очевидно не так много среди всех практикующих работников.

## **2. Уголовное дело по обвинению учёного-гидролога**

Широкую известность в России получило резонансное дело, в котором в качестве обвиняемого в убийстве был задержан учёный-гидролог Александр Цветков. Учёного

обвиняли в убийстве 20-летней давности. В 2023 году А. И. Цветков был задержан в аэропорту Домодедово в связи с тем, что сработала система распознавания лиц, которая нашла частичное (на 55%) сходство с фотороботом предполагаемого убийцы, составленным ещё в 2002 году.

При этом статистические значения 55% в научной сфере являются абсолютно нерелевантными, так как даже на конкурсные работы по тематике распознавания лиц принимаются работы с результативностью от 90 процентов, а занимают призовые места как правило работы со значениями 99.6 и выше [5].

В результате А. И. Цветков находился под арестом 11 месяцев, хотя принципиальная невозможность его нахождения на месте преступления была очевидна на основе иных установленных следствием обстоятельств по делу.

Совершенно очевидно, что при расследовании преступлений прошлых лет ошибочные результаты идентификации с использованием нейросетевых алгоритмов будут выступать в качестве убедительных, а часто и неопровержимых подтверждений причастности лиц, у которых просто не окажется возможности опровержения доводов обвинения.

Среди таких лиц, как показывают результаты зарубежных исследований [6, с. 6; 7], в первую очередь оказываются мигранты, расовым и национально-специфическим признакам внешности которых присвоены неправильные статистические веса в массиве, где из-за малой репрезентативности их доли в общей выборке формируются явная предвзятость нейросетевых алгоритмов.

Обстоятельствами, послужившими причиной прекращения уголовного преследования в отношении А.И. Цветкова также послужили результаты экспертных исследований.

По официальному заявлению уполномоченных лиц Главного следственного управления СКР по Москве: «В ходе расследования уголовного дела следствие тщательно проверило собранные доказательства, назначило и провело судебные экспертизы, которые установили непричастность Цветкова к инкриминируемым преступлениям. В связи с чем уголовное преследование в отношении Цветкова прекращено» [Встретила свобода: как доказали невиновность ученого в убийствах 20-летней давности. URL: <https://iz.ru/1661286/elena-balaian/vstretila-svoboda-kak-dokazali-nevinovnost-uchenogo-v-ubiistvakh-20-letnei-davnosti> (дата обращения: 28.11.2024)].

### **3. Обвинения в преступлениях и административных правонарушениях по результатам вмешательства в выборы**

22 января 2024 года Генеральная прокуратура штата Нью-Гэмпшир начала расследование по факту рассылки автоматических голосовых сообщений (так называемых «робозвонков»), полученных тысячами жителей штата. Сообщения содержали ложную информацию о необходимости «сохранить свой голос для ноябрьских выборов» и утверждали, что «ваш голос имеет значение в ноябре, а не в этот вторник». Голос президента США Д. Байдена в записанных сообщениях был клонирован с использованием нейросетевого алгоритма. Кроме того, было установлено, что сообщение было результатом спуфинга, то есть было замаскировано под исходящее от казначея политического комитета, поддерживавшего кампанию Байдена на праймериз Демократической партии США.

В ходе расследования было установлено, что организатором роботизированного обзвона (типовая процедура электорального процесса в США) выступал гражданин С. Крамер.

Заявления поступили от тринадцати потерпевших. Ему были предъявлены обвинения в соответствии с двумя статьями законодательства штата Нью-Гэмпшир:

13 тяжких преступлений по статье RSA 659:40, III (уголовное нарушение, связанное с подавлением избирательной активности). Статья гласит: «Никто не должен заниматься подавлением избирательной активности, сознательно пытаясь предотвратить или удержать другого человека от голосования или регистрации для голосования на основании мошеннической, обманчивой, вводящей в заблуждение или ложной информации». С. Крамеру инкриминируется использование клонированного с использованием нейросетевых технологий голоса Президента США и сообщение ложной информации с целью отговорить 13 избирателей от участия в праймериз 23 января 2024 года.

13 административных правонарушений по статье RSA 666:7-а (имитация кандидата). Статья устанавливает ответственность за ложное представление себя кандидатом в телефонном звонке. С. Крамер был обвинён в том, что лично или с привлечением других лиц совершал звонки 13 избирателям, выдавая себя за кандидата.

В результате по решению Федеральной комиссии по связи США телекоммуникационная компания Lingo Telecom, оборудование которой использовалось для совершения роботизированных звонков, выплатила штраф, а в отношении С. Крамера были возбуждены уголовные дела. Федеральная комиссия по связи США также выдвинула обвинения в отношении С. Крамера, оценив размер штрафа в 6 миллионов долларов [Consultant behind deepfaked Biden robocall fined \$6m as new charges filed. US elections 2024. Retrieved from <https://www.theguardian.com/us-news/article/2024/may/23/biden-robocall-indicted-primary> (дата обращения: 28.11.2024)]. Следствие по делу продолжается.

Данный кейс демонстрирует ситуацию, в которой правонарушители достигают мгновенного (как часто бывает в случае с выборами) и масштабного эффекта. В такой ситуации судебно-экспертное установление обстоятельств, даже в случае результативных выводов эксперта, будет с учётом сроков производства экспертиз и исследований неактуально, так как результат противоправного посягательства подобными методами не может быть предотвращён.

#### **4. Дипфейк-фонограмма в качестве доказательства в бракоразводном процессе**

Дипфейки предоставляются в качестве доказательств не только в рамках уголовных дел, но и по гражданским делам. Показательным случаем является, например, дело о правах на опеку над ребёнком, которое рассматривалось в Великобритании. В ходе бракоразводного процесса жена в качестве доказательства представила фонограмму, на которой запечатлён голос и речь мужа, угрожавшего жизни и здоровью жены и их ребёнка. Подобного рода доказательство, безусловно, уменьшало шансы супруга на опеку над ребёнком.

Назначенная судом фоноскопическая экспертиза впоследствии смогла установить, что после записи в файл были внесены изменения: в неё были вставлены синтезированные фрагменты [P. Ryan. Deepfake Audio Evidence used in U.K. Court to Discredit Dubai Dad. The National UAE. Retrieved from <https://www.thenationalnews.com/uae/courts/deepfake-audio-evidence-usedin-uk-court-to-discredit-dubai-dad-1.975764> (дата обращения: 28.11.2024)].

Фрагменты фонограммы с угрозами жена сгенерировала на основе голосовых сообщений



мужа в мессенджерах. В качестве массива данных для генерации по заданным параметрам (в терминологии нейросетевой генерации – клонирования) были использованы голосовые сообщения, большой объём которых позволил добиться высокого уровня сходства голоса и речи.

Ключевыми обстоятельствами, обуславливающим процесс и результаты доказывания по таким делам, фактически выступают компетентность злоумышленника и наличие у него/неё большого объёма образцов клонируемого голоса. То есть одним из инструментов выявления потенциально причастных к генерации лиц может стать решение экспертом-фоноскопистом диагностической задачи о том какой набор данных (например, по его объёму или амплитудно-частотным характеристикам фонограмм-образцов) был использован для генерации.

Немаловажную роль в установлении обстоятельств генерации играет также и осведомлённость эксперта о специфических особенностях современных сервисов генерации/клонирования голоса с целью ответа на вопрос о том, с использованием какой программы был синтезирован/клонирован тот или иной голос.

Представляется, что для разработки методик решения подобных диагностических задач в первую очередь необходимы организация и ведение актуальных фонотек с результатами генерации. Данное направление требует отдельной теоретической и практической разработки.

## **5. Судебная практика защиты исключительных прав на цифровые продукты**

Первым известным нам случаем из российской судебной практики является судебное разбирательство по арбитражному делу о взыскании компенсации за нарушение исключительного права на видеоролик в размере между ООО «Рефейс технолоджис» и ООО «Бизнес-Аналитика» [Материалы дела № А40-200471/23-27-1448. URL: <https://kad.arbitr.ru/Card/4d7f0305-69af-44fe-8841-a59e84aa7deb> (дата обращения: 28.11.2024)].

Судом было установлено, что ООО «Бизнес-Аналитика» в сети Интернет разместило аудиовизуальное произведение продолжительностью 31 секунду, исключительное право на которое принадлежит ООО «Рефейс технолоджис».

Ответчик обосновывал отсутствие исключительного права у истца на данный объект отсутствием доказательств наличия исключительного права на спорный видеоролик, а также тем, что «созданный видеоролик не является объектом авторского права по причине использования технологии Deep-fake».

Суд данные доводы отклонил, указав, что «поскольку технология Deep-fake – это дополнительный инструмент обработки (технического монтажа) видеоматериалов, а не способ их создания». Соответственно осуществление технического монтажа «исходных материалов видеоролика посредством технологии deep-fake сам по себе не свидетельствует о том, что видеоролик доступен для свободного использования (без согласия правообладателя), или о том, что группа лиц, обеспечившая написание сценария видеоролика, видеосъемку, его аудиосопровождение, не внесли личный творческий вклад в создание видеоролика и не признаются его авторами».

Однако, в связи с тем, что в настоящее время отсутствует законодательное определение дипфейка, данная позиция нам представляется не совсем верной. В опубликованных ранее рецензируемых научных статьях мы указывали, что «Дипфейк – цифровой продукт

в виде текста, графики, звука или их сочетания, сгенерированный полностью или частично при помощи нейросетевых технологий, для цели в ведения в заблуждение или преодоления пользователем систем контроля и управления доступом» [\[1, С. 38\]](#).

Таким образом, исходя из сути рассматриваемого дела спорное аудиовизуальное произведение не являлось дипфейком, а является сгенерированным цифровым продуктом, созданным с помощью нейросетевых технологий. Компания не скрывала способ его создания, противоправных целей, распространяя его, не имела. Кроме того, дипфейк – не технология, а результат использования нейросетевых технологий и не является инструментом технического монтажа.

Таким образом, ещё одной правовой проблемой является вопрос признания прав на сигналограммы, сгенерированные при помощи нейросетевых технологий. В данном кейсе суд решил признать право ООО «Рефейс технолоджис» нарушенным и взыскало с ООО «Бизнес-аналитика» компенсацию [Решение Арбитражный суд г. Москвы в составе судьи Крикуновой В.И. от 30 ноября 2023 г. по делу № А40-200471/23-27-1448 URL: <https://kad.arbitr.ru/Card/4d7f0305-69af-44fe-8841-a59e84aa7deb> (дата обращения: 28.11.2024)].

## **6. Категория дел о телефонных мошенничествах с использованием нейросетевых технологий**

Различные виды дипфейков повсеместно используются при совершении телефонных мошенничеств и мошенничеств с использованием мессенджеров. Особенности подобного рода преступлений уже получили рассмотрение в научной литературе. Например, в одной из публикаций рассматриваются некоторые мошеннические схемы, при которых «мошенники выдают себя за руководителей министерств и ведомств, ректоров и других высокопоставленных чиновников» [\[8, С. 25\]](#). Осуществляются данные схемы с использованием как голосовых, так и графических дипфейков. Реализуются подобные схемы как с целью неправомерного получения денежных средств от жертвы, так и для «кражи» персональных данных.

Активно используются как дипфейки, созданные на основе биометрических персональных данных близких жертве лиц (родственников, друзей, знакомых), так и высокопоставленных чиновников, популярных персон. Например, в одном из недавних дел мошенники создали телеграм-канал от имени популярного блогера Оксаны Самойловой, размещали в нем различные голосовые сообщения и видеозаписи от её имени (голосовые и графические дипфейки) с предложением заработать, в результате чего граждане, доверившись популярному блогеру, переводили значительные суммы денежных средств. В настоящий момент пострадавшие уже обратились в полицию с заявлением о мошенничестве [Сеют панику и подрывают веру в человечество. Россиян атакуют дипфейки URL: <https://regnum.ru/article/3909268>(дата обращения: 28.11.2024)].

## **7. Практика Верховного суда РФ**

Пленум Верховного суда РФ также обратил своё внимание на проблему дипфейков, упомянув в своих Постановлениях: «вводящие в заблуждение поддельных изображений, аудио- и аудиовизуальных материалов, в том числе созданных с помощью компьютерных технологий» (п.111), «распространение вводящих в заблуждение избирателей, участников референдума поддельных изображений, аудио- и аудиовизуальных материалов, в том числе созданных с помощью компьютерных технологий» (п.129) [Постановление Пленума Верховного Суда Российской Федерации № 24 от 27 июня 2024

г. «О некоторых вопросах, возникающих при рассмотрении судами административных дел о защите избирательных прав и права на участие в референдуме граждан Российской Федерации»]; «с использованием вводящих в заблуждение и выдаваемых за достоверные недостоверных изображений, аудио- и аудиовизуальной информации, в том числе созданных с помощью компьютерных технологий (п.20)» [Постановление Пленума Верховного Суда Российской Федерации № 17 от 25 июня 2024 г. «Об отдельных вопросах, возникающих у судов при рассмотрении дел об административных правонарушениях, посягающих на установленный порядок информационного обеспечения выборов и референдумов»].

Председатель судебного состава Судебной коллегии по административным делам Верховного суда России С. Кузьмичев, представляя Постановление перед голосованием [Верховный суд назвал использование дипфейков на выборах нарушением правил агитации. URL: <https://tass.ru/politika/21193603> (дата обращения: 18.11.2024)], а в след за ним и средства массовой информации сделали заявления о том, что в данных Постановлениях дано определение дипфейка.

В законопроекте, предусматривающем внесение в некоторые статьи Уголовного кодекса РФ дополнительного квалифицированного состава, предлагающего установить использование дипфейка в качестве квалифицирующего признака, также даётся сомнительная с практической точки зрения формулировка: «с использованием изображения или голоса (в том числе фальсифицированных или искусственно созданных) потерпевшего или иного лица, а равно с использованием биометрических персональных данных потерпевшего или иного лица» [Проект федерального закона № 718538-8 «О внесении изменений в Уголовный кодекс Российской Федерации» (внесён 16.09.2024 г. депутатом Государственной Думы Я. Е. Ниловым, Сенатором Российской Федерации А. К. Пушковым) // Система обеспечения законодательной деятельности. URL: <https://sozd.duma.gov.ru/bill/718538-8> (дата обращения: 24.11.2024)]. В заключении на проект данного федерального закона также говорится о неопределённости используемой в законопроекте формулировки «дипфейка»: «Используемая в законопроекте формулировка «фальсифицированные или искусственно созданные» является неопределённой и может привести к неоднозначному толкованию проектируемых положений» [Заключение на проект федерального закона № 718538-8 от 08.11.2024 г. URL: <https://sozd.duma.gov.ru/bill/718538-8> (дата обращения: 24.11.2024)]. Однако как сообщается в СМИ, данный законопроект «концептуально поддержан Верховным судом РФ и правительством России» [Кабмин и Верховный суд поддержали законопроект о наказании за дипфейки. URL: <https://tass.ru/ekonomika/22401345> (дата обращения: 24.11.2024)].

Однако отсутствие определения «дипфейка» в российском законодательстве может крайне негативно влиять на правоприменительную практику. Предложенная дефиниция дипфейка не позволяет однозначно квалифицировать действия, связанные с его использованием и распространением.

В ранее опубликованных нами статьях мы описывали перечень потенциальных угроз использования и распространения дипфейков [9, С. 174-176]. Данные кейсы прямо подтверждают, что описанные угрозы уже стали реальными, различные обстоятельства дел, в которых преступления были совершены с использованием дипфейков все чаще описываются и в научных статьях [10].

Результаты обобщения информации о рассмотренных в настоящей статье делах подтверждают не только необходимость разработки правовых механизмов регулирования

генеративного контента, но и совершенствования судебно-экспертного противодействия для борьбы с дипфейками. Но на современном этапе судебно-экспертное обеспечение как в нашей стране, так и за рубежом значительно отстаёт от темпов развития генеративных нейросетей. При этом в научном плане при производстве судебно-экспертных исследований экспертам приходится ориентироваться на иностранные научные работы, в которых на сегодняшний день уделяется значительно большее внимание созданию конкретных алгоритмов для обнаружения дипфейка [\[11, 12\]](#).

**8. Судебно-экспертное обеспечение по делам о противоправном распространении генеративного контента, созданного с помощью технологий искусственного интеллекта**

В ситуации противоправного распространения голосовых дипфейков необходимо задействовать инструментарий судебной фоноскопической экспертизы [\[13, 14, 15\]](#). При этом помимо методического обеспечения судебной фоноскопической экспертизы требуется ещё и правильная постановка вопросов, формулируемых в постановлении или определении.

В первую очередь правоприменитель должен обладать навыками формулирования экспертной задачи, особенно на современном этапе разработанности методического обеспечения, когда консультационная помощь сведущих лиц может быть получена только от наиболее квалифицированных специалистов, которые, как мы отмечали ранее, ожидаемо составляют меньшинство от общего числа профессионалов в области судебной фоноскопии.

Классической и знакомой большинству правоприменителей формулировкой для технического исследования фонограмм является вопрос: «Имеются ли в представленной фонограмме «Название фонограммы.mp3» признаки неситуационных изменений, в том числе монтажа и др.?».

Однако в случае со сгенерированным контентом изменения не вносятся в какую-либо фонограмму (хотя, как мы описывали ранее, и такие случаи возможны). Важно понимать, что в ситуациях с предоставлением в качестве доказательства дипфейка речь чаще всего идёт о создании новой фонограммы, отражающей речевое событие, которого никогда не существовало.

Таблица 1. Формулировки вопросов, выносимых на судебную фоноскопическую экспертизу

№ №	Существующая формулировка из методических рекомендаций	Некорректные формулировки	Предлагаемые формулировки
1.	Имеются ли в представленной фонограмме «Название фонограммы.mp3» признаки неситуационных изменений, в том числе монтажа и др.?	Имеются ли в представленной фонограмме «Название фонограммы.mp3» признаки нейросетевой генерации аудиосигнала?	Является ли голос и звучащая речь, записанные на фонограмме «Название фонограммы.mp3» синтезированными с использованием нейросетевых алгоритмов?

2.			была ли фонограмма «Название фонограммы.mp3» получена путем добавления/объединения фрагментов, которые синтезированы с использованием нейросетевых алгоритмов?
3.	Каковы технические характеристики и параметры устройства, с использованием которого производилась звукозапись?	Какие свойства, характеристики и параметры имеет фонограмма «Название фонограммы.mp3» на представленном носителе информации?	Каким программным обеспечением (или ПО какого устройства) создан файл «Название фонограммы.mp3»?

При этом с практической точки зрения стоит учесть, что вопрос №3 «Каким программным обеспечением (или ПО какого устройства) создан файл «Название фонограммы.mp3»?» в Таблице 1 с учётом специфики объекта исследования находится на стыке компетенций технического исследования фонограмм и компьютерно-технической экспертизы.

В целях разработки научно-методической базы для производства судебной фоноскопической экспертизы в отношении синтезированной речи, в том числе клонированных голосов, принципиально важной задачей является создание специализированного массива данных (датасэт), который будет включать в себя репрезентативные выборки с фонограммами как клонированных голосов, так и тех, на которых нейросеть была обучена. Данный подход, по нашему мнению, позволит выработать подходы к детекции дипфейков методами судебной фоноскопии. В рамках Национальной стратегии развития искусственного интеллекта сущность понятия «датасэт» отражает термин «набор данных», оно было уточнено в последней редакции Указа Президента РФ:

«д) набор данных - состав данных, которые структурированы или сгруппированы по определенным признакам, соответствуют требованиям законодательства Российской Федерации и необходимы для разработки программ для электронных вычислительных машин на основе искусственного интеллекта» [В редакции Указа Президента Российской Федерации от 15.02.2024 № 124. Указ Президента РФ от 15.02.2024 № 124 «О внесении изменений в Указ Президента Российской Федерации от 10 октября 2019 г. N 490 «О развитии искусственного интеллекта в Российской Федерации» и в Национальную стратегию, утверждённую этим Указом»].

В контексте судебно-экспертной деятельности подобный «набор данных» можно рассматривать как специализированный вид фонотеки, предназначенный для решения задач судебно-экспертного исследования. На сегодняшний день отсутствие подобных фонотек является значительным препятствием для развития данного направления в судебно-экспертных учреждениях и образовательных организациях. Данный фактор, безусловно, замедляет разработку и внедрение новых научно-методических подходов к исследованию синтезированной звучащей речи, а также затрудняет подготовку и

повышение квалификации экспертов-фоноскопистов и специалистов других экспертных специальностей. Создание таких фонотек могло бы стать важным шагом на пути к совершенствованию экспертных технологий и повышению их надёжности в условиях активного внедрения технологий ИИ в процесс генерации голосовых данных.

На данный момент алгоритмы обнаружения спуфинг атак базируются на нейросетевых технологиях, которые проходят обучение на ограниченных наборах данных. В качестве таких данных используются фонограммы, сгенерированные с использованием конкретных систем синтеза. Тем не менее, стремительное развитие нейросетевых технологий создаёт значительные сложности для разработчиков аппаратно-программных решений. Они не успевают адаптировать свои системы и своевременно внедрять обновления, что негативно сказывается на эффективности их работы. В результате автоматизированные методы детекции спуфинга, направленные на выявление синтезированной речи, пока не достигают высокой точности, что подчёркивает необходимость разработки алгоритмов, не привязанных к конкретной нейросети, и совершенствования инструментов для анализа подобных фонограмм.

В ноябре 2024 года глава Минцифры заявил об обсуждении возможности создания базы биометрических образцов телефонных мошенников в целях борьбы с ними [Минцифры хочет начать формировать базу биометрии телефонных мошенников. URL: <https://tass.ru/obschestvo/22440773> (дата обращения: 28.11.2024)].

В своём заявлении М. И. Шадаев не конкретизировал, какие именно биометрические данные планируется собирать, однако предполагаем, что с учётом характера телефонных мошенничеств (в том числе аудио, видео-звонки в различных мессенджерах), речь должна идти об образцах голоса и внешнего облика злоумышленников.

Однако эффективность данной меры в современных условиях вызывает определённые сомнения. Во-первых, первоначально требуется разработать критерии пригодности для добавления в массив подобных образцов биометрических данных. Так, например, большая часть акустических сигналов, прошедших через телефонный тракт, характеризуется весьма ограниченными параметрами качества фиксации особенностей голоса и речи, чем нередко пользуются телефонные мошенники.

Во-вторых, в связи с тем, что крайне часто мошенники используют голоса и изображения, созданные при помощи технологий ИИ, первоначально необходимо определить, является ли спорная сигналограмма, которую планируется добавить с условную «фоно/видеотеку» синтезированной или нет. На современном этапе как мы уже указывали ранее это крайне затруднительно.

Кроме того, непонятно как будет функционировать данная система, будут ли привлекаться специалисты, обладающие специальными знаниями в области фоноскопической, видеотехнической или портретной экспертизах, для её создания и индексации данных в ней. Даже если по этому вопросу будет принято положительное решение, открытым остаётся вопрос, могут ли подобные объёмы данных быть проанализированы силами действующих сотрудников судебно-экспертных учреждений и как соотносить эти трудозатраты с необходимостью выполнения их основных функций по производству судебных экспертиз и исследований.

Возможности поиска в различных фонотеках, если говорить о неустановленных лицах, с учётом современного уровня науки и техники крайне ограничены. Стоит отметить, что МВД России в 2018 году отказалось от ведения Учёта фонограмм речи (голоса) неустановленных лиц с 1 ноября 2018 года [Приказ МВД РФ от 11.09.2018 № 585 «О

внесении изменений в приказ МВД России от 10 февраля 2006 г. № 70 «Об организации использования экспертно-криминалистических учетов органов внутренних дел российской федерации» // СПС Консультант плюс]. Имея фонограмму с голосом неизвестного лица и фонотеку с биометрическими данными (в данном случае голос человека) телефонных мошенников крайне затруднительно осуществить поиск в целях сравнения имеющихся образцов голоса со спорной фонограммой.

Например, специализированное программное обеспечения для проведения сравнения дикторов Икар лаб 3 позволяет на основе алгоритма автоматического сравнения в режиме «1:N» осуществлять поиск целевого диктора во множестве заданных фонограмм. Модуль позволяет провести также автоматическую оценку технических характеристик сравниваемых фонограмм. Однако максимальное количество фонограмм (параметр N) – 50.

А предполагается, что база образцов биометрических данных телефонных мошенников будет значительно большего размера, что требует разработки принципиально нового подхода к анализу информации. Следует отметить, что в современных системах дактилоскопической регистрации работа алгоритмов происходит с гораздо меньшими объемами данных, чем файлы фонограмм, поэтому на современном этапе трудно представить себе какой аналог структурирования, разметки и поиска информации в базе данных может быть использован для фонограмм или видеофонограмм.

Кроме того, создание подобной базы ставит под угрозу и конституционные права граждан: статья 23 Конституции РФ [Конституция Российской Федерации (принята всенародным голосованием 12.12.1993 г. с изменениями, одобренными в ходе общероссийского голосования 01.07.2020 г.)] «2. Каждый имеет право на тайну переписки, телефонных переговоров, почтовых, телеграфных и иных сообщений. Ограничение этого права допускается только на основании судебного решения».

Для записи телефонных переговоров, даже с предполагаемыми телефонными мошенниками оператору сотовой связи необходимо будет получить согласие абонента, что с практической точки зрения создаст препятствие для оперативной и эффективной работы проектируемой системы.

Кроме того, в связи с тем, что пока нет никаких подробностей о реализации проекта, необходимо иметь представление, как и на каких основаниях биометрические данные телефонных мошенников будут заноситься в предполагаемую базу. Очевидно, что в действующей системе правовой регламентации данных правоотношений это возможно после вступления в силу приговора суда, что заранее увеличивает отставание таких данных от современных темпов технологического развития нейросетевых алгоритмов.

Таким образом, создание подобной базы биометрических данных телефонных мошенников пока вызывает больше вопросов, как технического, так и правового характера, чем решений уже накопившихся проблем.

По нашему мнению, необходимо создавать фонотеки не образцов предполагаемых мошенников, а фонотеки с результатами экспериментальных фоноскопических исследований. Для обеспечения судебного-экспертного противодействия распространению дипфейк-контента необходима выработка единых научно-методических подходов к исследованию сгенерированных при помощи нейросетевых технологий акустических сигналов (как с голосом и речью человека, так и с шумами фоновой обстановки, которые уже сейчас используются мошенниками для маскирования реальной обстановки дистантной коммуникации).

## Выводы

По результатам проведённого анализа установлено, что стремительное развитие технологий искусственного интеллекта, в частности генеративных нейросетей, приводит к новым вызовам для правоприменения и судебно-экспертной деятельности. Обобщение отечественной и зарубежной судебной практики, связанной с фактами предоставления дипфейков и других результатов нейросетевой генерации в качестве доказательств, продемонстрировало сложность и многоаспектность задач, стоящих перед экспертами и в последующем – перед правоприменителями.

В качестве ключевых проблем уже сейчас кристаллизировались в правоприменительной практике:

### 1. Неопределённость правового регулирования.

Отсутствие унифицированного определения дипфейка и регламентации генеративного контента создаёт правовую неопределённость. Это осложняет квалификацию противоправных действий, связанных с использованием дипфейков, что в последующем будет неизбежно затруднять доказывание обстоятельств в суде.

### 2. Технологическая сложность исследования.

Существующие экспертные методики не всегда позволяют достоверно определить факт генерации контента нейросетевыми алгоритмами. Это особенно актуально для фоноскопической экспертизы и анализа синтезированных/клонированных голосов.

### 3. Ограниченный доступ к технологическим возможностям для исследования.

Технологическая, финансовая и организационная доступность экспертиз и исследований остаётся ограниченной, что будет препятствовать реализации права на защиту и справедливое судебное разбирательство.

Несмотря на очевидно малый на данном этапе объем подобных дел, прогнозируя меры, необходимые для более эффективной правоприменительной и экспертной деятельности, можно сформулировать некоторые рекомендации, актуальные уже сейчас:

- Разработка научно-методической базы.

Необходимо создать специализированные фонотеки и датасеты с примерами синтезированного контента и доступом к образцам, использованным для генерации. Это позволит обеспечить подготовку и повышение квалификации экспертов, а также выработать новые методические подходы к выявлению дипфейков. Организационно это приведет и к упрощению таких исследований, и к снижению затрат на их производство.

- Междисциплинарное взаимодействие.

Важной задачей на первом этапе становится координация работы правоприменителей, экспертов и технических специалистов в области технологий искусственного интеллекта для создания универсальных алгоритмов детекции дипфейков, не зависящих от конкретных технологий генерации.

- Совершенствование законодательства.

Требуется внедрение чётких правовых дефиниций и механизмов регулирования использования генеративных технологий. Особое внимание следует уделить квалификации противоправных действий, совершённых с использованием дипфейков.



В качестве конкретных изменений в уголовно-процессуальное законодательство наиболее своевременными нам представляются следующие:

1. Вектор цифровизации уголовного процесса должен учитывать сам факт того, что в качестве доказательств в цифровой форме будут предоставляться подложные сведения в виде конкретных объектов – дипфейков. Так, на определенном этапе технологического развития и фактической доступности высокореалистичных технологий нейросетевой генерации вполне целесообразным представляется дополнение ст. 196 УПК РФ пунктом 6 следующего содержания:

«6) факт генерации цифровых данных, представленных в качестве доказательства по уголовному делу, когда возникает сомнение в их достоверности или необходимо установление факта использования биометрических персональных данных конкретного лица в процессе такой генерации».

2. Важным аспектом является создание условий для безопасного хранения, обработки и исследования синтезированного контента, чтобы исключить его видоизменение или использование в иных противоправных целях, когда доступ к нему не будет ограничен.

В связи с этим логично дополнить ст. 82 УПК РФ положением о порядке хранения цифровых доказательств, включая результаты нейросетевой генерации и использованную для их создания биометрическую информацию:

«Вещественные доказательства, представленные в цифровой форме, включая контент, сгенерированный с использованием технологий искусственного интеллекта, должны храниться с применением мер защиты, исключающих их утрату, модификацию или несанкционированный доступ».

Подобные правила в ст. 82 УПК РФ на современном этапе будут намного более востребованными, чем правила обращения с такими архаизмами как видео- или киноплёнка.

Большой объем доказательственной информации в цифровой форме поступает в уголовное судопроизводство достаточно давно. Но технологии нейросетевой генерации контента своим существованием, распространением и уровнем развития ставят под сомнение эффективность существующей уголовно-процессуальной регламентации для цели объективного установления обстоятельств по делам, где процесс доказывания связан с изучением информации в цифровой форме.

Иными словами, сомнения относительно природы происхождения того или иного объекта в цифровой форме возникают в судопроизводстве обоснованно и неизбежно. А правовые механизмы оценки или методические аспекты проверки доказательств на предмет их возможной нейросетевой генерации в настоящее время не разработаны или неэффективны (в силу недостаточного уровня их развития или высоких темпов устаревания на фоне бума развития новых нейросетевых технологий).

## Библиография

1. Бодров Н.Ф., Лебедева А.К. Понятие дипфейка в российском праве, классификация дипфейков и вопросы их правового регулирования // Юридические исследования. 2023. № 11. С.26-41. DOI: 10.25136/2409-7136.2023.11.69014 EDN: DYIHIR URL: [https://e-notabene.ru/lr/article\\_69014.html](https://e-notabene.ru/lr/article_69014.html)
2. Pfefferkorn R. «Deepfakes» in the Courtroom. Rochester // Boston University Public

- Interest Law Journal.Vol. 29. 2020. No. 2, Pp. 244-276. Retrieved from [https://papers.ssrn.com/sol3/papers.cfm?abstract\\_id=4321140](https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4321140).
3. Парыгина Н. Н. Латентная диффамация: правонарушение с «творческим» подходом // Закон. 2024. № 7. С. 171-180. DOI 10.37239/0869-4400-2024-21-7-171-180.
4. Мыскина К. М. О некоторых проблемах оплаты судебных экспертиз в гражданском процессе // Вестник Университета имени О.Е. Кутафина (МГЮА). 2024. № 3(115). С. 137-143.
5. Chingovska I. et al. The 2nd competition on counter measures to 2D face spoofing attacks // Proc. IAPR Int. Conf. Biometrics (ICB), Jun. 2013. Pp. 1-6. Retrieved from [https://www.researchgate.net/publication/303156337\\_The\\_2nd\\_competition\\_on\\_counter\\_measures\\_to\\_2D\\_face\\_spoofing\\_attacks](https://www.researchgate.net/publication/303156337_The_2nd_competition_on_counter_measures_to_2D_face_spoofing_attacks).
6. Frisella M. et al. Quantifying Bias in a Face Verification System // Computer Sciences & Mathematics Forum. 2022 volume 3. Retrieved from [https://www.researchgate.net/publication/360168993\\_Quantifying\\_Bias\\_in\\_a\\_Face\\_Verification\\_System](https://www.researchgate.net/publication/360168993_Quantifying_Bias_in_a_Face_Verification_System)
7. Yucer S. et al. Exploring Racial Bias within Face Recognition via per-subject Adversarially-Enabled Data Augmentation // IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW). 2020. Pp. 83-92. Retrieved from [https://www.researchgate.net/publication/340805716\\_Exploring\\_Racial\\_Bias\\_within\\_Face\\_Recognition\\_via\\_per-subject\\_Adversarially-Enabled\\_Data\\_Augmentation](https://www.researchgate.net/publication/340805716_Exploring_Racial_Bias_within_Face_Recognition_via_per-subject_Adversarially-Enabled_Data_Augmentation)
8. Алихаджиева И. С. О новых способах совершения преступлений с использованием персональных данных // Вестник Прикамского социального института. 2024. № 1 (97). С. 22-30.
9. Бодров Н.Ф., Лебедева А.К. Понятие дипфейка (deepfake) в российском праве, его классификация и проблемы правового регулирования // Юридический вестник ДГУ. 2023. Т 48, № 4(68). С. 173-181.
10. Поздняк И. Н. Цифровые угрозы в современном мире: технология deepfake // Судебная экспертиза Беларуси. 2024. № 2(19). С. 72-77.
11. Paulo Max Gil Innocencio Reis, Rafael Oliveira Ribeiro. A forensic evaluation method for DeepFake detection using DCNN-based facial similarity scores // Forensic Science International, 2024. Volume 358. <https://doi.org/10.1016/j.forsciint.2023.111747>. Retrieved from <https://www.sciencedirect.com/science/article/pii/S0379073823001974>.
12. Öztürk S. B., Özyer B. and Temiz Ö. Detection of Voices Generated by Artificial Intelligence with Deep Learning Methods // 32nd Signal Processing and Communications Applications Conference (SIU), Mersin, Turkiye, 2024. Pp. 1-4, doi: 10.1109/SIU61531.2024.10601078. Retrieved from <https://ieeexplore.ieee.org/abstract/document/10601078>.
13. Бодров Н. Ф., Лебедева А. К. Дипфейк как объект судебной экспертизы // Национальные и международные тенденции и перспективы развития судебной экспертизы: сборник докладов Научно-практической конференции с международным участием. Нижний Новгород: ННГУ, 2024. С. 42-50.
14. Зубов Г. Н., Зубова П. И. Фальсификация звуковой информации с использованием технологий искусственного интеллекта. Особенности технического исследования // Вестник криминалистики. 2023. № 3. С. 5-26.
15. Спиридонов М. С. Технологии искусственного интеллекта в уголовно-процессуальном доказывании // Journal of Digital Technologies and Law. 2023. Т. 1, № 2. С. 481-497. doi 10.21202/jdtl.2023.20

## Результаты процедуры рецензирования статьи

*В связи с политикой двойного слепого рецензирования личность рецензента не*

раскрывается.

Со списком рецензентов издательства можно ознакомиться [здесь](#).

Предметом исследования в представленной на рецензирование статье является, как это следует из ее наименования, судебная практика установления обстоятельств в случаях противоправного распространения генеративного контента, созданного с помощью технологий искусственного интеллекта. Заявленные границы исследования соблюдены ученым.

Методология исследования в тексте статьи не раскрывается.

Актуальность избранной автором темы исследования несомненна и обосновывается им следующим образом: "В настоящее время технологии искусственного интеллекта (далее – ИИ) для создания генеративного контента развиваются крайне стремительно и задействуются в самых различных областях общественной жизни. Одновременно формируется и практика рассмотрения судебных дел, в которых в качестве доказательств используются цифровые следы, представляющие собой сгенерированный при помощи ИИ контент. Данный контент – это цифровые продукты в различном виде: в виде текста, графики, звука или их сочетания. Безусловно, активное развитие нейросетевых технологий в сфере создания генеративного контента предоставляет огромнейшие возможности для его противоправного использования и распространения. Актуальность выбранной темы обусловлена тем, что всё больше случаев, когда синтезированный контент используется в целях введения в заблуждение или преодоления пользователем систем контроля и управления доступом [1, С. 38], в рамках дел о клевете, различных мошенничествах, при рассмотрении граждански дел (например, бракоразводный процесс). В подобных ситуациях перед правоприменителями встаёт задача установления обстоятельств, связанных с созданием, использованием и распространением сгенерированного при помощи нейросетевых технологий контента: – технологические сложности выявления подобного генеративного контента, доказательства факта его создания как в целом при помощи нейросетевых технологий, так и конкретных нейросетей, использованных для его создания; – целесообразность назначения судебной экспертизы" и др. Дополнительно ученому необходимо перечислить фамилии ведущих специалистов, занимавшихся исследованием поднимаемых в статье проблем, а также раскрыть степень их изученности.

Научная новизна работы проявляется в ряде заключений автора: "Вопрос оплаты за производство судебной экспертизы может в ряде случаев быть существенной преградой для реализации права на защиту от обвинения или возражения на иск. Так, например, в отечественной гражданско-процессуальной практике расходы на оплату судебной экспертизы возлагаются судом, как правило, на сторону, ходатайствовавшую о её производстве [4]. Согласно положениям ст. 96 ГПК РФ, денежные суммы, подлежащие выплате экспертам, предварительно вносятся на счёт, открытый в порядке, установленном бюджетным законодательством РФ, стороной, заявившей соответствующую просьбу. В уголовном судопроизводстве финансовые расходы часто возникают ещё на стадии внепроцессуальных исследований, когда для обоснования заявления требуется заключение специалиста, оплата подготовки которого происходит за счёт собственных средств заявителя. Производство судебной фоноскопической экспертизы также связано с некоторыми рисками. Во-первых, в ситуации, когда результатом экспертного исследования станет вывод в форме НПВ (не представляется возможным ответить на поставленный судом или следователем вопрос), заявитель может быть фактически лишён возможностей последующего доказывания. Во-вторых, в ситуации, когда судебно-экспертные методики не содержат соответствующих разделов, посвящённых выявлению признаков синтеза, шанс выявить сам факт синтеза весьма

мал, так как у эксперта при производстве исследования не будет методического подспорья для выявления и описания каких-либо типовых признаков генерации или обстоятельств, подлежащих установлению при проверке версии о синтезе. В-третьих, в ситуации, когда экспертная практика по такого рода делам ещё не сформировалась, логично предположить, что выявление факта генерации окажется под силу только наиболее высококвалифицированным экспертам, которых очевидно не так много среди всех практикующих работников"; "Однако отсутствие определения «дипфейка» в российском законодательстве может крайне негативно влиять на правоприменительную практику. Предложенная дефиниция дипфейка не позволяет однозначно квалифицировать действия, связанные с его использованием и распространением. Результаты обобщения информации о рассмотренных в настоящей статье делах подтверждают не только необходимость разработки правовых механизмов регулирования генеративного контента, но и совершенствования судебно-экспертного противодействия для борьбы с дипфейками. Но на современном этапе судебно-экспертное обеспечение как в нашей стране, так и за рубежом значительно отстаёт от темпов развития генеративных нейросетей"; "В первую очередь правоприменитель должен обладать навыками формулирования экспертной задачи, особенно на современном этапе разработанности методического обеспечения, когда консультационная помощь сведущих лиц может быть получена только от наиболее квалифицированных специалистов, которые, как мы отмечали ранее, ожидаемо составляют меньшинство от общего числа профессионалов в области судебной фоноскопии. Классической и знакомой большинству правоприменителей формулировкой для технического исследования фонограмм является вопрос: «Имеются ли в представленной фонограмме «Название фонограммы.mp3» признаки неситуационных изменений, в том числе монтажа и др.?». Однако в случае со сгенерированным контентом изменения не вносятся в какую-либо фонограмму (хотя, как мы описывали ранее, и такие случаи возможны). Важно понимать, что в ситуациях с предоставлением в качестве доказательства дипфейка речь чаще всего идёт о создании новой фонограммы, отражающей речевое событие, которого никогда не существовало" и др. Таким образом, статья вносит определенный вклад в развитие отечественной правовой науки и, безусловно, заслуживает внимания потенциальных читателей.

Научный стиль исследования выдержан автором в полной мере.

Структура работы логична. Во вводной части статьи ученый обосновывает актуальность избранной им темы исследования. Основная часть работы состоит из следующих разделов: "1. Уголовное дело по обвинению директора школы"; "2. Уголовное дело по обвинению учёного-гидролога"; "3. Обвинения в преступлениях и административных правонарушениях по результатам вмешательства в выборы"; "4. Дипфейк-фонограмма в качестве доказательства в бракоразводном процессе"; "5. Судебная практика защиты исключительных прав на цифровые продукты"; "6. Категория дел о телефонных мошенничествах с использованием нейросетевых технологий"; "7. Практика Верховного суда РФ"; "8. Судебно-экспертное обеспечение по делам о противоправном распространении генеративного контента, созданного с помощью технологий искусственного интеллекта". В заключительной части работы содержатся выводы по результатам проведенного исследования.

Содержание статьи соответствует ее наименованию, но не лишено недостатков формального характера.

Так, автор пишет: "Актуальность выбранной темы обусловлена тем, что всё больше случаев, когда синтезированный контент используется в целях введения в заблуждение или преодоления пользователем систем контроля и управления доступом [1, С. 38], в рамках дел о клевете, различных мошенничествах, при рассмотрении граждански дел

(например, бракоразводный процесс)" - "гражданских дел" (опечатка).

Ученый отмечает: "Однако, как нам представляется, центральное место в доказывании в подобных делах, будет занимать заключение эксперта. Так как первоочередной проблемой является сложность в доказывании того, является ли представленный цифровой продукт, сгенерированным при помощи нейросетевых технологий или нет" - "Однако, как нам представляется, центральное место в доказывании по подобным делам будет занимать заключение эксперта, поскольку первоочередной проблемой становится сложность в доказывании того, является ли представленный цифровой продукт сгенерированным при помощи нейросетевых технологий или нет" (см. на пунктуацию).

Автор пишет: "Дипфейки все чаще используется при осуществлении диффамации, как указывает, например, Парыгина Н. Н. «Создать ложное представление о том, что тот или иной человек что-либо сказал или сделал, теперь беспрецедентно легко» [3, С. 178]" - "Дипфейки все чаще используются при осуществлении диффамации, как указывает, например, Н. Н. Парыгина: «Создать ложное представление о том, что тот или иной человек что-либо сказал или сделал, теперь беспрецедентно легко» [3, С. 178]".

Таким образом, статья нуждается в дополнительном вычитывании - в ней встречаются опечатки, пунктуационные и стилистические ошибки (приведенный в рецензии перечень опечаток и ошибок не является исчерпывающим!).

Библиография исследования представлена 15 источниками (научными статьями), в том числе на английском языке. С формальной и фактической точек зрения этого достаточно. Автору удалось раскрыть тему исследования с необходимой полнотой и глубиной. Работа выполнена на достаточно высоком академическом уровне.

Апелляция к оппонентам имеется, как общая, так и частная (Я. Е. Нилов, А. К. Пушков, М. И. Шадаев и др.), и вполне достаточна. Научная дискуссия ведется автором корректно. Положения работы обоснованы в должной степени и проиллюстрированы примерами.

Выводы по результатам проведенного исследования имеются ("По результатам проведенного анализа установлено, что стремительное развитие технологий искусственного интеллекта, в частности генеративных нейросетей, приводит к новым вызовам для правоприменения и судебно-экспертной деятельности. Обобщение отечественной и зарубежной судебной практики, связанной с фактами предоставления дипфейков и других результатов нейросетевой генерации в качестве доказательств, продемонстрировало сложность и многоаспектность задач, стоящих перед экспертами и в последующем – перед правоприменителями. В качестве ключевых проблем уже сейчас кристаллизовались в правоприменительной практике: 1. Неопределённость правового регулирования. Отсутствие унифицированного определения дипфейка и регламентации генеративного контента создаёт правовую неопределённость. Это осложняет квалификацию противоправных действий, связанных с использованием дипфейков, что в последующем будет неизбежно затруднять доказывание обстоятельств в суде. 2. Технологическая сложность исследования. Существующие экспертные методики не всегда позволяют достоверно определить факт генерации контента нейросетевыми алгоритмами. Это особенно актуально для фоноскопической экспертизы и анализа синтезированных/клонированных голосов.

3. Ограниченный доступ к технологическим возможностям для исследования.

Технологическая, финансовая и организационная доступность экспертиз и исследований остаётся ограниченной, что будет препятствовать реализации права на защиту и справедливое судебное разбирательство. Несмотря на очевидно малый на данном этапе объем подобных дел, прогнозируя меры, необходимые для более эффективной правоприменительной и экспертной деятельности, можно сформулировать некоторые рекомендации, актуальные уже сейчас: - Разработка научно-методической базы.

Необходимо создать специализированные фонотеки и датасеты с примерами синтезированного контента и доступом к образцам, использованным для генерации. Это позволит обеспечить подготовку и повышение квалификации экспертов, а также выработать новые методические подходы к выявлению дипфейков. Организационно это приведет и к упрощению таких исследований, и к снижению затрат на их производство. - Междисциплинарное взаимодействие. Важной задачей на первом этапе становится координация работы правоприменителей, экспертов и технических специалистов в области технологий искусственного интеллекта для создания универсальных алгоритмов детекции дипфейков, не зависящих от конкретных технологий генерации. - Совершенствование законодательства.

Требуется внедрение чётких правовых дефиниций и механизмов регулирования использования генеративных технологий. Особое внимание следует уделить квалификации противоправных действий, совершённых с использованием дипфейков. В качестве конкретных изменений в уголовно-процессуальное законодательство наиболее своевременными нам представляются следующие: 1. Вектор цифровизации уголовного процесса должен учитывать сам факт того, то в качестве доказательств в цифровой форме будут предоставляться подложные сведения в виде конкретных объектов – дипфейков. Так, на определенном этапе технологического развития и фактической доступности высокореалистичных технологий нейросетевой генерации вполне целесообразным представляется дополнение ст. 196 УПК РФ пунктом 6 следующего содержания: «6) факт генерации цифровых данных, представленных в качестве доказательства по уголовному делу, когда возникает сомнение в их достоверности или необходимо установление факта использования биометрических персональных данных конкретного лица в процессе такой генерации» и др.), они четкие, конкретные, обладают свойствами достоверности, обоснованности и, несомненно, заслуживают внимания научного сообщества.

Интерес читательской аудитории к представленной на рецензирование статье может быть проявлен прежде всего со стороны специалистов в сфере административного права, информационного права, уголовного процесса при условии ее небольшой доработки: раскрытии методологии исследования, дополнительном обосновании актуальности его темы (в рамках сделанного замечания), устранении нарушений в оформлении статьи.

## **Результаты процедуры повторного рецензирования статьи**

*В связи с политикой двойного слепого рецензирования личность рецензента не раскрывается.*

*Со списком рецензентов издательства можно ознакомиться [здесь](#).*

### **РЕЦЕНЗИЯ**

на статью на тему «Анализ судебной практики установления обстоятельств в случаях противоправного распространения генеративного контента, созданного с помощью технологий искусственного интеллекта».

Предмет исследования.

Предложенная на рецензирование статья посвящена актуальным вопросам установления обстоятельств в случаях противоправного распространения генеративного контента, созданного с помощью технологий искусственного интеллекта. Как отмечается в самой статье, «Судебная практика в рассматриваемой сфере ещё только формируется, сталкиваясь безусловно с рядом сложностей, как правового, так и технологического характера. В данной статье рассмотрена отечественная и зарубежная практика

судебного и следственного установления обстоятельств по делам о противоправном распространении синтезированного контента (дипфейков)». В качестве конкретного предмета исследования выступили, прежде всего, положения российского законодательства, материалы судебной практики, мнения ученых.

Методология исследования.

Цель исследования прямо в статье не заявлена. При этом она может быть ясно понята из названия и содержания работы. Цель может быть обозначена в качестве рассмотрения и разрешения отдельных проблемных аспектов вопроса о судебной практике установления обстоятельств в случаях противоправного распространения генеративного контента, созданного с помощью технологий искусственного интеллекта. Исходя из поставленных цели и задач, автором выбрана методологическая основа исследования. Как отмечено в самой статье, «Методологическую основу исследования составляют всеобщий диалектический метод, общенаучные (описание, сравнение, обобщение, моделирование и др.) и частнонаучные методы. Мы проанализируем те, пока немногочисленные, кейсы, когда синтезированный контент использовался в качестве доказательства в гражданских, арбитражных и уголовных делах. По нашему мнению, в ближайшей перспективе объём подобных дел в судопроизводстве будет неуклонно возрастать, а первичный анализ и обобщение специфики подобных дел смогут быть использованы для предотвращения судебных ошибок».

В частности, автором используется совокупность общенаучных методов познания: анализ, синтез, аналогия, дедукция, индукция, другие. В частности, методы анализа и синтеза позволили обобщить и разделить выводы различных научных подходов к предложенной тематике, а также сделать конкретные выводы из материалов судебной практики.

Наибольшую роль сыграли специально-юридические методы. В частности, автором активно применялся формально-юридический метод, который позволил провести анализ и осуществить толкование норм действующего законодательства. Например, следующий вывод автора: «создание подобной базы ставит под угрозу и конституционные права граждан: статья 23 Конституции РФ [Конституция Российской Федерации (принята всенародным голосованием 12.12.1993 г. с изменениями, одобренными в ходе общероссийского голосования 01.07.2020 г.)] «2. Каждый имеет право на тайну переписки, телефонных переговоров, почтовых, телеграфных и иных сообщений. Ограничение этого права допускается только на основании судебного решения». Для записи телефонных переговоров, даже с предполагаемыми телефонными мошенниками оператору сотовой связи необходимо будет получить согласие абонента, что с практической точки зрения создаст препятствие для оперативной и эффективной работы проектируемой системы».

Следует положительно оценить возможности эмпирического метода исследования, связанного с изучением материалов судебной практики. В частности, указано, что «Пленум Верховного суда РФ также обратил своё внимание на проблему дипфейков, упомянув в своих Постановлениях: «вводящие в заблуждение поддельных изображений, аудио- и аудиовизуальных материалов, в том числе созданных с помощью компьютерных технологий» (п.111), «распространение вводящих в заблуждение избирателей, участников референдума поддельных изображений, аудио- и аудиовизуальных материалов, в том числе созданных с помощью компьютерных технологий» (п.129) [Постановление Пленума Верховного Суда Российской Федерации № 24 от 27 июня 2024 г. «О некоторых вопросах, возникающих при рассмотрении судами административных дел о защите избирательных прав и права на участие в референдуме граждан Российской Федерации»]; «с использованием вводящих в заблуждение и выдаваемых за достоверные недостоверных изображений, аудио- и аудиовизуальной информации, в том

числе созданных с помощью компьютерных технологий (п.20)» [Постановление Пленума Верховного Суда Российской Федерации № 17 от 25 июня 2024 г. «Об отдельных вопросах, возникающих у судов при рассмотрении дел об административных правонарушениях, посягающих на установленный порядок информационного обеспечения выборов и референдумов»]».

Таким образом, выбранная автором методология в полной мере адекватна цели исследования, позволяет изучить все аспекты темы в ее совокупности.

Актуальность.

Актуальность заявленной проблематики не вызывает сомнений. Имеется как теоретический, так и практический аспекты значимости предложенной темы. С точки зрения теории тема установления обстоятельств в случаях противоправного распространения генеративного контента, созданного с помощью технологий искусственного интеллекта, сложна и неоднозначна. Сложно спорить с автором статьи в том, что «В настоящее время технологии искусственного интеллекта (далее – ИИ) для создания генеративного контента развиваются крайне стремительно и задействуются в самых различных областях общественной жизни. Одновременно формируется и практика рассмотрения судебных дел, в которых в качестве доказательств используются цифровые следы, представляющие собой сгенерированный при помощи ИИ контент. Данный контент – это цифровые продукты в различном виде: в виде текста, графики, звука или их сочетания». Приводимые автором в статье примеры из судебной практики наглядно демонстрирует этот вопрос.

Тем самым, научные изыскания в предложенной области стоит только поприветствовать. Научная новизна.

Научная новизна предложенной статьи не вызывает сомнений. Во-первых, она выражается в конкретных выводах автора. Среди них, например, такой вывод:

«В качестве ключевых проблем уже сейчас кристаллизировались в правоприменительной практике: 1. Неопределённость правового регулирования. Отсутствие унифицированного определения дипфейка и регламентации генеративного контента создаёт правовую неопределённость. Это осложняет квалификацию противоправных действий, связанных с использованием дипфейков, что в последующем будет неизбежно затруднять доказывание обстоятельств в суде. 2. Технологическая сложность исследования. Существующие экспертные методики не всегда позволяют достоверно определить факт генерации контента нейросетевыми алгоритмами. Это особенно актуально для фоноскопической экспертизы и анализа синтезированных/клонированных голосов. 3. Ограниченный доступ к технологическим возможностям для исследования. Технологическая, финансовая и организационная доступность экспертиз и исследований остаётся ограниченной, что будет препятствовать реализации права на защиту и справедливое судебное разбирательство».

Указанный и иные теоретические выводы могут быть использованы в дальнейших научных исследованиях.

Во-вторых, автором предложены идеи по совершенствованию действующего законодательства. В частности,

«В связи с этим логично дополнить ст. 82 УПК РФ положением о порядке хранения цифровых доказательств, включая результаты нейросетевой генерации и использованную для их создания биометрическую информацию: «Вещественные доказательства, представленные в цифровой форме, включая контент, сгенерированный с использованием технологий искусственного интеллекта, должны храниться с применением мер защиты, исключающих их утрату, модификацию или несанкционированный доступ». Подобные правила в ст. 82 УПК РФ на современном этапе будут намного более востребованными, чем правила обращения с такими



архаизмами как видео- или киноплёнка».

Приведенный вывод может быть актуален и полезен для правотворческой деятельности. Таким образом, материалы статьи могут иметь определенных интерес для научного сообщества с точки зрения развития вклада в развитие науки.

Стиль, структура, содержание.

Тематика статьи соответствует специализации журнала «Юридические исследования», так как она посвящена правовым проблемам, связанным с формированием судебной практики по вопросам, связанным с дипфейками и искусственным интеллектом.

Содержание статьи в полной мере соответствует названию, так как автор рассмотрел заявленные проблемы, в целом достиг цели исследования.

Качество представления исследования и его результатов следует признать в полной мере положительным. Из текста статьи прямо следуют предмет, задачи, методология и основные результаты исследования.

Оформление работы в целом соответствует требованиям, предъявляемым к подобного рода работам. Существенных нарушений данных требований не обнаружено.

Библиография.

Следует высоко оценить качество использованной литературы. Автором активно использована литература, представленная авторами из России и из-за рубежа (Бодров Н.Ф., Лебедева А.К., Зубов Г.Н., Зубова П.И., Öztürk S.B., Özyer B., Temiz Ö. и другие). Хотелось бы отметить использование автором большого количества материалов судебной практики, что позволило придать исследованию правоприменительную направленность. Таким образом, труды приведенных авторов соответствуют теме исследования, обладают признаком достаточности, способствуют раскрытию различных аспектов темы.

Апелляция к оппонентам.

Автор провел серьезный анализ текущего состояния исследуемой проблемы. Все цитаты ученых сопровождаются авторскими комментариями. То есть автор показывает разные точки зрения на проблему и пытается аргументировать более правильную по его мнению.

Выводы, интерес читательской аудитории.

Выводы в полной мере являются логичными, так как они получены с использованием общепризнанной методологии. Статья может быть интересна читательской аудитории в плане наличия в ней систематизированных позиций автора применительно к вопросам совершенствования законодательства по отношению к противоправному распространению генеративного контента, созданного с помощью технологий искусственного интеллекта.

На основании изложенного, суммируя все положительные и отрицательные стороны статьи

«Рекомендую опубликовать»