Galactica Media: Journal of Media Studies. 2026. No 4 | ISSN: 2658-7734
AI, Technology & Culture | https://doi.org/10.46539/gmd.v8i4.727

The Existential Dimensions of the Deepfake Phenomenon in Digital Culture

Liudmila V. Baeva*

Astrakhan Tatischev State University; Institute of Philosophy, Russian Academy of Science.
Astrakhan, Russia. Email: baevaludmila[at]mail.ru ORCID https://orcid.org/0000-0003-0439-525X

* — corresponding author

Received: 9 January 2026 | Revised: 7 March 2026 | Accepted: 15 March 2026

Abstract

Digital culture and neural networks give rise to ambivalent phenomena, including deepfakes, which pose a high risk given the widespread accessibility of this technology. The use of deepfakes to construct false realities for destructive purposes reshapes questions concerning human existence and transcendence, the openness and authenticity of communication, and the safeguarding of personal uniqueness. This article examines the existential, communicative, and sociocultural aspects of the use of deepfake technology. It also characterizes deepfakes as instruments of information and hybrid warfare directed against both individuals and social systems.

This study conceptualizes the deepfake phenomenon within a system of diverse interrelations and levels of personal being—technological, psychological, mass-mediated, political, ethical, and aesthetic—thereby demonstrating both its multifaceted capabilities and the substantial threats it poses, especially when applied to destructive social practices.

The findings may be used to inform regulatory policy in media production and artificial intelligence, as well as humanities-based assessments of contemporary technological solutions.

Keywords

Deepfake; Digital Culture; Existentiality and Transcendence; Media Product; Virality; Reality and Quasi-Reality; Manipulation; Disinformation; Uniqueness

This work is licensed under a Creative Commons “Attribution” 4.0 International License
© 2026 Liudmila V. Baeva

Galactica Media: Journal of Media Studies. 2026. № 4 | ISSN: 2658-7734
ИИ, технологии и культура | https://doi.org/10.46539/gmd.v8i4.727

Экзистенциальные аспекты феномена дипфейка в цифровой культуре

Баева Людмила Владимировна*

Астраханский государственный университет им. В.Н. Татищева; Институт философии РАН. Астрахань, Россия. Email: baevaludmila[at]mail.ru ORCID https://orcid.org/0000-0003-0439-525X

* — корреспондирующий автор

Рукопись получена: 9 января 2026 | Пересмотрена: 7 марта 2026 | Принята: 15 марта 2026

Аннотация

Цифровая культура и нейросети порождают амбивалентные феномены, в том числе дипфейки, обладающие высоким рискогенным потенциалом в условиях широкой доступности соответствующих технологий. Конструирование ложной реальности с помощью дипфейков с деструктивными целями обозначает новые аспекты проблем экзистенции и трансценденции личности, открытости и подлинности ее коммуникации, сохранения ее уникальности. В статье рассматриваются экзистенциальные, коммуникативные и социокультурные аспекты использования технологии дипфейка. Охарактеризовано применение дипфейка как инструмента информационной и гибридной войны, направленной как против отдельной личности, так и против социальных систем.

Исследование позволяет осмыслить феномен дипфейка в системе разнообразных взаимосвязей и уровней бытия личности — технологического, психологического, массмедийного, политического, этического и эстетического, — и выявить многообразие его возможностей, а также высокий рискогенный потенциал, обусловленный использованием технологии в деструктивных социальных практиках.

Результаты исследования могут быть использованы при разработке мер нормативного регулирования в сфере медиапроизводства и ИИ, а также при проведении гуманитарной экспертизы современных технологических решений.

Ключевые слова

дипфейк; цифровая культура; экзистенция и трансценденция; медиапродукт; вирусный эффект; реальность и квазиреальность; манипуляция; дезинформация; уникальность

Это произведение доступно по лицензии Creative Commons “Attribution” («Атрибуция») 4.0
© 2026 Баева Людмила Владимировна

– Ну как можно поверить в невозможное?
– Постепенно
(Кэрролл Л. «Алиса в зазеркалье»)

Введение

Цифровая культура отличается высокой динамичностью и амбивалентностью, а также появлением новых социокультурных явлений в ответ на быстрое развитие и глобальное распространение технологий. Феномены, порождаемые цифровой культурой, обладают неодинаковыми характеристиками, масштабами и интенциями, служат различным целям и играют все более значимую (во многих отношениях доминирующую) роль в жизненном мире современного человека. С развитием технологий искусственного интеллекта и «глубокого обучения» в современной культуре сформировался поток информационной продукции принципиально нового типа, создаваемой генеративными нейросетями под управлением человека. Одним из таких явлений стали так называемые «дипфейки» (от англ. Deepfake: Deep — глубокий, Fake — подделка) — аудиовизуальные медиапродукты, созданные с помощью генеративного ИИ, в которых образы персонажей искусственно сконструированы, умышленно искажены или сфальсифицированы, но при этом практически неотличимы от реальных изображений и видеозаписей; такие медиапродукты, как правило, используются в деструктивных и манипулятивных практиках. Эти медиапродукты наносят существенный моральный вред личности, с одной стороны, а с другой — способствуют росту социальной напряженности, возникновению конфликтов и разжиганию ненависти. Диапазон их воздействия простирается от развлечения пользователей с помощью забавных видеороликов до введения общественности в заблуждение, способного повлечь политические или экономические последствия. При этом потенциал этой технологии еще далеко не полностью реализован, и по мере совершенствования технологии и усиления информационных войн рискогенный потенциал и разрушительное воздействие дипфейков будут возрастать. В 2025 году масштабы производства дипфейков в мире выросли в несколько раз, а само явление приобрело массовый и резонансный характер. Технологии, ранее доступные только специалистам, стали доступны рядовым пользователям через нейросетевые сервисы. Этот сдвиг приобрел революционное значение не только в технологическом, но и в социальном плане. Проблематика дипфейков носит междисциплинарный характер и находится на стыке различных областей знания, что актуализирует необходимость ее философского анализа. Такой анализ позволяет осмыслить данный феномен в системе многочисленных взаимосвязей и последствий, а также выработать общие методологические подходы к решению связанных с ним проблем.

В связи с этим цель исследования состоит в характеристике дипфейка как медиапродукта современной цифровой культуры и оценке рисков, которые

он создает для экзистенции и идентичности личности, а также для безопасности социальной коммуникации. Задачами исследования являются:

- анализ рисков для уникальности личности и ее существования в современном цифровом социуме;
- выявление деструктивных тенденций, порождаемых применением технологии дипфейка в массовой культуре, политических и социальных процессах;
- поиск способов регулирования возникающих при этом проблем.

Для решения поставленных задач мы будем опираться на существующие научные разработки, посвященные правовым и этическим аспектам создания и использования дипфейков, а также на положения теории экзистенциальной безопасности цифровой среды и теории коммуникации.

Исследование дипфейков в социогуманитарных науках

Изучение дипфейков как технологического продукта, а также методов их создания и генерации посредством программирования и обучения нейросетей выходит далеко за пределы предмета нашего исследования, в котором основное внимание будет сосредоточено на экзистенциальных, социальных, культурных, политических и нравственных аспектах этого сравнительно нового явления. В современных социальных и гуманитарных науках дипфейки уже стали предметом исследований, посвященных прежде всего связанным с ними правовым и этическим проблемам, а также вопросам медиакоммуникации. Можно выделить *информационно-правовой подход*, в рамках которого рассматриваются вопросы правового регулирования и мер ответственности за применение дипфейк-технологии (Архиереев, 2025; Бодров, 2024; Добробаба, 2022; Калятин, 2022; Киселёв, 2021; Малышева, 2025; Шубенкова, 2022; Bermudez Bermudez, 2025 и др.).

В рамках *информационно-политологического подхода* дипфейки рассматриваются как инструмент информационной борьбы как на внешнеполитической арене, так и в электоральных процессах (Виноградов, 2024; Воронин, 2024; Ефремова, 2024; Иванов, 2020; Лукина, 2023; Фалалеев, 2021; Nascimento, 2024; Dobber, 2021; Li, 2025 и др.).

С позиции *медиакоммуникативного подхода* дипфейк исследуется как медиакоммуникационный продукт, характеризующийся специфическими механизмами продвижения, распространения и восприятия в медиасреде, а также воздействия на аудиторию (Bainotti, 2022; Barari, 2025; Mishra, 2022; Булгарова, 2025 и др.).

В рамках *психологического подхода* дипфейк изучается как средство манипуляции сознанием личности, способное вводить ее в заблуждение, порождать неуверенность, страх и другие негативные состояния (Levine, 2019; Fox, 2009

и др.). Исследователи анализируют реакции людей с различными типами личности и уровнями когнитивных способностей на дипфейки и их распространение, отмечая, что люди с более высокими когнитивными способностями реже принимают дипфейки за достоверные и распространяют их (Ahmed & Chua, 2023). Исследователи также отмечали, что феномены виртуальной реальности способны встраиваться в память и вызывать ложные воспоминания (Segovia & Bailenson, 2009). Этот вывод в полной мере применим и к дипфейкам, малоотличимым от реальности.

Исследование дипфейка с позиции *социального взаимодействия* помогает глубже понять, какие социальные проблемы затрагивает эта технология, какие конфликты она способна провоцировать и какие социальные группы наиболее восприимчивы к аудио- и видеодипфейкам (Hancock & Bailenson, 2021; Nightingale & Farid, 2022 и др.).

Особое значение для нашего исследования имеет *философский подход*, в рамках которого рассматриваются прежде всего эпистемические и этические проблемы, связанные с феноменом дипфейков. Так, Л. Флориди, исследуя эстетические и творческие аспекты работы ИИ, называл сгенерированные художественные объекты, имитирующие подлинники, «эктипами» и подчеркивал, что они демонстрируют разрыв связи между автором и произведением. При этом исследователь высказывал надежду, что системы ИИ будут не только создавать копии, но и смогут их выявлять (Floridi, 2018). Д. Фейллиз также указывал, что помимо формирования ложных убеждений и подрыва обоснованности истинных убеждений, дипфейки могут препятствовать формированию у людей истинных убеждений (Fallis, 2021). В этом отношении технология дипфейков может привести к резкому росту связанных с ней эпистемологических издержек.

В целом можно констатировать устойчивый рост числа исследований, посвященных проблеме дипфейков в различных социогуманитарных науках, а также сформировавшееся среди исследователей представление о том, что это явление обладает растущим рискогенным потенциалом и требует пристального внимания, государственного регулирования, повышения медиаграмотности населения, а также разработки IT-сектором систем выявления, маркировки или блокировки дипфейков.

Далее рассмотрим ряд основных проблем, связанных с развитием дипфейков, и проанализируем их с позиций социокультурного, экзистенциального и медиакоммуникативного подходов на основе принципа дополнительности.

Риски личной безопасности и дипфейки в цифровой культуре

Проблема защиты личного пространства и частной жизни приобрела особую значимость в последние десятилетия в связи с развитием цифровой

коммуникации, социальных сетей и практик регистрации персональных данных. Риски утраты конфиденциальности и контроля над авторством, а также нарушения безопасности персональных данных стали своеобразной платой за свободный доступ к цифровым ресурсам, платформам и сетям. Первые практики создания дипфейков были связаны с незаконным использованием персональных данных, прежде всего изображения лица и голоса. Технологической основой создания дипфейков стали генеративно-состязательные сети (Generative Adversarial Networks, GAN), разработанные И. Гудфеллоу в 2014 г. Первые дипфейки появились в медиасреде в 2017 году, когда пользователь Reddit начал загружать видео, созданные с использованием методов «глубокого обучения» и содержавшие контент откровенного характера, полученный посредством наложения лиц известных актрис на порнографические материалы. Таким образом, дипфейк как инструмент изначально использовался для оскорбления личности и привлечения внимания массовой аудитории.

Современные исследователи констатируют, что по мере развития технологий ИИ отличить сфабрикованный контент от реального становится все сложнее. В этих условиях дипфейк-технология становится новой угрозой личной безопасности, поскольку она может применяться для кражи персональных данных, кибербуллинга, кибермошенничества, вымогательства и угроз. Особую опасность преступное использование дипфейков представляет для наиболее уязвимых групп населения — детей, пожилых людей и граждан, находящихся в трудной жизненной ситуации, поскольку они в большей степени подвержены манипулятивному воздействию.

Дипфейки активно применяются для введения человека в заблуждение и совершения в отношении него противоправных действий. Можно выделить следующие основные направления использования дипфейков против личности:

- посягательство на честь и достоинство;
- вымогательство;
- провоцирование межличностных конфликтов, агрессии и ненависти.

При создании дипфейка мишенью становятся индивидуальность личности и проявление ее Я во внешней среде, намеренно трансформируемое в образ с искаженными характеристиками. При использовании этой технологии ни один человек не обладает абсолютной защитой: каждый может стать жертвой обмана либо быть использован для введения в заблуждение других людей. В отношении публичных личностей характерны две основные практики: во-первых, использование их изображений для создания скандального и провокационного контента и, во-вторых, распространение дезинформации с помощью сфальсифицированных аудиовизуальных образов этих лиц. Для немедийных лиц дипфейк чаще всего становится инструментом травли

или вымогательства. Применение этой технологии способно причинить личности различные виды вреда — от морального и репутационного до материального.

Дипфейки выступают источником экзистенциальной угрозы для личности как уникального индивида, обладающего особым образом, посредством которого он манифестирует себя во внешнем мире, а также свободой выбора, мысли и вербального самовыражения. Эта уникальность является *основой* социальной коммуникации, выстраиваемой на доверии, симпатии, уважении и других формах межличностного отношения. Утрата контроля над выражением собственной индивидуальности и ее искажение создают риск разрыва социальных связей, а в крайних случаях — доведения «жертвы» дипфейка до суицида.

Правовое регулирование использования дипфейков при совершении преступлений против личности все еще находится на стадии формирования. В российском уголовном законодательстве действует норма, предусматривающая «ответственность за незаконные использование и (или) передачу, сбор и (или) хранение компьютерной информации, содержащей персональные данные» (ст. 272.1 УК РФ). В гражданском законодательстве предусмотрены нормы, направленные на охрану изображения гражданина и его частной жизни (ст. 152.1 «Охрана изображения гражданина», ст. 152.2 «Охрана частной жизни гражданина» ГК РФ). Если охрана изображения человека урегулирована законодательно, то специальное правовое регулирование использования или подделки голоса пока отсутствует. Исследователи правовых аспектов этой проблемы отмечают, что масштабы злонамеренного использования технологии дипфейков против личности будут возрастать, что обусловливает необходимость совершенствования законодательства (Киселев, 2021; Добробаба, 2022). При этом ряд исследователей полагает, что правовые основания для такого регулирования уже существуют, однако их необходимо дополнить механизмами обязательной маркировки и указания на искусственный характер медиаобъекта, а также его блокировки или удаления при наличии потенциальной угрозы (Архиреев, 2025).

Сферы применения дипфейков расширяются быстрее, чем развивается их правовое регулирование, что усиливает напряженность и недоверие в межличностном взаимодействии в цифровой среде.

Риски и угрозы, связанные с дипфейками в массовой культуре

Проникновение ИИ и дипфейков в массовую культуру и индустрию развлечений приобрело революционный характер и стало особенно заметным с 2021 года. В 2024–2025 годах на видеохостингах преобладающим видом контента становятся видеоматериалы, сгенерированные нейросетями. В лавинообразном информационном потоке дипфейки становятся частью массовой

культуры, распространяясь как локально, так и *вирусно*. Вирусность представляет собой одну из ключевых черт современной медиакультуры. Такой эффект способен принести всемирную известность создателю даже одного медиапродукта, набравшего миллионы просмотров, а также лишить общественного доверия и опорочить в глазах широкой аудитории того или иного деятеля, несмотря на его честность и прежние заслуги. Дипфейки в этом отношении оказываются весьма эффективными инструментами, активно применяемыми в различных сферах. Как отмечают исследователи применения этой технологии в киноиндустрии:

> «создавая гиперреалистичные визуальные эффекты, deepfakes расширяют рамки творчества, одновременно бросая вызов традиционным нормам кинопроизводства и демонстрируя дальнейший поворот индустрии к их интеграции с новыми технологиями и ценностями» (Булгарова и др., 2025).

Воссоздание образов ушедших из жизни культовых героев и актеров, а также создание ремейков фильмов и клипов с использованием технологии дипфейков расширяют возможности киноиндустрии и режиссерского творчества. Такие медиапродукты вызывают восхищение и интерес широкой аудитории, обеспечивая не только эмоциональный, но и коммерческий эффект. Сгенерированные образы умерших актеров используются при съемке новых частей фильмов с их участием (например, при воссоздании образа В. Галкина в фильме «Диверсант» 2022 г.). Технологии «омоложения» и «старения» активно внедряются в современный кинематограф вместо замены актеров (Т. Круз, Б. Питт, Х. Форд и др.). Созданные с помощью технологии дипфейков образы популярных кинозвезд начинают использовать в рекламе (Б. Уиллис для компании «МегаФон», Л. Куравлев в роли Жоржа Милославского в рекламе «Сбера»), что обусловило необходимость лицензирования использования внешности актера. Одним из первых подобных прецедентов стало лицензирование образа американского актера Б. Уиллиса.

Несмотря на позитивный отклик аудитории на новый сегмент массового развлекательного контента, у представителей творческих профессий ИИ вызывает определенные опасения. Технология дипфейков, получившая широкое применение в рекламе, кинематографе, музыкальной и развлекательной индустриях, была воспринята актерами как угроза творческим профессиям, поскольку результаты труда их представителей могут подменяться сгенерированными медиаобразами. Забастовки голливудских киноактеров в США в 2023–2024 годах широко освещались в СМИ, а само протестное движение получило значительный резонанс (Blanchet, 2025). Главное опасение актеров связано с обесцениванием их труда и возможностью замещения его результатов продуктами, сгенерированными с помощью ИИ. В 2025 году протесты против применения ИИ в киноиндустрии поддержали актеры Болливуда в Индии, выступившие против использования их изображений в дискредитирующих видеоматериалах (Reddy, 2025). В определенном смысле исполь-

зование дипфейков для замещения актеров и исполнителей либо воспроизведение их образов без согласия может рассматриваться как форма экзистенциальной угрозы личности, связанная с подрывом основ ее профессионального существования в том качестве, в котором она позиционирует себя в профессиональном сообществе.

С позиции медиатехнологий и эстетики создание дипфейков может рассматриваться как новый вид творчества и проявление профессионального мастерства создателей. Обучение нейросетей и совершенствование технологий создания дипфейков как цифровых медиапродуктов имеют не только коммерческие и технологические, но и правовые, эстетические и социальные аспекты. Требует ли правовой защиты авторство при создании дипфейка? Является ли он результатом творческой деятельности личности? Кто несет ответственность за создание такого объекта, его вирусное распространение и вызываемые им социальные последствия? Эти вопросы имеют выраженный междисциплинарный характер и требуют комплексного осмысления и выработки соответствующих решений. Однако еще более острые проблемы применение технологии дипфейков порождает в сфере публичной политики.

Дипфейки в политической культуре

Одной из главных мишеней создателей дипфейков являются политические лидеры, обладающие высокой степенью публичности. Именно публичность привлекает создателей фейковых видео, поскольку способствует увеличению числа их просмотров. Кроме того, политическая борьба в период избирательных кампаний формирует спрос на подобную продукцию. Как отмечает Н. Луман:

> «успех массмедиа в рамках всего общества основывается на достижении ими признания их тематик, которые сохраняют независимость от того, позитивную или негативную позицию занимают [массмедиа] по отношению к информации, смысловым интерпретациям, прорисовывающимся оценкам» (Луман, 2005, с. 25).

Дипфейк, как и дезинформация, направлен на подрыв общественного доверия, обесценивание достижений и положительного имиджа политика. Опасность представляют как дипфейки, в которых публичная личность намеренно высмеивается или оскорбляется, так и реалистичные видеоматериалы, в которых ее фальсифицированный образ произносит или совершает то, чего в действительности не было, тем самым дискредитируя эту личность. Даже специалисты зачастую не могут распознать видеоподделку; для массовой аудитории, уровень медиаграмотности которой остается средним или низким, эта задача практически недоступна. Исследователи отмечают, что чем ниже уровень когнитивных способностей человека, тем выше вероятность того, что он поверит дипфейку и впоследствии распространит его в социальных сетях (Ahmed & Chua, 2023). Когнитивные способности во многом определяют

степень развития критического мышления, которое позволяет сопоставлять визуально воспринимаемый сгенерированный образ с ранее известными достоверными сведениями и отличать ложное от действительного.

В современной политической практике создание дипфейков с участием политиков уже неоднократно становилось источником скандалов и деструктивных последствий. Дипфейки активно применялись в социальных сетях для разжигания ненависти и паники в ходе войны Израиля и ХАМАС, а также конфликта Израиля и Ирана, в том числе с целью дестабилизации политической власти в Иране (например, дипфейк, в котором были показаны иранцы с лозунгами «Мы любим Израиль»). Они также использовались в ходе военных действий на Украине (дипфейк 2022 г., в котором В. Зеленский призывал военных сложить оружие, фальсифицированное видео 2023 года, в котором президент В. Путин якобы сообщал о введении военного положения). В региональной медиасреде дипфейки с участием губернаторов стали использоваться для подрыва их репутации или провоцирования паники среди населения (например, медиафейк о «самоубийстве» В. Федорищева, губернатора Самарской области) (Крымова, 2025).

На фоне растущей угрозы распространения дипфейков представители государственных органов, публичные деятели, ученые и ведущие технологические компании неоднократно призывали к регулированию использования нейросетей и ИИ в целом. Одной из первых подобных инициатив стало обращение представителей науки и технологической отрасли, изложенное в документе «Открытое письмо. Приоритетные направления исследований в области надежного и полезного искусственного интеллекта» (An Open Letter..., 2015). Эта инициатива была продолжена авторами коллективного меморандума «Приостановить гигантские эксперименты с искусственным интеллектом: Открытое письмо» (Pause Giant AI Experiments..., 2023), основная идея которого состоит в том, что человек не должен создавать системы, если не уверен в их положительном воздействии и управляемости возникающих рисков.

Международный союз электросвязи (МСЭ) ООН в июле 2025 г. представил доклад, в котором подчеркивалось, что генеративный ИИ обладает трансформационным потенциалом и порождает такие риски, как дезинформация и дипфейки, определенные в качестве важнейших глобальных рисков современности. В докладе МСЭ были выделены различные типы вредоносного информационного контента, которые трудно отличить от реального (таблица 1).

В зависимости от типа субъектов, распространяющих дезинформацию, исследователи подразделяют дипфейки на государственные, корпоративные и гражданские (Фалалеев и др., 2021). Государственные и корпоративные источники обеспечивают более широкий охват аудитории, чем гражданские сообщества, а их деятельность оказывает более значительное воздействие.

Table 1. Types of disinformation, misinformation, and harmful information ("AI Standards for Global Impact...", 2025, p. 57)

Таблица 1. Типы дезинформации, ложной информации и вредоносной информации ("AI Standards for Global Impact...", 2025, с. 57)

Тип контента	Назначение	Характеристика
Фальсифицированный контент	Обычно на 100% ложный и предназначен для обмана и причинения вреда.	Отличить реальный контент от фальсифицированного крайне сложно. Воздействие сложных дипфейков, используемых для продвижения фальсифицированного контента, которое может серьезно подорвать доверие к получаемым сообщениям
Манипулированный контент	Подлинная информация или изображения, которые были искажены	В таких типах контента часто манипулируют подлинным контентом, подделывая изображение, используя сенсационные заголовки или кликбейт.
Контент-подделка	Выдача себя за подлинные источники, очень часто с использованием бренда известного агентства или авторитетного новостного агентства	Эта форма дезинформации использует доверие людей к определенной организации, бренду или даже к человеку
Вводящий в заблуждение контент	Вводящая в заблуждение информация создается путем переформулирования историй в заголовках	Обычно это использует фрагменты цитат для поддержки более широкой точки зрения, часто приводя статистические данные таким образом, чтобы соответствовать реальности
Ложный контекст	Переформулированный контент с ложным содержанием	Изображения с ложным контекстом – это низкотехнологичная, но все же мощная форма дезинформации
Сатира и пародия	Юмористические, но ложные сюжеты, выдаваемые за правду	В них нет явного намерения причинить вред, но аудитория может быть обманута и перестает различать правда и вымысел. Технология активно используется для преднамеренного распространения слухов и теорий заговора.

В политической сфере дипфейки стали применяться практически сразу после появления этой технологии в 2017 году (примерами служат дипфейки с участием Б. Обамы); за последние годы их число многократно возросло. Политические лидеры ведущих государств становятся объектами дипфейков, которые создаются как для персонального троллинга, так и для намеренной дестабилизации социальной системы. Дипфейк становится новым инструментом информационной, социально-психологической и гибридной войны, который получает все более широкое применение. Мгновенное распространение медиафейков в социальных сетях может привести к подрыву доверия к политическому лидеру и государственной системе в целом, формированию убежденности в достоверности ложной информации, а также к росту социальной напряженности, ненависти и паники. Этот вид медиапродукции чрезвычайно быстро трансформировался из потенциальной угрозы в реальную и может расцениваться как оружие, создающее угрозу национальной безопасности.

В электоральном процессе дипфейк становится мощным оружием в борьбе за власть, при этом моральные нормы взаимодействия между политическими силами практически стираются. Оскорбление чести и достоинства, искажение передаваемой информации и другие деструктивные практики становятся средствами «черного пиара», который получает дополнительный эффективный инструмент деструктивного воздействия.

Особый аспект, связанный с применением дипфейков, касается финансовых рисков. В этой сфере, как известно, даже изменения интонации в выступлениях ключевых политических и коммерческих лидеров может быть достаточно, чтобы вызвать значительные колебания рынка. Выделяют следующие основные типы рисков в этой сфере (“Deepfakes in the Financial Sector…”, 2024):

- Рыночный риск: ложная информация может использоваться для воздействия на финансовые рынки и манипулирования ими.
- Риск информационной безопасности: атака с использованием дипфейков может стать первым шагом, который позволит злоумышленникам проникнуть в системы учреждения и похитить конфиденциальные данные.
- Риск мошенничества: дипфейки как компонент социальной инженерии могут быть использованы для обмана клиентов учреждения или совершения атаки на само финансовое учреждение в рамках мошеннической схемы.
- Репутационный риск: дезинформация с использованием дипфейков может подорвать доверие потребителей и авторитет учреждения, нанеся долгосрочный ущерб его бренду и репутации.

Распространение дипфейков в финансовой сфере может вызвать паническую реакцию держателей акций, кризис доверия и даже обвал финансовых рынков. Противодействие дипфейкам становится, таким образом, не только задачей правоохранительных органов, но и банков, социальных сетей и цифровых платформ.

В целом следует подчеркнуть, что на политическом уровне, в условиях информационных войн и развития технологий управления страхом, мотивацией, уровнем агрессивности и другими психологическими состояниями, дипфейки становятся политическим оружием массового воздействия. Для финансовой сферы использование этой технологии становится новым источником нестабильности и угрозой финансовой безопасности граждан.

Дипфейк: удвоенная реальность

Феномены цифровой культуры во многом отражают основные направления жизнедеятельности человека, охватывая коммуникативно-социальную, экономическую, политическую, когнитивную, образовательную и рекреационную сферы. Наряду с легальными практиками цифровая культура порождает явления противоправного и деструктивного характера, связанные с достижением различных целей: извлечением прибыли, причинением вреда личности, социальной дестабилизацией и др.

Дипфейк как социокультурное явление и разновидность цифрового медиапродукта имеет как общие черты с иными явлениями цифровой среды, так и явную специфику. Способность к глобальному распространению, открытость для передачи, доступность восприятия широкой аудиторией, синтетический характер и игровая форма — общие черты, сближающие дипфейк с другими феноменами цифровой культуры. При этом следует подчеркнуть уникальность дипфейка, обусловленную тем, что его создание осуществляется во взаимодействии человека и нейросети, способной к самокоррекции. Для совершенствования изображений и устранения неточностей в копировании оригинала в программную систему загружается исходный материал и запускается *процесс* редактирования, в ходе которого генеративная система обнаруживает неточности, а затем устраняет их на основе самообучения, постепенно повышая реалистичность изображения.

Другим важным отличием является высокая реалистичность искусственно сгенерированного контента, который намеренно выдается за реальный. В связи с этим возникает новый онтологический аспект проблемы, связанный со стиранием границы между реальным и нереальным. Иные продукты цифровой культуры, прежде всего симуляторы и видеоигры, также отличаются реалистичностью и потенциально могут способствовать стиранию этой границы. В случае дипфейк-технологии проблема затрагивает онтологический статус самого объекта: с одной стороны, он не тождествен тому, чем предстает перед аудиторией, а с другой — формирует у нее иска-

женное представление о реальности. Это не обезличивание или омассовление, ставшие традиционными угрозами Я в условиях цифровой среды, а новая форма разрыва сущности и явления, индивида и его облика в окружающей среде. Трансформация Я в иной, искусственно сконструированный образ изменяет и сами условия восприятия, «производя инаковость как управляемый, эстетизированный и коммерциализированный поток» (Aliev, 2005, p. 167).

В условиях цифровизации проблема множественной идентичности уже сравнительно широко изучена, однако с появлением дипфейков возникают ее новые аспекты. Жертвы дипфейков оказываются оторванными от своей группы, если их представляют как лиц, предавших ее интересы (что характерно для политических манипуляций). Их идентичность становится принципиально неопределенной, поскольку аудитория не способна установить границу между реальностью и ее медийным искажением.

С этих позиций актуализируется проблема уникальности личности и ее трансцендирования во внешний мир. Дипфейк с фальсифицированным изображением конкретного человека, выдаваемым за реальное, нарушает не только частную жизнь изображенного субъекта, но и его экзистенциальные параметры — индивидуальность, свободу, форму коммуникации с Другим, проявление самости в мире. Личность оказывается лишенной надежной защиты от копирования и искажения собственного образа, что требует совершенствования законодательства и формирования нового уровня защиты прав личности, ее биологических и социальных характеристик.

Дипфейк по характеру воздействия в определенном смысле может быть сопоставлен с технологией клонирования, которая по этическим и социальным причинам была подвергнута ограничению (в форме моратория на проведение экспериментов на человеке), поскольку могла бы вызвать хаос и создать глобальную угрозу человечеству. Дипфейк способен вызвать аналогичные эффекты, притом, что его производство не требует колоссальных затрат и профессиональных знаний и доступно практически любому пользователю нейросетей. Соответственно, возможные последствия и риски могут быть значительно более масштабными.

В гносеологическом плане дипфейк может быть отнесен к квазиобъектам, функция которых состоит в подмене реального ложным, то есть фальсифицированным. Получение знаний об объекте в этих условиях неизбежно сталкивается с дилеммой истинности или ложности эмпирического восприятия, которое во многом рассматривалось как критерий истинности знания. Даже если аргументы сторон опираются на очевидные факты, которые реципиент получает из цифровой среды, с широким применением дипфейков они перестают быть убедительными и, более того, могут вводить в опасное заблуждение. Это порождает недоверие, с одной стороны, к СМИ, социальным сетям и нейросетям, а с другой — к субъектам, которые становятся жертвами

дипфейков и подвергаются копированию. С появлением подобных технологий доверие, репутация, честь и достоинство оказываются под возрастающей угрозой; одновременно усиливается скептическое, недоверчивое и отчужденное отношение аудитории к информационным потокам.

В эстетическом отношении дипфейк может рассматриваться как результат творческой деятельности и сложный продукт высоких технологий, содержание которого при злонамеренном применении оказывается тем более разрушительным, чем совершеннее его форма. Низменные цели создания объектов, оскорбляющих достоинство личности, превращают даже технически совершенные медиапродукты в источники зла и ненависти.

Цифровые продукты и медиапродукты амбивалентны по своей сути: сами по себе они не являются ни ценностью, ни деструктивной силой, однако играют ключевую роль в том, какую реальность и каким образом воспринимает адресат коммуникации. С позиции теории коммуникации дипфейк как продукт «глубокого обучения» выступает новым участником коммуникативных связей. В классической модели коммуникации выделяются адресант, сообщение и адресат; М. Маклюэн акцентировал также определяющую роль медиасреды. Н. Луман, разделяя тезис об определяющей роли массмедиа, уточняет их функции с позиции конструктивизма, показывая, что они удваивают реальность, передавая ее особым символическим образом. Конструирование реальности с помощью медиа коренным образом меняется с появлением ИИ. В случае создания дипфейков речь идет не о воссоздании реальности или ее конструировании как определенной интерпретации, а о ее намеренном ложном конструировании. В этом отношении в структуре коммуникации появляется новый агент: помимо адресанта, адресата, сообщения — собственно видео- или аудиодипфейка — и цифровой среды, посредством которой осуществляется передача, возникает генератор информации (нейросеть на основе глубокого обучения). Его роль состоит не просто в удвоении реальности, а в создании цифровой антиреальности, которая будет существовать наряду с другими формами, притом, что ее воздействие на аудиторию в силу «вирусного эффекта» может стать доминирующим. Как отмечал Н. Луман, «массмедиа воздействуют на формирование социально-избыточных коммуникаций в рамках всего общества» (2005, с. 38). Поэтому массмедиа распространяют информацию максимально широко, доводя ее до каждого. Луман указывал, что «массмедиа поддерживают общество в состоянии бодрости», порождают в нем постоянную готовность считаться с возможными неожиданностями и даже помехами (2005, с. 40–41). Если рассматривать дипфейки с этих позиций, то они могут выступать своеобразным раздражителем и провокатором, заставляющим аудиторию критически анализировать воспринимаемый в сети контент, ожидать обмана или подделки и проявлять максимальную наблюдательность. В этом смысле дипфейки действительно поддер-

живают аудиторию в состоянии «бодрости» и стимулируют поиск достоверной информации.

Социальное конструирование с помощью технологии дипфейков приобретает беспрецедентный масштаб. На смысловую и ценностную среду современного субъекта теперь воздействуют не только тексты и «дешевые подделки» — примитивные видео с участием публичных персон, создававшиеся до появления нейросетей, — но и высококачественные видео- и аудиосообщения, сгенерированные ИИ от имени наиболее известных лиц. Передаваемая информация, даже несмотря на то, что дипфейк в дальнейшем может быть раскрыт, оказывает «вирусное» воздействие, встраивается в память, создает ложные воспоминания и неосознанно усваивается адресатом. В этом отношении дипфейк оказывает многоуровневое воздействие, сочетая эффект мгновенной «информационной бомбы» с «отложенным воздействием» на сознание и мотивацию индивида.

Выводы

Традиционно технологии, как и знания, воспринимались как ценностно-нейтральные, находящиеся вне сферы этической оценки. Однако с появлением оружия массового поражения, в том числе химического и биологического, к анализу технологий стали применять принцип «не навреди», поскольку они обладают потенциально высокой разрушительной силой и вследствие этого нуждаются в этической и гуманитарной экспертизе.

Информационно-коммуникационные продукты с появлением генеративного ИИ выходят на качественно новый уровень развития. Их воздействие на жизнь человека и общества становится все более масштабным и во многих отношениях определяющим. Феномены цифровой культуры, *порожденные* человеком, дополняются явлениями, созданными искусственными самообучающимися системами, что приводит к смешению различных уровней реальности, при котором информация удваивается, копируется, искажается и конструируется в интересах тех или иных социальных агентов. Технология дипфейков становится еще одной формой конструирования реальности, рискогенный потенциал которой обусловлен тем, что создаваемая с ее помощью реальность выдается за подлинную и тем самым ставит под сомнение существование и статус первичной реальности. Дипфейк по своей сути — это «отмена» подлинности. Дипфейк как коммуникативный феномен, предполагающий генерацию встроенного в акт коммуникации ложного контекста или сообщения, оказывает двойной негативный эффект: на субъекта, чей образ используется, и на адресата, причем оба становятся жертвами обмана. Он разрушает экзистенциальное пространство объекта дипфейка, нарушая его право на защиту чести и достоинства от посягательств, с одной стороны, и создает ложные смыслы и знания для наблюдателя, то есть аудитории, ориентируя их мышление, память и поведение

по той траектории, которая задана генератором этого продукта. Несмотря на возможные позитивные формы применения технологии дипфейка, ее деструктивный потенциал оказался значительно выше, а соответствующие практики получили более широкое распространение. Регулирование этого процесса с помощью законодательства, хотя и необходимо, уже не способно полностью остановить распространение дипфейков. Технологические меры, в частности создание фильтров и детекторов видеофейков, могут ограничить этот процесс. Подобно тому как антивирусные системы были внедрены в каждый гаджет, в ближайшем будущем должны появиться приложения для проверки получаемой информации на наличие признаков генеративной обработки. В нормативном отношении сгенерированные продукты нуждаются в обязательной маркировке, указывающей на искусственный характер изображения или фотографии. Отсутствие такой маркировки может рассматриваться по аналогии с подделкой документов и материальных ценностей и влечь усиленные санкции с учетом причиненного жертве морального вреда. В противном случае индивид, получающий информацию в цифровой среде, неизбежно будет оставаться легкой мишенью, каким бы уровнем защиты личного пространства и какими бы знаниями он ни обладал.

Заявления автора

Согласие с политиками и положениями сетевого издания

Автор данной рукописи был уведомлен и согласен с действующими политиками и положениями сетевого издания, изложенными на веб-сайте.

Оригинальность и эксклюзивность рукописи:

Автор подтверждает, что рукопись является оригинальной, ранее не публиковалась ни полностью, ни частично и не находится на рассмотрении для публикации в других изданиях.

Авторство и вклад (CRediT)

Концептуализация, проведение исследование и написание текста статьи — Баева Л. В.

Политика в отношении препринтов и предыдущих версий

Автор подтверждает, что нет препринта / рабочей версии

Конфликт интересов

Конфликтов нет

Финансирование

Данная работа не профинансирована ни одним спонсором.

Использование генеративного ИИ и вспомогательных инструментов

Инструменты искусственного интеллекта не указаны в качестве авторов; любой вклад (редактирование текста, перевод, иллюстрации, код) описан в разделе «Методы»; точность, оригинальность и соблюдение прав проверены.

Разрешения на использование (тексты, изображения, карты, иллюстрации)

Все материалы, защищённые авторским правом, воспроизводятся с разрешения правообладателя; соблюдаются условия лицензии (включая Creative Commons).

Благодарности и вклады без авторства

Автор благодарит редакционную коллегию журнала за возможность представления статей по междисциплинарной тематике в области медиаисследований и вклад в развитие этой тематической области.

Список литературы

Ahmed, S., & Chua, H. W. (2023). Perception and Deception: Exploring Individual Responses to Deepfakes Across Different Modalities. *Heliyon*, 9(10), e20383. https://doi.org/10.1016/j.heliyon.2023.e20383

AI Standards for Global Impact: From Governance to Action. (2025). ITU Telecommunication Standardization Bureau. https://www.itu.int/dms_pub/itu-t/opb/ai4g/T-AI4G-AI4GOOD-2025-4-PDF-E.pdf

Aliev, R. T. (2025). Post-Symbolic Coloniality: The Metasimulacrum of Otherness in Digital Culture. *Galactica Media: Journal of Media Studies*, 7(2), 166–186. https://doi.org/10.46539/gmd.v7i2.649

An Open Letter. Research priorities for robust and beneficial artificial intelligence. (2015). Future of Life Institute. https://futureoflife.org/2015/10/27/ai-open-letter/

Bainotti, L. (2022). Defines and Discusses Deepfakes. In *Visual Media Analysis for Online Discourses: The Case of Deepfakes*. SAGE Publications, Ltd. https://doi.org/10.4135/9781529608380.n2

Barari, S., Lucas, C., & Munger, K. (2025). Political Deepfakes Are as Credible as Other Fake Media and (Sometimes) Real Media. *The Journal of Politics*, 87(2), 510–526. https://doi.org/10.1086/732990

Bermudez Bermudez, Y. A. (2025). Deepfakes and Artificial Intelligence in Social Engineering: Emerging Threats in 21st-Century Cyberfraud. *IUSTA*, (63), 54–71. https://doi.org/10.15332/25005286.11068

Blanchet, A. (2025, February 6). *Hollywood's stand against AI: a blueprint for collective bargaining in the digital age*. https://www.equaltimes.org/hollywood-s-stand-against-ai-a?lang=en

Deepfakes in the Financial Sector: Understanding the Threats, Managing the Risks. (2024). FS-ISAC. https://www.fsisac.com/hubfs/Knowledge/AI/DeepfakesInTheFinancialSector-UnderstandingTheThreatsManagingTheRisks.pdf

Dobber, T., Metoui, N., Trilling, D., Helberger, N., & De Vreese, C. (2021). Do (Microtargeted) Deepfakes Have Real Effects on Political Attitudes? *The International Journal of Press/Politics*, 26(1), 69–91. https://doi.org/10.1177/1940161220944364

Fallis, D. (2021). The Epistemic Threat of Deepfakes. *Philosophy & Technology*, 34(4), 623–643. https://doi.org/10.1007/s13347-020-00419-2

Floridi, L. (2018). Artificial Intelligence, Deepfakes and a Future of Ectypes. *Philosophy & Technology*, 31(3), 317–321. https://doi.org/10.1007/s13347-018-0325-3

Fox, J., & Bailenson, J. N. (2009). Virtual Self-Modeling: The Effects of Vicarious Reinforcement and Identification on Exercise Behaviors. *Media Psychology*, 12(1), 1–25. https://doi.org/10.1080/15213260802669474

Hancock, J. T., & Bailenson, J. N. (2021). The Social Impact of Deepfakes. *Cyberpsychology, Behavior, and Social Networking*, 24(3), 149–152. https://doi.org/10.1089/cyber.2021.29208.jth

Ilyas, M., Hashmat, D. S., & Bibi, F. (2025). The Ethics of Deepfake Technology in Media: A Review Study. *Journal of Media Horizons*, 6(3), 385–398.

Levine, T. R. (2019). *Duped: Truth-Default Theory and the Social Science of Lying and Deception*. University of Alabama Press. https://doi.org/10.2307/jj.30297464

Li, S. (2025). The Social Harms of AI-Generated Fake News: Addressing Deepfake and AI Political Manipulation. *Digital Society & Virtual Governance*, 1(1), 72–88. https://doi.org/10.6914/dsvg.010105

Mishra, S., Shukla, P. K., & Agrawal, R. (2022). DeepFakes, Media, and Societal Impacts. In L. Gaur, *DeepFakes* (1st ed., pp. 115–120). CRC Press. https://doi.org/10.1201/9781003231493-9

Nascimento, N. L. (2024). *The Role of Critical Thinking in Combating Deepfakes in Politics and Misinformation*. SocArXiv. https://doi.org/10.31235/osf.io/cxjaq

Nightingale, S. J., & Farid, H. (2022). AI-Synthesized Faces Are Indistinguishable from Real Faces and More Trustworthy. *Proceedings of the National Academy of Sciences*, 119(8), e2120481119. https://doi.org/10.1073/pnas.2120481119

Pause Giant AI Experiments: An Open Letter. (2024). Future of Life Institute. https://futureoflife.org/open-letter/pause-giant-ai-experiments/

Reddy, S. (2025). *Stars Strike Back at Deepfakes*. Times of India. https://timesofindia.indiatimes.com/india/stars-strike-back-at-deepfakes/articleshow/125982843.cms

Segovia, K. Y., & Bailenson, J. N. (2009). Virtually True: Children's Acquisition of False Memories in Virtual Reality. *Media Psychology*, 12(4), 371–393. https://doi.org/10.1080/15213260903287267

Sutrop, M. (2020). Challenges of Aligning Artificial Intelligence with Human Values. *Acta Baltica Historiae et Philosophiae Scientiarum*, 8(2), 54–72. https://doi.org/10.11590/abhps.2020.2.04

Архиреев, Н. В. (2025). Правовое регулирование создания и использования дипфейков. *Правовое государство: теория и практика*, 21(2), 22–30. https://doi.org/10.33184/pravgos-2025.2.3

Бодров, Н. Ф., & Лебедева, А. К. (2023). Понятие дипфейка в российском праве, классификация дипфейков и вопросы их правового регулирования. *Юридические исследования*, (11), 26-41. https://doi.org/10.25136/2409-7136.2023.11.69014

Булгарова, Б. А., Уллах Мд, Т., Мондал, П., & Ма, Ф. (2025). От зарождения к этическим границам: Двойное влияние искусственного интеллекта и технологии Deepfeke на кинематограф (2017-2024). *Litera*, (3), 244–256. https://doi.org/10.25136/2409-8698.2025.3.73652

Виноградов, В. А., & Кузнецова, Д. В. (2024). Зарубежный опыт правового регулирования технологии «дипфейк». *Право. Журнал Высшей школы экономики*, 17(2), 215–240. https://doi.org/10.17323/2072-8166.2024.2.215.240

Воронин, И. А., & Гавра, Д. П. (2024). Дипфейки: Современное понимание, подходы к определению, характеристики, проблемы и перспективы. *Российская школа связей с общественностью*, (33), 28–47.

Добробаба, М. Б. (2022). Дипфейки как угроза правам человека. *Lex Russica*, 75(11), 112–119. https://doi.org/10.17803/1729-5920.2022.192.11.112-119

Ефремова, М. А., & Русскевич, Е. А. (2024). Дипфейк (deepfake) и уголовный закон. *Вестник Казанского Юридического Института МВД России*, 15(2), 97–105. https://doi.org/10.37973/VESTNIKKUI-2024-56-13

Иванов, В. Г., & Игнатовский, Я. Р. (2020). Deepfakes: Перспективы применения в политике и угрозы для личности и национальной безопасности. *Вестник Российского университета дружбы народов. Серия: Государственное и муниципальное управление*, 7(4), 379-386. https://doi.org/10.22363/2312-8313-2020-7-4-379-386

Калятин, В. О. (2022). Дипфейк как правовая проблема: Новые угрозы или новые возможности? *Закон*, (7), 87–103. https://doi.org/10.37239/0869-4400-2022-19-7-87-103

Киселёв, А. С. (2021). О необходимости правового регулирования в сфере искусственного интеллекта: Дипфейк как угроза национальной безопасности. *Московский юридический журнал*, (3), 54–64. https://doi.org/10.18384/2310-6794-2021-3-54-64

Крымова, Ю. (2025). *Дипфейки с главами регионов заполонили сеть. Кто из чиновников уже стал жертвой мошенников и как распознать обман*. RGRU. https://rg.ru/2025/10/08/reg-dfo/vam-zvonit-mer.html

Лукина, Ю. В. (2023). Использование дипфейков в общественно-политической жизни. *Русская политология*, 2(27), 41–48.

Луман, Н. (2005). *Реальность массмедиа*. Праксис.

Малышева, Ю. Ю. (2025). Мошеннические проявления фейков и дипфейков: Проблемы противодействия. *Российско-азиатский правовой журнал*, (1), 105–109. https://doi.org/10.14258/ralj(2025)1.16

Свиридова, Е. А. (2024). Правила использования технологий дипфейк в праве США и КНР: адаптация зарубежного опыта правового регулирования. *Современное право*, (3), 119–123. https://doi.org/10.25799/NI.2024.96.67.019

Фалалеев, М. А., Ситдикова, Н. А., & Нечай, Е. Е. (2021). Дипфейк как феномен политической коммуникации. *Вестник Забайкальского государственного университета*, 27(6), 101–106. https://doi.org/10.21209/2227-9245-2021-27-6-101-106

Шубенкова, К. В. (2022). Правовые проблемы охраны цифрового образа личности в РФ. *Гуманитарные, социально-экономические и общественные науки*, (10), 226–230. https://doi.org/10.23672/r5500-3607-6625-k

References

Ahmed, S., & Chua, H. W. (2023). Perception and Deception: Exploring Individual Responses to Deepfakes Across Different Modalities. *Heliyon*, 9(10), e20383. https://doi.org/10.1016/j.heliyon.2023.e20383

AI Standards for Global Impact: From Governance to Action. (2025). ITU Telecommunication Standardization Bureau. https://www.itu.int/dms_pub/itu-t/opb/ai4g/T-AI4G-AI4GOOD-2025-4-PDF-E.pdf

Aliev, R. T. (2025). Post-Symbolic Coloniality: The Metasimulacrum of Otherness in Digital Culture. *Galactica Media: Journal of Media Studies*, 7(2), 166–186. https://doi.org/10.46539/gmd.v7i2.649

An Open Letter. Research priorities for robust and beneficial artificial intelligence. (2015). Future of Life Institute. https://futureoflife.org/2015/10/27/ai-open-letter/

Arkhiereev, N. V. (2025). Legal Regulation of Creating and Using Deepfakes. *The Rule-of-Law State Theory and Practice*, 21(2), 22–30. https://doi.org/10.33184/pravgos-2025.2.3 (In Russian).

Bainotti, L. (2022). Defines and Discusses Deepfakes. In *Visual Media Analysis for Online Discourses: The Case of Deepfakes*. SAGE Publications, Ltd. https://doi.org/10.4135/9781529608380.n2

Barari, S., Lucas, C., & Munger, K. (2025). Political Deepfakes Are as Credible as Other Fake Media and (Sometimes) Real Media. *The Journal of Politics*, 87(2), 510–526. https://doi.org/10.1086/732990

Bermudez Bermudez, Y. A. (2025). Deepfakes and Artificial Intelligence in Social Engineering: Emerging Threats in 21st-Century Cyberfraud. *IUSTA*, (63), 54–71. https://doi.org/10.15332/25005286.11068

Blanchet, A. (2025, February 6). *Hollywood's stand against AI: a blueprint for collective bargaining in the digital age*. https://www.equaltimes.org/hollywood-s-stand-against-ai-a?lang=en

Bodrov, N. F., & Lebedeva, A. K. (2023). The Concept of Deepfake in Russian Law, Classification of Deepfake and Issues of Their Legal Regulation. *Legal Studies*, (11), 26–41. https://doi.org/10.25136/2409-7136.2023.11.69014 (In Russian).

Bulgarova, B. A., Ullah Md, T., Mondal, P., & Ma, F. (2025). From Emergence to Ethical Frontiers: AI and Deepfake Technology's Dual Impact on Cinema (2017–2024). *Litera*, (3), 244–256. https://doi.org/10.25136/2409-8698.2025.3.73652 (In Russian).

Deepfakes in the Financial Sector: Understanding the Threats, Managing the Risks. (2024). FS-ISAC. https://www.fsisac.com/hubfs/Knowledge/AI/DeepfakesInTheFinancialSector-UnderstandingTheThreatsManagingTheRisks.pdf

Dobber, T., Metoui, N., Trilling, D., Helberger, N., & De Vreese, C. (2021). Do (Microtargeted) Deepfakes Have Real Effects on Political Attitudes? *The International Journal of Press/Politics*, 26(1), 69–91. https://doi.org/10.1177/1940161220944364

Dobrobaba, M. B. (2022). Deepfakes as a Threat to Human Rights. *Lex Russica*, 75(11), 112–119. https://doi.org/10.17803/1729-5920.2022.192.11.112-119 (In Russian).

Efremova, M., & Russkevich, E. (2024). Deepfake And Criminal Law. *Bulletin of the Kazan Law Institute of MIA Russia*, 15(2), 97–105. https://doi.org/10.37973/VESTNIKKUI-2024-56-13 (In Russian).

Falaleev, M., Sitdikova, N., & Nechay, E. (2021). Deepfake as a Phenomenon of Political Communication. *Transbaikal State University Journal*, 27(6), 101–106. https://doi.org/10.21209/2227-9245-2021-27-6-101-106 (In Russian).

Fallis, D. (2021). The Epistemic Threat of Deepfakes. *Philosophy & Technology*, 34(4), 623–643. https://doi.org/10.1007/s13347-020-00419-2

Floridi, L. (2018). Artificial Intelligence, Deepfakes and a Future of Ectypes. *Philosophy & Technology*, 31(3), 317–321. https://doi.org/10.1007/s13347-018-0325-3

Fox, J., & Bailenson, J. N. (2009). Virtual Self-Modeling: The Effects of Vicarious Reinforcement and Identification on Exercise Behaviors. *Media Psychology*, 12(1), 1–25. https://doi.org/10.1080/15213260802669474

Hancock, J. T., & Bailenson, J. N. (2021). The Social Impact of Deepfakes. *Cyberpsychology, Behavior, and Social Networking*, 24(3), 149–152. https://doi.org/10.1089/cyber.2021.29208.jth

Ilyas, M., Hashmat, D. S., & Bibi, F. (2025). The Ethics of Deepfake Technology in Media: A Review Study. *Journal of Media Horizons*, 6(3), 385–398.

Ivanov, V. G., & Ignatovskiy, Y. R. (2020). Deepfakes: Prospects for Political Use and Threats to the Individual and National Security. *RUDN Journal of Public Administration*, 7(4), 379–386. https://doi.org/10.22363/2312-8313-2020-7-4-379-386 (In Russian).

Kalyatin, V. O. (2022). Deepfake as a Legal Problem: New Threats or New Opportunities? *Zakon*, (7), 87–103. https://doi.org/10.37239/0869-4400-2022-19-7-87-103 (In Russian).

Kiselev, A. S. (2021). On the Expansion of Legal Regulation in the Field of Artificial Intelligence: Deepfake as a Threat to National Security. *Bulletin of the Moscow State Regional University (Jurisprudence)*, (3), 54–64. https://doi.org/10.18384/2310-6794-2021-3-54-64 (In Russian).

Krymova, Yu. (2025). *Deepfakes Featuring Regional Leaders Have Flooded the Internet. Which Officials Have Already Fallen Victim to Scammers, and How to Spot the Deception*. RGRU. https://rg.ru/2025/10/08/reg-dfo/vam-zvonit-mer.html (In Russian).

Levine, T. R. (2019). *Duped: Truth-Default Theory and the Social Science of Lying and Deception*. University of Alabama Press. https://doi.org/10.2307/jj.30297464

Li, S. (2025). The Social Harms of AI-Generated Fake News: Addressing Deepfake and AI Political Manipulation. *Digital Society & Virtual Governance*, 1(1), 72–88. https://doi.org/10.6914/dsvg.010105

Luhmann, N. (2005). *The reality of the mass media*. Praksis. (In Russian)

Lukina, Yu. V. (2023). Use of Deepfakes in Public and Political Life. *Russian Political Science*, 2(27), 41–48. (In Russian).

Malysheva, Yu. Yu. (2025). Fraudulent manifestations of fakes and deepfakes: Problems of counteraction. *Russian-Asian Law Journal*, (1), 105–109. https://doi.org/10.14258/ralj(2025)1.16 (In Russian)

Mishra, S., Shukla, P. K., & Agrawal, R. (2022). DeepFakes, Media, and Societal Impacts. In L. Gaur, *DeepFakes* (1st ed., pp. 115–120). CRC Press. https://doi.org/10.1201/9781003231493-9

Nascimento, N. L. (2024). *The Role of Critical Thinking in Combating Deepfakes in Politics and Misinformation*. SocArXiv. https://doi.org/10.31235/osf.io/cxjaq

Nightingale, S. J., & Farid, H. (2022). AI-Synthesized Faces Are Indistinguishable from Real Faces and More Trustworthy. *Proceedings of the National Academy of Sciences*, 119(8), e2120481119. https://doi.org/10.1073/pnas.2120481119

Pause Giant AI Experiments: An Open Letter. (2024). Future of Life Institute. https://futureoflife.org/open-letter/pause-giant-ai-experiments/

Reddy, S. (2025). *Stars Strike Back at Deepfakes*. Times of India. https://timesofindia.indiatimes.com/india/stars-strike-back-at-deepfakes/articleshow/125982843.cms

Segovia, K. Y., & Bailenson, J. N. (2009). Virtually True: Children's Acquisition of False Memories in Virtual Reality. *Media Psychology*, 12(4), 371–393. https://doi.org/10.1080/15213260903287267

Shubenkova, K. V. (2022). Legal problems of protecting the digital image of the individual in the Russian Federation. *Humanities, Social-Economic and Social Sciences*, (10), 226–230. https://doi.org/10.23672/r5500-3607-6625-k (In Russian).

Sutrop, M. (2020). Challenges of Aligning Artificial Intelligence with Human Values. *Acta Baltica Historiae et Philosophiae Scientiarum*, 8(2), 54–72. https://doi.org/10.11590/abhps.2020.2.04

Sviridova, E. A. (2024). Rules for the use of deepfake technologies in the law of the USA and the People's Republic of China: Adaptation of foreign experience in legal regulation. *Modern Law*, (3), 119–123. https://doi.org/10.25799/NI.2024.96.67.019 (In Russian)

Vinogradov, V. A., & Kuznetsova, D. V. (2024). Foreign Experience in Legal Regulating Deepfake Technology. *Law Journal of the Higher School of Economics*, (2), 215–240. https://doi.org/10.17323/2072-8166.2024.2.215.240 (In Russian).

Voronin, I. A., & Gavra, D. P. (2024). Deepfakes: Modern Understanding, Approaches to Definition, Characteristics, Problems and Prospects. *Russian School of Public Relations*, (33), 28–47. (In Russian).