Влияние технологий «deepfake» на оценку доказательств: кризис доверия

Хубиева Милана Муратовна
студентка, МГЮА им. О.Е. Кутафина,
milanakhubieva777@gmail.com

Технологии синтетического аудиовизуального контента (deepfake, далее – дипфейк) создают беспрецедентный вызов уголовному судопроизводству, ставя под сомнение фундаментальные принципы оценки доказательств. В статье исследуется феномен алгоритмически генерируемых синтетических медиа как качественно новый деструктивный фактор, способный системно искажать доказательственную информацию. На основе анализа зарубежной и российской практики выявляются системные риски использования дипфейков, особое внимание уделено техническим, процессуальным и психологическим барьерам их выявления. Методологическую основу исследования составляют диалектический метод, общенаучные методы анализа, синтеза, индукции, дедукции, системно-структурный подход, а также частнонаучные методы: формально-юридический, сравнительно-правовой и правового моделирования. Эмпирическая база включает материалы зарубежной и отечественной судебной практики, доктринальные источники на стыке уголовно-процессуального права, криминалистики и компьютерных наук. Автор обосновывает необходимость замены презумпции доверия к аудиовизуальным доказательствам на повышенный стандарт проверки их достоверности. Предлагается комплекс мер: законодательное закрепление обязательного хеширования цифровых доказательств; введение обязательной комплексной судебной экспертизы при мотивированных сомнениях в аутентичности; дополнение статьи 204 УПК РФ требованиями к раскрытию экспертных методик; институциональное укрепление экспертных и следственных подразделений; системное повышение цифровой компетентности судей, следователей и адвокатов. Делается вывод, что своевременное и комплексное противодействие угрозам, порождаемым технологиями синтетического аудиовизуального контента, является ключевым условием сохранения доверия общества к правосудию в цифровую эпоху.

Ключевые слова: уголовный процесс, доказательства, оценка доказательств, относимость доказательств, достоверность доказательств, допустимость доказательств, фальсификация доказательств, deepfake, цифровые технологии, судебная экспертиза, цифровая криминалистика.

Современное уголовное судопроизводство стоит на пороге технологической революции, бросающей вызов его классическим, многовековым основам. Стремительное развитие технологий генеративного искусственного интеллекта, и в частности методов создания синтетических аудиовизуальных материалов, известных также под наименованием «дипфейк», порождает качественно новый деструктивный фактор уголовного судопроизводства – алгоритмически генерируемые аудио- и видеоподлоги целенаправленно искажающие устанавливаемые по делу обстоятельства и потенциально способные дестабилизировать всю систему доказывания. Указанное явление требует незамедлительного научного осмысления и адекватной процессуальной реакции.

Актуальность темы исследования обусловлена тремя ключевыми факторами. Во-первых, отмечается экспоненциальное развитие технологий синтеза медиаданных по созданию фотореалистичного контента, идентифицировать который без применения специальных средств становится невозможным даже для подготовленного эксперта. В 2026 году исследователи Бристольского университета Саймон Кларк и Стефан Левандовски провели три пререгистрированных эксперимента общей выборкой в 673 человека. Участники выносили суждение по шкале от –3 («определённо не виновен») до +3 («определённо виновен»), где ноль соответствовал невозможности дать ответ. Несмотря на прямое предупреждение о том, что демонстрируемое видео является дипфейком, средняя оценка виновности в экспериментальной группе составила +0,87 (склонность к ответу «скорее виновен»), тогда как в контрольной группе, наблюдавшей нейтральный ролик, балл составляет –0,40. Более того, 53,8 % участников, заявивших о доверии к предупреждению, всё равно основывали своё мнение на содержании видео.

Во-вторых, уголовно-процессуальное законодательство Российской Федерации, основанное на традиционных подходах к оценке доказательств (ст. 17, 88 Уголовно-процессуального кодекса Российской Федерации (далее – УПК РФ)), оказалось не готово к вызовам, порожденным цифровой трансформации. В условиях стремительного развития технологий уголовное судопроизводство оказывается в позиции, где нормы права отстают от технологического развития. Уже сегодня известны случаи использования дипфейков для шантажа, распространения порнографических материалов и дискредитации публичных лиц.

В-третьих, основная опасность заключается не только в риске фальсификации доказательств, но и в системных трудностях их своевременного выявления и оспаривания в рамках судебного разбирательства, что ставит под удар такие фундаментальные принципы уголовного судопроизводства, как независимость судей, состязательность и свобода оценки доказательств.

Цель настоящей статьи – выявить системные барьеры выявления дипфейков и предложить процессуальный механизм их преодоления.

В настоящее время вопросы, связанные с дипфейками, находятся преимущественно в фокусе внимания специалистов в области компьютерных наук и цифровой безопасности. В отечественной юридической науке данная проблематика только начинает изучаться, что проявляется в единичных публикациях, посвященных, главным образом, общему описанию явления и его потенциальных угроз. При этом комплексный процессуальный анализ

проблем допустимости, проверки и оценки подобных материалов, а также системных барьеров на пути их разоблачения в суде остается недостаточно исследованным. Существует явный разрыв между технологическим вызовом и его уголовно-процессуальным осмыслением.

Технологии дипфейк представляют собой продукт развития нейросетевых алгоритмов, способных с высокой степенью правдоподобия подменять изображение или голос человека на фото- и видеоматериалах. В результате создаётся контент, практически неотличимый от подлинного. Если на ранних этапах подобные «подделки» можно было выявить благодаря примитивным артефактам, таким как искажённые тени или статичность мимики, то современные генеративные модели обеспечивают уровень реалистичности, который затрудняет распознавание даже для экспертов.

В уголовном процессе это явление имеет двойственную природу: с одной стороны, дипфейк может быть представлен как доказательство, отражающее «события преступления», с другой – он становится инструментом преступления, когда используется для шантажа, мошенничества, распространения порнографии или клеветы, выступая одновременно и предметом преступного деяния, и потенциальным источником ложной доказательственной информации.

Наиболее серьёзные риски проявляются в следующих аспектах:

1. Подмена личности – создание с помощью дипфейк-видео или аудиозаписей с «участием» лица, фактически не принимавшего участия в событии. Например, искусственное «признание» в совершении преступления.

2. Формирование ложного алиби – моделирование ситуации, в которой подозреваемый якобы находился в другом месте.

3. Дискредитация свидетеля или потерпевшего – создание компрометирующих материалов, ставящих под сомнение достоверность их показаний.

4. Манипулирование общественным мнением – особенно в делах с участием присяжных заседателей, для которых визуальная убедительность материала может перевесить рациональную оценку доказательств.

Реальная практика уже подтверждает опасность использования подобных технологий. Так, в 2019 году в Великобритании был зафиксирован случай, когда злоумышленники с помощью синтезированной голосовой имитации убедили руководителя одной из энергетических компаний перевести средства на подставной счёт. Согласно опубликованным данным, потерпевший был уверен, что разговаривает по телефону с вышестоящим руководителем материнской компании, тогда как в действительности голос принадлежал мошеннику, использовавшему технологию генеративного искусственного интеллекта для клонирования речи. Аналогичный инцидент произошёл в Объединённых Арабских Эмиратах в начале 2020 года: с использованием синтезированного голоса директора банка было похищено 35 млн долларов США.

Особую проблему составляет восприятие материалов, созданных с использованием дипфейков, судами. Исследования показывают, что даже специалисты, не говоря уже о судьях или присяжных, не способны с уверенностью отличить подделку от оригинала. Как показал описанный ранее эксперимент Кларка и Левандовски, даже прямое предупреждение о подделке не устраняет влияние синтетического контента на формирование убеждения. В условиях, когда визуальная информация традиционно считается наиболее надёжной, это создаёт риск серьёзных судебных ошибок.

В российском праве отсутствует специальное регулирование, посвящённое материалам, созданным с использованием дипфейков. Оценка доказательств в уголовном процессе традиционно строится на основе внутреннего убеждения суда, сформированного по результатам всестороннего и объективного исследования всех материалов дела. УПК РФ ограничивается общими положениями: любое доказательство должно быть получено законным путём и подлежит проверке с точки зрения относимости, допустимости и достоверности (ст. 74–88 УПК РФ), при этом закон не устанавливает иерархии доказательств. В свою очередь, на практике именно аудио- и видеозаписи нередко воспринимаются как обладающие повышенной убедительностью. Суды проверяют главным образом процессуальную форму, а вопрос о технической подлинности содержания почти не затрагивается. Например, при рассмотрении дел, где видеозапись выступает ключевым доказательством, суды обычно опираются на заключения экспертов о целостности носителя и отсутствии монтажа. Современные технологии дипфейк позволяют создавать материалы без явных следов редактирования, что делает стандартные методы проверки малоэффективными, а это, в свою очередь, создаёт риск судебных ошибок при сохранении прежнего уровня доверия к таким записям.

В контексте анализа новых форм фальсификации показательно рассмотренное в 2023 году в Республике Татарстан уголовное дело, где работник коллекторского агентства, вместо стандартных угроз, создал с помощью графического редактора серию порнографических коллажей, в которые интегрировал фотографии супруги должника и её знакомых. Затем он рассылал эти изображения не только жертве, но и её детям, друзьям, разрушая социальные связи. Данный эпизод, несмотря на применение относительно простых технологий монтажа, представляет собой важный переходный этап к более сложным синтетическим медиа и раскрывает системные проблемы правовой оценки подобных действий. Как отмечают К.В. Старостенко и А.С. Коновалов, дипфейк-технологии эксплуатируют «эмоциональную суггестию» – шокирующий контент вызывает сильные эмоции, снижая критическое восприятие. В данном случае правоприменитель квалифицировал действия по трём статьям: клевета (ст. 128.1 УК РФ), незаконное изготовление порнографических материалов (ст. 242 УК РФ) и самоуправство (ст. 330 УК РФ). Примечательно, что суд признал цифровой коллаж не просто способом клеветы, а самостоятельным преступным продуктом по ст. 242 УК РФ.

Данный эпизод с использованием статичного фотомонтажа, пусть и технологически примитивного по меркам современных дипфейков, тем не менее отчётливо демонстрирует уязвимость правовой системы: даже относительно грубые цифровые подделки ставят правоприменителя в тупик. Можно предположить, что при переходе злоумышленников к полноценному синтетическому аудиовизуальному контенту риски кратно возрастут.

Подтверждение этому можно найти и в судебной практике последних лет. Так, приговором Самарского районного суда от 6 июня 2025 года по делу № 1-48/2025 пятеро сотрудников транспортной полиции были осуждены за превышение должностных полномочий (п. «г, е» ч. 3 ст. 286 УК РФ). Ключевым доказательством обвинения выступала аудиозапись оперативно-розыскного мероприятия «опрос», на которой, по версии следствия, один из подсудимых признавался в хищении вещественных доказательств. Подсудимый последовательно заявлял, что запись является сфальсифицированной с использованием дипфейк-технологий синтеза речи, однако суд отклонил этот довод, сославшись на два обстоятельства. Во-первых, проведённая по делу фоноскопическая экспертиза не смогла идентифицировать голос подсудимого из-за низкого качества фонограммы, но и не подтвердила факт монтажа или синтеза. Во-вторых, что показательно, пока-

зания свидетелей – должностных лиц, знавших подсудимого по службе, – были признаны судом достаточными для отождествления голоса в обход экспертного заключения.

Данный пример наглядно иллюстрирует две системные проблемы, обозначенные в настоящем исследовании. Во-первых, суд не учёл принципиальные ограничения существующих методов экспертного исследования: фоноскопическая экспертиза не смогла ни подтвердить, ни исключить синтетическое происхождение записи, однако суд не применил повышенный стандарт проверки достоверности, фактически презюмируя подлинность доказательства, не получившего надлежащего научного подтверждения. Во-вторых, дело демонстрирует, что в отсутствие законодательно закреплённой процедуры верификации – включая обязательное хеширование и комплексную экспертизу – суды вынуждены возвращаться к «классическим» методам оценки, а именно к показаниям свидетелей. В эпоху высококачественного синтеза голоса подобный подход создаёт очевидный риск судебной ошибки.

На текущий момент специальные процессуальные нормы не установлены. Основным, а в ряде случаев единственным, инструментом проверки достоверности цифровых материалов остаётся судебная экспертиза. В российских экспертных учреждениях отсутствуют ведомственно утверждённые унифицированные методики выявления дипфейк-контента, что может приводить к формализации заключений. Систематическое применение глубокого технологического анализа в уголовном процессе не является общепринятой практикой.

Обращение к зарубежной практике тоже не даёт готовых решений. Проблема находится в фокусе внимания многих правопорядков, однако единообразное процессуальное решение пока не выработано: законодатели и суды действуют разрозненно, а универсальной модели проверки аутентичности аудиовизуальных доказательств в эпоху дипфейк до сих пор не предложено.

В Соединённых Штатах Америки действует стандарт Daubert, требующий от экспертных методик подтверждённой научной состоятельности, что потенциально позволяет сторонам оспаривать заключения, полученные с использованием непрозрачных алгоритмов детекции дипфейков. Вместе с тем на сегодняшний день отсутствуют прецеденты, в которых федеральные суды на основании стандарта Daubert исключали бы заключения экспертов по причине «непроверяемости» применённого программного обеспечения. Дискуссия остаётся в доктринальной плоскости. Одновременно на федеральном уровне обсуждаются законопроекты, направленные на криминализацию распространения синтетического контента, способного повлиять на правосудие, однако их принятие не снимает вопроса о процессуальных механизмах проверки таких материалов в суде.

Отдельного внимания заслуживает опыт Китайской Народной Республики, где с 2019 года действуют правила об обязательной маркировке синтетического аудио- и видеоконтента, а её нарушение влечёт ответственность вплоть до уголовной. Данный подход решает задачу общей превенции, но малоэффективен в ситуациях, когда дипфейк создаётся злоумышленником именно с целью введения в заблуждение правоприменителя – вероятно, что такой субъект не станет добровольно маркировать свой продукт.

Европейские государства, в частности, Германия, ориентируются на развитие возможностей судебной компьютерно-технической экспертизы и обсуждают повышение стандартов проверки материалов, полученных из непроверенных цифровых источников, однако соответствующие процессуальные нормы пока не кодифицированы.

Таким образом, мировой опыт подтверждает актуальность проблемы, но не предлагает готовых процессуальных моделей, применимых для механического заимствования в российском уголовном процессе.

Главная проблема заключается в том, что дипфейк-материалы разрушают устоявшуюся презумпцию достоверности аудиовизуальных доказательств. Процесс доказывания в уголовном судопроизводстве, регламентированный главой 11 УПК РФ, исторически формировался вокруг материальных и «осязаемых» доказательств. Появление высококачественных синтетических медиа, созданных с помощью технологий искусственного интеллекта, ставит перед правоприменителем ряд практически неразрешимых на текущий момент задач. Если раньше видеозапись считалась объективным отражением действительности, то теперь она может быть доказательством, достоверность которого не поддаётся верификации традиционными методами. Для суда это означает необходимость кардинально менять подход к оценке: вместо доверия к видео – применение повышенного стандарта проверки достоверности, требующего подтверждения аутентичности материала иными источниками. Трудности носят системный характер и начинаются с фундаментальных технических и экспертных барьеров.

Во-первых, ключевой проблемой является «эффект черного ящика», присущий как процессу создания, так и детекции дипфейков. Нейросетевые модели, генерирующие контент, работают по алгоритмам, не всегда понятным даже их разработчикам, что создает непреодолимый барьер при попытке разъяснить суду в рамках ст. 17 УПК РФ и ст. 204 УПК РФ природу выявленных признаков подлога. Заключение эксперта рискует превратиться в перечень технических параметров и вероятностей, лишенных наглядности и убедительности для лиц, не обладающих специализированными познаниями в области машинного обучения.

Например, эксперт в заключении указывает, что «метрика FID (Fréchet Inception Distance) для кадров видео аномально низка, а анализ спектрограммы аудиодорожки выявил артефакты кодирования, характерные для модели VQ-VAE-2». Для суда и сторон обвинения и защиты такая формулировка – юридически бессодержательный набор слов. В отличие от традиционной фальсификации, где эксперт может наглядно показать присяжным следы клея на склеенной аудиокассете или несовпадение теней на фотоколлаже, цифровой фальсификат таких осязаемых улик не оставляет. Всё, что от него остаётся, – статистические аномалии, различимые лишь на уровне математических моделей, но не воспринимаемые на глаз или на слух.

Второй фактор – это непрекращающееся соперничество генеративных и детектирующих алгоритмов. Как только исследователи находят способ выявлять артефакты одного поколения дипфейк, создатели подделок тут же выпускают следующее, лишённое этих признаков. Такого рода «состязание в скорости» делает любой статичный набор экспертных методик заведомо устаревшим к моменту его утверждения. Методы детекции, основанные на выявлении специфических артефактов ранних моделей дипфейк (например, отсутствие моргания, неестественная работа лицевой мускулатуры, артефакты в области зубов), стремительно устаревают. Как систематизирует итальянский специалист в области цифровой криминалистики Ирэн Америни в обзоре «Deepfake Media Forensics: Status and Future Challenges» (2025), если ранние детекторы опирались на явные артефакты (отсутствие моргания, неестественная мимика), то новейшие генеративные модели – диффузионные (DDPM) и StyleGAN3 – создают контент, где такие маркеры уже отсутствуют. Автор отмечает, что это приводит к феномену

«Impostor Bias» (предвзятости самозванца) – устойчивому скептицизму в отношении любых аудиовизуальных материалов, даже подлинных. С точки зрения уголовного процесса это означает, что эксперт, использующий методику, основанную на выявлении старых артефактов, с высокой вероятностью даст ложное заключение о подлинности видео, созданного по последнему слову техники. Динамика соревновательного развития между генерацией и детекцией дипфейков делает судебную экспертизу принципиально нестабильной: её выводы устаревают быстрее, чем может быть обновлена ведомственная методика.

В 2023 г. в Стэнфордском университете была продемонстрирована генеративная модель, синтезирующая видеопортреты с физиологически безупречным морганием – признаком, который ранее считался надёжным индикатором подделки. Экспериментально стало подтверждено: эксперт, пользующийся методиками, рассчитанными на предшествующие поколения дипфейк-алгоритмов, при исследовании современного синтетического контента с высокой вероятностью придёт к ошибочному выводу о его аутентичности. Сама объективность и достоверность экспертного заключения (ст. 88 УПК РФ) ставится в прямую зависимость от скорости обновления его методической базы, что противоречит презумпции стабильности судебной экспертизы.

В-третьих, отсутствие единых, законодательно или ведомственно утвержденных стандартов и методик проведения судебной экспертизы цифровых фальсификатов создает правовой вакуум и угрозу произвола. В отличие от традиционных видов экспертиз, например, дактилоскопической или почерковедческой, где существуют устоявшиеся протоколы и системы критериев, экспертиза дипфейков находится в зачаточном состоянии. Это приводит к ситуации, когда заключения разных экспертных учреждений по одному и тому же материалу могут быть прямо противоположными, так как основаны на разных, зачастую проприетарных алгоритмах и субъективной интерпретации их результатов.

В качестве примера можно привести ситуацию, когда сторона обвинения представляет видеозапись, полученную с камеры наблюдения, на которой запечатлен подсудимый. Защита ходатайствует о назначении экспертизы, утверждая, что запись – дипфейк. Одна экспертная организация, используя детектор «А», выдает заключение о подлинности видео, другая, используя более современный детектор «Б», – о его синтетической природе. У суда отсутствуют легальные критерии для предпочтения одной методики другой, поскольку ни одна из них не имеет официального статуса. Это парализует процесс оценки доказательств по внутреннему убеждению и превращает судебное разбирательство в «битву экспертов», исход которой определяется не истиной, а ресурсными возможностями сторон для привлечения более технологичных подрядчиков.

Технические и экспертные барьеры не являются лишь техническими неполадками. Они трансформируются в серьезные процессуальные проблемы, подрывающие основы доказательственного права: принципы относимости, допустимости, достоверности. Невозможность достоверно и наглядно продемонстрировать подлог в зале суда де-факто легитимизирует цифровой фальсификат, превращая его в «идеальное» орудие для манипуляции правосудием.

Наряду с техническими и процессуальными ограничениями существенным препятствием для выявления контента, созданного с использованием технологий дипфейк, выступают когнитивные искажения, свойственные участникам уголовного судопроизводства, которые создают условия, при которых синтетический контент может оставаться нераспознанным и использоваться для формирования ложных выводов.

Во-первых, фундаментальной проблемой является «эффект фотографической достоверности». Аудиовизуальная информация на подсознательном уровне воспринимается человеческим мозгом как объективное и неоспоримое свидетельство, фиксация реальности. Эта эвристика, сформированная в доцифровую эпоху, делает судей, прокуроров и присяжных чрезвычайно уязвимыми. Внутреннее убеждение, формируемое в соответствии со ст. 17 УПК РФ, оказывается основано на иллюзии, так как непосредственно исследуется не реальное событие, а его высокотехнологичная симуляция. Экспериментальные исследования подтверждают, что видеосвидетельства оказывают исключительно сильное психологическое воздействие, превосходящее даже логические контраргументы. В контексте российского процесса это вступает в противоречие с принципом оценки доказательств по внутреннему убеждению, основанному на их совокупности, однако на практике видеодоказательство нередко приобретает непропорциональное влияние, несоразмерное его доказательственной силе

Во-вторых, существует риск недооценки сложности выявления дипфейков со стороны правоприменителей: распространено мнение, что любую подделку можно быстро выявить стандартным программным обеспечением или беглым осмотром специалиста. Из этого заблуждения вырастает недооценка рисков и формальный подход к назначению экспертиз. Возможна ситуация, когда суд отклоняет ходатайство защиты о комплексной экспертизе, ограничиваясь вопросом: «имеются ли видимые признаки монтажа?». На такой вопрос при качественном дипфейк последует отрицательный ответ – и формальный повод для отказа найден. Представим реальную ситуацию: защита настаивает на экспертизе видеозаписи, указывая на её ключевую роль для обвинения и ссылаясь на современные технологии фальсификации. Судья, ссылаясь на ст. 284 УПК РФ, отказывает, мотивируя это тем, что «сторона защиты не представила конкретных доказательств, ставящих под сомнение подлинность видеозаписи». Тем самым на защиту фактически возлагается бремя доказывания отрицательного факта, а необходимость применять к таким доказательствам повышенный стандарт проверки достоверности остаётся за скобками.

Третье препятствие – низкий уровень цифровой грамотности в профессиональной юридической среде и слабая осведомлённость о реальных возможностях генеративных нейросетей. Без понимания базовых принципов их работы затруднительно ни сформулировать эксперту корректные вопросы, ни оценить полноту и обоснованность полученного заключения, ни, наконец, заподозрить подлог на ранней стадии. Пример из следственной практики: следователь, знакомый лишь с классическими признаками монтажа (скачки кадров, рассинхронизированность аудио), изучает предъявленную видеозапись. Дипфейк-видео этих артефактов не содержит – оно выглядит как целостный, технически безупречный поток данных. Следователь, не обнаруживая классических признаков монтажа, может прийти к выводу о достоверности записи, не назначая дополнительной проверки. Таким образом, возникает когнитивный парадокс: чем выше качество синтетической подделки, тем выше вероятность её некритического принятия правоприменителем.

В-четвёртых, в делах с участием присяжных срабатывает «эффект ореола». Убедительная видеозапись, представленная обвинением, подсознательно повышает кредит доверия и ко всем остальным доказательствам стороны обвинения – даже к тем, которые сами по себе слабы либо носят косвенный характер.

Сочетание этих психологических ловушек и процессуальных пробелов создаёт идеальные условия, повышающие риски злоупотреблений. Высококачественный дипфейк, попадая в суд, не просто вводит участников процесса в заблуждение – он эксплуатирует саму систему формирования судейского убеждения, заложенную в УПК РФ. Преодоление выявленных барьеров требует комплекса мер, включающего: техническое перевооружение, масштабную образовательную работу и, что важнее всего, внесение в процессуальные нормы поправок, прямо учитывающих уязвимость человеческого восприятия перед цифровыми симулякрами.

Внедрение технологий генеративного искусственного интеллекта в уголовное судопроизводство создаёт не частную проблему для отдельных институтов криминалистики или процессуального права, а системный риск, затрагивающий ключевые характеристики правосудия как публичного института. Потенциальные последствия использования дипфейков не исчерпываются возможностью вынесения ошибочного приговора по отдельному делу, но могут влиять на базовые предпосылки справедливого судебного разбирательства.

Начнём с принципа состязательности (ст. 15 УПК РФ). Он предполагает равенство процессуальных возможностей обвинения и защиты, однако при технологическом дисбалансе фактическое равенство возможностей сторон может отсутствовать. Тот, у кого есть ресурсы для создания качественного синтетического видео и для его квалифицированной детекции, получает несправедливое преимущество. Защита же, не имеющая доступа к современным средствам анализа, оказывается заведомо слабее. В таких условиях, принцип состязательности может утрачивать практическое содержание. Представим ситуацию: в деле о коррупции в распоряжение следствия через оперативно-розыскные каналы поступает видеозапись, на которой высокопоставленное лицо якобы получает взятку. Сторона защиты заявляет, что запись является высококачественным дипфейком, созданным неустановленными лицами с целью дискредитации обвиняемого, и ходатайствует о назначении комплексной экспертизы. Защита, не располагая доступом к специализированным экспертным учреждениям, оснащённым современными средствами детекции, лишена возможности эффективно оспорить такое доказательство. Суд, лишенный альтернативного аргументированного заключения, вынужден положить его в основу приговора. Технология становится инструментом недобросовестной процессуальной борьбы, нивелируя фактическое равенство сторон.

Во-вторых, искажается суть принципа непосредственности исследования доказательств (ст. 240 УПК РФ). Смысл данного принципа заключается в том, что суд обязан исследовать доказательства непосредственно: заслушивать показания, осматривать вещественные доказательства, оглашать протоколы. Дипфейк искажает существо данного принципа, подменяя подлинное доказательство его искусственной симуляцией. Судья и присяжные, полагая, что непосредственно воспринимают реальность, на самом деле воспринимают искусственно созданную аудиовизуальную симуляцию. Принцип состязательности формально соблюдается, но его содержательное наполнение – доверие к непосредственному восприятию – используется против правосудия.

В-третьих, меняется содержание принципа свободы оценки доказательств (ст. 17 УПК РФ). Внутреннее убеждение судьи формируется в результате беспристрастного анализа всей совокупности доказательств, однако, как было показано выше, убедительность дипфейк основана на когнитивных уязвимостях. Судебное познание, как и любое человеческое восприятие, подвержено когнитивным искажениям при столкновении с высокореалистичными синтетическими образами. Следовательно, его внутреннее убеждение может быть сформировано под влиянием заведомо ложного, но обладающего высокой степенью психологической убедительности доказательства. Легитимность самого решения суда ставится под сомнение, если оно основано на сфальсифицированном «факте».

Риск судебной ошибки существовал всегда, однако дипфейк придает ему качественно новый характер. Речь идет не о случайной ошибке, вызванной недостатком доказательств или добросовестным заблуждением, а о целенаправленном, труднообнаружимом и потому крайне опасном внедрении в процесс ложной информации, обладающей высокой степенью реалистичности. Если традиционные технологии служили средством фиксации и установления фактических обстоятельств, то дипфейк может использоваться для их искажения. Не исключена ситуация, когда через несколько лет потребуется массовый пересмотр приговоров, вынесенных на основе нераспознанных в свое время цифровых фальсификатов, что приведет к существенному снижению институционального доверия к судебной власти.

Легитимность судебной власти в значительной степени обусловлена уровнем общественного доверия к ней. Уверенность граждан в том, что суд является институтом, способным установить объективную истину и вынести справедливое решение, – основа легитимности судебной власти. Широкое распространение информации о случаях использования дипфейков в судебных процессах, независимо от их реального масштаба, приведет к значительному снижению этого доверия. Любой обвинительный приговор может быть подвергнут сомнению со стороны общественности: «не основан ли он на синтетическом аудиовизуальном материале?». У стороны защиты появляется процессуальный ресурс – ссылаться на саму возможность технической подделки. И даже если эта вероятность объективно низка, сам факт такого апеллирования формирует в правосознании устойчивый скепсис: любой судебный акт может быть поставлен под сомнение.

Распространение подобного отношения объективно затрудняет отправление правосудия, а значит, деструктивный потенциал дипфейков не исчерпывается способностью обмануть конкретного судью или присяжных. Наиболее значимым представляется другое: технология подтачивает системообразующие принципы уголовного процесса – состязательность, непосредственность, свободу оценки доказательств. Как справедливо отмечает И.Г. Смирнова, «категория доверия имеет ярко выраженную уголовно-процессуальную природу, поскольку является базисом процесса доказывания в целом и оценки доказательств в частности». Следовательно, когда дипфейк ставит под сомнение достоверность доказательства, он подрывает не частный элемент, а само основание уголовно-процессуальной деятельности. Отвечать на вызовы генеративного искусственного интеллекта нужно не точечными поправками, а выстраиванием целостной, многоуровневой системы защиты, сочетающей три компонента: превентивные технологические барьеры, адаптацию процессуальных институтов и планомерное повышение цифровой компетентности юридического сообщества.

Ключевой задачей является недопущение проникновения фальсифицированных материалов в уголовное дело, а также создание условий для их быстрой и объективной верификации. В рамках проведенного исследования предлагается следующее:

Первым и минимально необходимым шагом должно стать законодательное закрепление обязательного хеширования цифровых доказательств при их приобщении к материалам дела. Профессор И. Америни выделяет два

типа аутентификации: пассивную, через анализ артефактов после факта и активную – встраивание верифицируемых меток при создании контента. Для уголовного процесса наиболее перспективна активная аутентификация. Предлагается при изъятии цифровых носителей производить расчёт хеш-сумм с сертифицированным ПО и, по возможности, вносить их в распределённый блокчейн-реестр. Однако, как подчёркивает исследователь, такой метод защищает только от посторонних изменений после изъятия, но не от изначальной подделки, исходя из чего хеширование должно быть лишь первым уровнем. Второй уровень – обязательная комплексная экспертиза с использованием нескольких независимых алгоритмов детекции, например, одновременно CNN и частотный анализ, как это предлагается в проекте FF4ALL. Статью 166 УПК РФ следует дополнить частью следующего содержания: «При изъятии электронных носителей информации либо копировании файлов, имеющих значение для уголовного дела, следователь (дознаватель) обязан произвести расчёт хеш-суммы каждого файла с использованием сертифицированных средств криптографической защиты и отразить полученное значение в протоколе следственного действия. Копия протокола незамедлительно вручается лицу, у которого произведено изъятие, либо его представителю». Данная мера позволит в последующем достоверно установить, подвергался ли файл модификации после его поступления в распоряжение следствия.

При этом необходимо ясно осознавать границы данного механизма: хеширование удостоверяет лишь неизменность файла с момента его фиксации, но не способно выявить изначальную подделку, созданную до изъятия. Предложение о внесении хеш-сумм в распределённый реестр (блокчейн) может рассматриваться как перспективное направление для ведомственных информационных систем, однако на текущем этапе его реализация сопряжена со значительными инфраструктурными и бюджетными издержками, а также требует решения вопроса о равном доступе сторон защиты и обвинения к такой системе, что исключает её немедленное внедрение.

Вторым шагом является институциональное усиление кадрового потенциала органов расследования и суда. Предлагается поэтапное введение в штат следственных подразделений, специализирующихся на расследовании преступлений в сфере высоких технологий, должностей специалистов по цифровой криминалистике. Их процессуальный статус должен определяться ст. 58 и 168 УПК РФ. В функции такого специалиста при осмотре места происшествия или производстве выемки электронных носителей должны входить:

1) создание криминалистического образа носителя на месте;

2) расчёт и фиксация хеш-сумм всех значимых файлов;

3) экспресс-анализ аудиовизуальных материалов с использованием аккредитованного программного обеспечения на предмет выявления явных признаков синтетического происхождения;

4) консультирование следователя при формулировании вопросов для последующей судебной экспертизы.

Результаты такого первичного исследования подлежат отражению в протоколе следственного действия.

Аналогичная потребность существует и в судебной системе: при рассмотрении дел, в которых фигурируют цифровые аудиовизуальные доказательства, суду должна быть обеспечена возможность привлечения такого специалиста для дачи разъяснений в порядке ст. 58 УПК РФ. Это позволит судье более компетентно оценивать обоснованность ходатайств о назначении экспертизы и формулировать вопросы экспертному учреждению.

Реализация данного предложения потребует разработки образовательных программ на базе ведомственных учебных заведений и корректировки штатных расписаний, однако эти издержки оправданы масштабом потенциального вреда от судебных ошибок, вызванных нераспознанными синтетическими доказательствами.

Третьим шагом выступает дифференцированное перераспределение бремени доказывания, исходя из повышенных рисков фальсификации цифрового контента, при избежании излишней процессуальной нагрузки. В основу предлагаемого механизма должен быть положен принцип разумной достаточности и состязательности. Речь идет о следующем подходе:

Представляется обоснованным установить повышенный стандарт проверки достоверности цифровых аудио- и видеодоказательств, дифференцированный в зависимости от их процессуального значения. В отношении ключевых доказательств – то есть тех, на которых исключительно или преимущественно строится обвинение и которые не подтверждены иными независимыми источниками, – при заявлении стороной мотивированного ходатайства о возможной фальсификации с использованием технологий синтетического медиаконтента суду надлежит назначать обязательную комплексную экспертизу. Если сторона обвинения основывает обвинение на аудиовизуальном доказательстве, а защита заявляет мотивированное ходатайство о его фальсификации с использованием дипфейк-технологий, обвинение обязано обеспечить проведение комплексной экспертизы; уклонение от проверки влечёт исключение данного доказательства из основы обвинительного приговора. В отношении иных цифровых доказательств достаточным является требование об обязательном раскрытии метаданных, хеш-суммы и сведений об устройстве записи. Отсутствие указанных верификационных данных может служить основанием для признания доказательства недопустимым по ходатайству стороны или инициативе суда (ст. 235 УПК РФ).

Предлагается дополнить статью 88 УПК РФ частью 2.1 следующего содержания:

«Достоверность аудиовизуальных материалов, полученных в электронной форме и представленных в качестве доказательств, при наличии мотивированных сомнений в их аутентичности, вызванных возможным применением технологий синтетического медиаконтента, подлежит обязательной проверке путём проведения судебной экспертизы. Отсутствие такой проверки исключает возможность использования указанных материалов в качестве основы обвинительного приговора или иного итогового судебного решения».

Четвертым шагом необходимо ввести обязательную многоуровневую судебную экспертизу. Для проверки цифровых доказательств должна назначаться комплексная экспертиза, объединяющая компетенции эксперта-IT-специалиста (цифрового криминалиста) (проверка метаданных, истории файла, признаков перекодирования, анализ на соответствие хеш-суммам), эксперта-лингво-акустика (анализ речевой и неречевой составляющей аудиодорожки на предмет артефактов синтеза), эксперта-видеотехника/фототехника (анализ видеоряда на предмет физической и биологической непротиворечивости (естественность мимики, моргания, синхронность аудио и видео, соответствие освещения и физики движения)).

В организационном плане представляется необходимым создание в системе государственных судебно-экспертных учреждений специализированных лабораторий, оснащённых несколькими независимыми программными комплексами детекции синтетического контента и обеспеченных возможностью их регулярного обновления. Однако ещё более значимой является процессуальная регламентация самого порядка проведения таких экспертиз.

В условиях, когда жёстко утверждённая методика неизбежно устаревает быстрее, чем проходит законодательная процедура её актуализации, предлагается ввести не фиксированный набор методик, а процессуальный стандарт комплексной цифровой экспертизы. Для этого статью 204 УПК РФ следует дополнить требованием об обязательном отражении в заключении эксперта следующих сведений:

– перечень использованных программных и аппаратных средств, их версии и даты актуальности;

– описание применённых алгоритмов детекции и их происхождения (проприетарные, открытые, коммерческие);

– проводилась ли верификация полученных результатов альтернативными методами;

– оценка достоверности выводов с указанием известных ограничений использованных технологий.

Такой подход не сковывает экспертное сообщество устаревающей методикой, но при этом создаёт правовые условия для оценки научной обоснованности заключения судом и сторонами – в том числе через призму критериев проверяемости и рецензируемости, известных зарубежной практике (стандарт Daubert).

Проведённое исследование позволяет сделать вывод, что распространение технологий синтетического аудиовизуального контента порождает качественно новый вызов для уголовного судопроизводства. Специфика этого вызова заключается в синергетическом эффекте, возникающем при наложении факторов технической непрозрачности алгоритмов генерации и детекции, процессуальных ограничений и когнитивных уязвимостей субъектов доказывания. В совокупности названные обстоятельства способны деформировать реализацию принципов состязательности, непосредственности и свободной оценки доказательств.

В сложившихся условиях минимально необходимыми мерами представляются:

1) законодательное закрепление обязательного хеширования цифровых доказательств при их приобщении к материалам дела;

2) введение повышенного стандарта проверки достоверности аудиовизуальных материалов посредством комплексной судебной экспертизы, назначаемой при наличии мотивированных сомнений в их аутентичности;

3) дополнение статьи 204 УПК РФ требованиями к раскрытию использованных экспертом программных средств и методик;

4) институциональное укрепление экспертных и следственных подразделений специалистами по цифровой криминалистике;

5) системное повышение цифровой грамотности судей, прокуроров, следователей и адвокатов.

Адаптация уголовного процесса к цифровой эпохе требует не отказа от фундаментальных принципов, а их последовательного укрепления через создание механизмов, гарантирующих достоверность доказательственной базы. Предложенный комплекс мер позволяет с разумной степенью надёжности выявлять синтетические доказательства либо, при невозможности такого выявления, исключать их из основы итогового судебного решения, обеспечивая баланс публичных и частных интересов в условиях, когда доверие к аудиовизуальному контенту более не может презюмироваться.

Литература

1. Америни И., Барни М., Баттиато С. [и др.]. Медиакриминалистика дипфейков: текущее состояние и будущие вызовы // J Imaging. 2025. Т. 11, № 3. С. 73.

2. Касперский Е. В. Deepfake, дипфейк // Лаборатория Касперского : [сайт]. URL: https://encyclopedia.kaspersky.ru/glossary/deepfake/ (дата обращения: 10.12.2025).

3. Кольцова О. Ю., Киркиж Э. А. Влияние интернета на участие в протестах // Полития. Анализ. Хроника. Прогноз. 2016. № 1 (80). С. 90–110.

4. Таранова Ю. В., Гавра Д. П. Применение искусственного интеллекта в территориальном брендинге: возможности и вызовы // Российская школа связей с общественностью. 2024. № 32.

5. Воронин И. А., Гавра Д. П. Дипфейки: современное понимание, подходы к определению, характеристики, проблемы и перспективы // Российская школа связей с общественностью. 2024. № 33. URL: https://cyberleninka.ru/article/n/dipfeyki-sovremennoe-ponimanie-podhody-k-opredeleniyu-harakteristiki-problemy-i-perspektivy (дата обращения: 07.12.2025).

6. Смирнова И. Г. Доверие как аксиологическая детерминанта процесса доказывания по уголовным делам: эмодзи как объект оценки судом // Всероссийский криминологический журнал. 2024. № 5 (83). С. 504–513.

7. Софийчук Н. В. К вопросу о значении доверия в уголовно-процессуальном регулировании // Вестник Университета имени О. Е. Кутафина (МГЮА). 2024. № 1 (113). С. 67–74.

8. Старостенко К. В., Коновалов А. С. Дипфейк как инновационный инструмент политического манипулирования в начале XXI века // Известия Саратовского университета. Новая серия. Серия: Социология. Политология. 2025. Т. 25, вып. 2. С. 195–201.

9. Aktions plan Deepfake Wien. – Wien, 2022. – 30 S. – (III-740 der Beilagen XXVII. GP – Bericht – 02 Hauptdokument).

10. Albahar M., Almalki J. Deepfakes: Threats and countermeasures systematic review // Journal of Theoretical and Applied Information Technology. 2019. Vol. 97. № 22. P. 3242–3250.

11. Appel M., Prietzel F. The detection of political deepfakes // Journal of Computer-Mediated Communication. 2022. Vol. 27. № 4.

12. Bateman J. Deepfakes and synthetic media in the financial system: Assessing threat scenarios. – Washington : Carnegie Endowment for International Peace, 2022.

13. Brewster T. Fraudsters Cloned Company Director's Voice In $35 Million Bank Heist, Police Find // Forbes. 2021. 14 Oct. URL: https://www.forbes.com/sites/thomasbrewster/2021/10/14/huge-bank-fraud-uses-deep-fake-voice-tech-to-steal-millions/ (дата обращения: 17.01.2026).

14. Boulianne S., Lee S. Conspiracy beliefs, misinformation, social media platforms, and protest participation // Media and Communication. 2022. Vol. 10. № 4. P. 30–41.

15. Case Against 'Deepfake Mom' Falls Apart After Prosecutors Backtrack // The Daily Dot. URL: https://www.dailydot.com/debug/deepfake-mom-cheerleaders-prosecutors-backtrack/ (дата обращения: 10.12.2025).

16. Clark S., Lewandowsky S. The continued influence of AI-generated deepfake videos despite transparency warnings // Communications Psychology. 2026. Vol. 4. Article 13.

17. Delfino R. A. Pornographic deepfakes: The case for federal criminalization of revenge porn's next tragic act // Fordham Law Review. 2019. Vol. 88. P. 887.

18. Dobber T. et al. Do (microtargeted) deepfakes have real effects on political attitudes? // The International Journal of Press/Politics. 2021. Vol. 26. № 1. P. 69–91.

19. Dolhansky B. et al. The deepfake detection challenge (dfdc) dataset // arXiv.org. 2020. URL: https://arxiv.org/abs/2006.07397 (дата обращения: 10.12.2025).

20. Goodfellow I. et al. Generative adversarial nets // Advances in neural information processing systems. 2014. Vol. 27.

21. Kietzmann J. et al. Deepfakes: Trick or treat? // Business Horizons. 2020. Vol. 63. № 2. P. 135–146.

22. Kshetri N. The economics of deepfakes // Computer. 2023. Vol. 56. № 8. P. 89–94.

23. Masood M. et al. Deepfakes generation and detection: State-of-the-art, open challenges, countermeasures, and way forward // Applied intelligence. 2023. Vol. 53. № 4. P. 3974–4026.

24. Mirsky Y., Lee W. The creation and detection of deepfakes: A survey // ACM computing surveys (CSUR). 2021. Vol. 54. № 1. P. 1–41.

25. Stupp C. Fraudsters Used AI to Mimic CEO's Voice in Unusual Cybercrime Case // The Wall Street Journal. 2019. 30 Aug. URL: https://www.wsj.com/articles/fraudsters-use-ai-to-mimic-ceos-voice-in-unusual-cybercrime-case-11567157402 (дата обращения: 17.01.2026).

26. Text – H.R.3230 – 116th Congress (2019-2020): DEEP FAKES Accountability Act. URL: https://www.congress.gov/bill/116th-congress/house-bill/3230/text (дата обращения: 10.12.2025).

27. Vromen A., Xenos M. A., Loader B. Young people, social media and connective action: From organisational maintenance to everyday political talk // Journal of Youth Studies. 2015. Vol. 18. № 1. P. 80–100.

28. Westerlund M. The emergence of deepfake technology: A review // Technology innovation management review. 2019. Vol. 9. № 11.

29. Vincent J. Watch Jordan Peele Use AI to Make Barack Obama Deliver a PSA About Fake News // The Verge. 2018. 17 April. URL: https://www.theverge.com/tldr/2018/4/17/17247334/ai-fake-news-video-barack-obama-jordan-peele-buzzfeed (дата обращения: 10.12.2025).

The impact of deepfake technologies on evidence assessment: a crisis of trust

Khubieva M.M.
Kutafin Moscow State Law University (MSAL)

Synthetic audiovisual content (deepfake) technologies pose an unprecedented challenge to criminal justice, undermining fundamental principles of evidence evaluation. The article examines algorithmically generated synthetic media as a qualitatively new destructive factor in criminal proceedings, systematically distorting evidentiary information. Drawing on foreign practice and Russian law enforcement, the study identifies systemic risks of deepfake use, with particular attention paid to technical, procedural, and psychological barriers to their detection. The methodological basis of the research comprises the dialectical method of scientific cognition, general scientific methods of analysis, synthesis, induction, deduction, and the systemic-structural approach, as well as specific legal methods: formal-legal, comparative-legal, and legal modeling. The empirical base includes foreign and domestic judicial practice materials and interdisciplinary doctrinal sources at the intersection of criminal procedure law, criminalistics, and computer science. The author argues that the historically formed presumption of trust in audiovisual evidence must give way to a heightened standard of reliability verification. A comprehensive set of measures is proposed: legislative codification of mandatory hashing of digital evidence; introduction of a mandatory comprehensive judicial examination where there are reasoned doubts as to authenticity; amendment of Article 204 of the Russian Criminal Procedure Code with requirements for disclosure of expert methodologies; institutional strengthening of expert and investigative units with digital forensics specialists; and systematic improvement of digital literacy among judges, investigators, and lawyers. The conclusion is drawn that timely and comprehensive counteraction to the threats posed by synthetic audiovisual content technologies is a key condition for maintaining public trust in justice in the digital era.

Keywords: criminal proceedings, evidence, evaluation of evidence, relevance of evidence, reliability of evidence, admissibility of evidence, falsification of evidence, deepfake, digital technologies, forensic expertise, and digital forensics.

References

1. Amerini I., Barney M., Battiatto S. [et al.]. Media Forensics of Deepfakes: Current State and Future Challenges // J Imaging. 2025. Vol. 11, No. 3. P. 73.

2. Kaspersky E. V. Deepfake // Kaspersky Lab : [website]. URL: https://encyclopedia.kaspersky.ru/glossary/deepfake/ (accessed: 10.12.2025).

3. Koltsova O. Yu., Kirkizh E. A. The Influence of the Internet on Protest Participation // Politiya. Analysis. Chronicle. Forecast. 2016. No. 1 (80). P. 90–110.

4. Taranova Yu. V., Gavra D. P. Application of Artificial Intelligence in Territorial Branding: Opportunities and Challenges // Russian School of Public Relations. 2024. No. 32.

5. Voronin I. A., Gavra D. P. Deepfakes: Modern Understanding, Approaches to Definition, Characteristics, Problems and Prospects // Russian School of Public Relations. 2024. No. 33. URL: https://cyberleninka.ru/article/n/dipfeyki-sovremennoe-ponimanie-podhody-k-opredeleniyu-harakteristiki-problemy-i-perspektivy (accessed: 07.12.2025).

6. Smirnova I. G. Trust as an Axiological Determinant of Evidence in Criminal Cases: Emoji as an Object of Court Evaluation // Russian Criminological Journal. 2024. No. 5 (83). P. 504–513.

7. Sofiychuk N. V. On the Importance of Trust in Criminal Procedure Regulation // Bulletin of the Kutafin Moscow State Law University (MSAL). 2024. No. 1 (113). P. 67–74.

8. Starostenko K. V., Konovalov A. S. Deepfake as an Innovative Tool of Political Manipulation in the Early 21st Century // Izvestiya of Saratov University. New Series. Series: Sociology. Political Science. 2025. Vol. 25, iss. 2. P. 195–201.

9. Aktions plan Deepfake Wien. – Wien, 2022. – 30 S. – (III-740 der Beilagen XXVII. GP – Bericht – 02 Hauptdokument).

10. Albahar M., Almalki J. Deepfakes: Threats and Countermeasures Systematic Review // Journal of Theoretical and Applied Information Technology. 2019. Vol. 97. No. 22. P. 3242–3250.

11. Appel M., Prietzel F. The Detection of Political Deepfakes // Journal of Computer-Mediated Communication. 2022. Vol. 27. No. 4.

12. Bateman J. Deepfakes and Synthetic Media in the Financial System: Assessing Threat Scenarios. – Washington : Carnegie Endowment for International Peace, 2022.

13. Brewster T. Fraudsters Cloned Company Director's Voice In $35 Million Bank Heist, Police Find // Forbes. 2021. Oct. 14. URL: https://www.forbes.com/sites/thomasbrewster/2021/10/14/huge-bank-fraud-uses-deep-fake-voice-tech-to-steal-millions/ (accessed: 17.01.2026).

14. Boulianne S., Lee S. Conspiracy Beliefs, Misinformation, Social Media Platforms, and Protest Participation // Media and Communication. 2022. Vol. 10. No. 4. P. 30–41.

15. Case Against 'Deepfake Mom' Falls Apart After Prosecutors Backtrack // The Daily Dot. URL: https://www.dailydot.com/debug/deepfake-mom-cheerleaders-prosecutors-backtrack/ (accessed: 10.12.2025).

16. Clark S., Lewandowsky S. The Continued Influence of AI-Generated Deepfake Videos Despite Transparency Warnings // Communications Psychology. 2026. Vol. 4. Article 13.

17. Delfino R. A. Pornographic Deepfakes: The Case for Federal Criminalization of Revenge Porn's Next Tragic Act // Fordham Law Review. 2019. Vol. 88. P. 887.

18. Dobber T. et al. Do (Microtargeted) Deepfakes Have Real Effects on Political Attitudes? // The International Journal of Press/Politics. 2021. Vol. 26. No. 1. P. 69–91.

19. Dolhansky B. et al. The Deepfake Detection Challenge (DFDC) Dataset // arXiv.org. 2020. URL: https://arxiv.org/abs/2006.07397 (accessed: 10.12.2025).

20. Goodfellow I. et al. Generative Adversarial Nets // Advances in Neural Information Processing Systems. 2014. Vol. 27.

21. Kietzmann J. et al. Deepfakes: Trick or Treat? // Business Horizons. 2020. Vol. 63. No. 2. P. 135–146.

22. Kshetri N. The Economics of Deepfakes // Computer. 2023. Vol. 56. No. 8. P. 89–94.

23. Masood M. et al. Deepfakes Generation and Detection: State-of-the-Art, Open Challenges, Countermeasures, and Way Forward // Applied Intelligence. 2023. Vol. 53. No. 4. P. 3974–4026.

24. Mirsky Y., Lee W. The Creation and Detection of Deepfakes: A Survey // ACM Computing Surveys (CSUR). 2021. Vol. 54. No. 1. P. 1–41.

25. Stupp C. Fraudsters Used AI to Mimic CEO's Voice in Unusual Cybercrime Case // The Wall Street Journal. 2019. Aug. 30. URL: https://www.wsj.com/articles/fraudsters-use-ai-to-mimic-ceos-voice-in-unusual-cybercrime-case-11567157402 (accessed: 17.01.2026).

26. Text – H.R.3230 – 116th Congress (2019-2020): DEEP FAKES Accountability Act. URL: https://www.congress.gov/bill/116th-congress/house-bill/3230/text (accessed: 10.12.2025).

27. Vromen A., Xenos M. A., Loader B. Young People, Social Media and Connective Action: From Organisational Maintenance to Everyday Political Talk // Journal of Youth Studies. 2015. Vol. 18. No. 1. P. 80–100.

28. Westerlund M. The Emergence of Deepfake Technology: A Review // Technology Innovation Management Review. 2019. Vol. 9. No. 11.

29. Vincent J. Watch Jordan Peele Use AI to Make Barack Obama Deliver a PSA About Fake News // The Verge. 2018. Apr. 17. URL: https://www.theverge.com/tldr/2018/4/17/17247334/ai-fake-news-video-barack-obama-jordan-peele-buzzfeed (accessed: 10.12.2025).