Известия Саратовского университета. Новая серия. Серия: Социология. Политология. 2025. Т. 25, вып. 2. С. 195–201
Izvestiya of Saratov University. Sociology. Politology, 2025, vol. 25, iss. 2, pp. 195–201
https://soziopolit.sgu.ru https://doi.org/10.18500/1818-9601-2025-25-2-195-201, EDN: TTKKLN

Научная статья
УДК 32.019.51

Дипфейк как инновационный инструмент политического манипулирования в начале XXI века

К. В. Старостенко, А. С. Коновалов ✉

Орловский государственный университет имени И. С. Тургенева, Россия, 302026, г. Орел, ул. Комсомольская, д. 95

Старостенко Константин Викторович, доктор политических наук, профессор, заведующий кафедрой общей и прикладной политологии, pilotskv@bk.ru, https://orcid.org/0000-0003-1850-8606, AuthorID: 647263

Коновалов Андрей Сергеевич, аспирант кафедры общей и прикладной политологии, andrei070199@mail.ru, https://orcid.org/0009-0000-5219-7982, AuthorID: 1264608

Аннотация. Статья посвящена анализу такого феномена, как дипфейк, в контексте его применения в политической сфере с точки зрения манипуляции обществом. В первую очередь, авторы раскрывают сущность явления дезинформации, проводя краткую историческую ретроспективу, затем дают определение понятию «дипфейк» и объясняют принцип работы данной технологии как с технической, так и с когнитивной точек зрения. В ходе анализа некоторых наглядных кейсов, предлагаются наиболее доступные на сегодняшний день для среднестатистического пользователя способы обнаружения дипфейков. В конечном итоге делается вывод о том, что из-за стремительного технического прогресса и, как следствие, увеличения сгенерированного контента перспектива информационного пространства видится весьма удручающей. Это связано с тем, что искусственно созданный материал с каждым днём становится всё неотличимее от подлинного, и в обозримом будущем человек неизбежно столкнётся с невозможностью обнаружения подделки без специальных технических средств, что может в корне подорвать доверие граждан даже к официальным новостным источникам.
Ключевые слова: политика, дипфейк, искусственный интеллект, нейросеть, информация, дезинформация, пропаганда

Для цитирования: *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. https://doi.org/10.18500/1818-9601-2025-25-2-195-201, EDN: TTKKLN

Article

Deepfake as an innovative tool of political manipulation in the early XXI century

K. V. Starostenko, A. S. Konovalov ✉

Orel State University named after I. S. Turgenev, 95 Komsomolskaya St., Orel 302026, Russia

Konstantin V. Starostenko, pilotskv@bk.ru, https://orcid.org/0000-0003-1850-8606, AuthorID: 647263

Andrey S. Konovalov, andrei070199@mail.ru, https://orcid.org/0009-0000-5219-7982, AuthorID: 1264608

Abstract. The article is devoted to the analysis of such phenomenon as deepfake in the context of its application in the political sphere from the point of view of manipulation of society. First of all, the authors reveal the essence of the phenomenon of disinformation by conducting a brief historical retrospective. Then they define the concept of "deepfake" and explain how this technology works from both technical and cognitive points of view. In the course of analyzing some illustrative cases, they propose the most affordable ways for the average user to detect deepfakes today. In the end, they conclude that due to the rapid technical progress and the resulting increase in generated content, the outlook for the information space is quite bleak. This is due to the fact that every day artificially created material becomes more and more indistinguishable from the genuine, and in the foreseeable future, a person will inevitably face the impossibility of detecting the fake without special technical means, which can fundamentally undermine the trust of citizens even to official news sources.
Keywords: politics, deepfake, artificial intelligence, neural network, information, disinformation, propaganda

For citation: Starostenko K. V., Konovalov A. S. Deepfake as an innovative tool of political manipulation in the early XXI century. *Izvestiya of Saratov University. Sociology. Politology,* 2025, vol. 25, iss. 2, pp. 195–201 (in Russian). https://doi.org/10.18500/1818-9601-2025-25-2-195-201, EDN: TTKKLN

Информация всегда была важнейшим фактором, оказывающим непосредственное влияние на социальные, экономические и, что самое главное, политические процессы. Однако в нынешнюю цифровую эпоху, когда одним из приоритетнейших научных направлений является исследование искусственного интеллекта и, как следствие, технологий глубокого обучения, появились новые способы создания и распространения ложных сообщений. Особое место среди них занимают дипфейки. Но прежде чем перейти к исследованию этого феномена, необходимо обратиться к истории, дабы прояснить сущность такого явления, как дезинформация.

Дезинформация представляет собой сознательное и преднамеренное распространение фальшивой или искажённой информации с целью ввести аудиторию в заблуждение и манипулировать её восприятием. В отличие от случайных ошибок или недостоверных данных, дезинформация всегда предполагает определённую цель – будь то политическое влияние, социальная дестабилизация или извлечение экономической выгоды. История этого явления уходит корнями в глубь веков, и самым известным по сей день произведением, посвящённым военной стратегии, безусловно, остаётся трактат «Искусство войны» (V в. до н.э.), где обман рассматривается как один из ключевых инструментов достижения победы. Тем не менее приёмы дезинформации использовались ещё задолго до этого; в своей работе Сунь-цзы лишь систематизировал и углубил уже существующие военные практики, такие как, например, фальшивые отступления или искажение данных о численности войск и их местоположении [1].

Продолжая историческую ретроспективу, стоит отметить, что Средневековье и Ренессанс также ознаменовались активным использованием и усовершенствованием методов дезинформации, в значительной степени благодаря религиозным институтам. В то время церковь обладала монополией на знания, что позволяло ей контролировать распространение различного рода сведений и ограничивать доступ к альтернативным взглядам. В политической же сфере широко применялись слухи и интриги, направленные на дискредитацию оппонентов и укрепление власти. В эпоху Просвещения, с развитием печатных изданий, методы стали ещё более разнообразными; так, сатирические памфлеты и панегирики служили в качестве усиления манипуляции общественным мнением.

Эпоха мировых войн наглядно продемонстрировала, какую важную роль может играть информация в руках крупных политических акторов. В этот период как отдельные государства, так и целые блоки государств активно использовали дезинформацию в контексте пропаганды для поддержания морального духа населения, деморализации врага, а также для создания ложных представлений о событиях, направляя общественное мнение в нужное русло. Так, например, в годы Второй мировой войны плакаты, радиопередачи и кинофильмы стали ключевыми средствами донесения официальной позиции правительства до народа. Холодная война, с её глобальным противостоянием Востока и Запада, охарактеризовалась ещё более мощной информационной конфронтацией, в ходе которой операции по дезинформации стали неотъемлемой частью политической стратегии великих держав. В свою очередь, радио и телевидение, ставшие массово доступными, позволили этим операциям выйти на качественно новый уровень взаимодействия с аудиторией и достичь невиданного ранее эффекта.

Анализируя в этом направлении специфику начала нового тысячелетия, можно отметить, что современные методы манипуляции благодаря наличию глобальной сети Интернет характеризуются широкомасштабным и скоростным распространением различного контента, а также трудностями его контроля из-за высокой степени анонимности источников в условиях децентрализованных информационных потоков. В этой связи одним из наиболее значимых инструментов современной дезинформации являются дипфейки. Этот термин происходит от английского «deepfake», что представляет собой сочетание слов «deep» (глубокий) и «fake» (подделка). Сама этимология слова указывает на его происхождение вследствие широкого распространения технологий глубокого обучения, которые используются для создания поддельного, но чрезвычайно реалистичного контента.

Таким образом, дипфейки можно определить как синтетические медиапродукты, создаваемые с использованием технологий искусственного интеллекта, в частности глубокого обучения, которые позволяют генерировать объекты в формате текста, изображения, аудио и видео. Этот феномен приобрел широкую популярность именно благодаря своей высокой степени правдоподобности, настолько высокой, что даже самые опытные специалисты порой не

способны верифицировать его должным образом. Это связано с тем, что основой для создания дипфейков служат нейронные сети[1], которые, в свою очередь, обеспечивают возможность детального анализа поведенческих особенностей реальных людей, что позволяет создавать материал, вполне достоверный с точки зрения восприятия целевой аудиторией. Стоит обратить внимание, что данный процесс осуществляется благодаря генеративно-состязательным сетям (GAN), представляющим собой инновационную архитектуру, организованную вокруг взаимодействия двух основных компонентов: генератора и дискриминатора. Генератор отвечает за создание поддельных объектов, максимально приближенных к действительным, тогда как дискриминатор оценивает их подлинность [2]. Процесс такого конкурентного обучения напоминает механизм эволюции, где посредством многочисленных итераций генерируемый контент приобретает всё большую степень совершенства.

В этой связи довольно существенным аспектом повышения реалистичности дипфейков являются так называемые автоэнкодеры [3] и технологии синтеза речи [4], которые помогают в процессе производства мультимедийных объектов высокой степени визуальной и аудиальной детализации. Автоэнкодер – это вид нейронной сети, используемый для обучения эффективному кодированию и декодированию данных; основная его цель заключается в представлении данных в наиболее компактной форме без существенной потери информации. Это достигается путём обучения сети на задаче восстановления входных данных из их сжатого представления, минимизируя разницу между оригиналом и реконструкцией. Синтез речи, в свою очередь, позволяет моделям воспроизводить голосовые данные, передавая такие тонкие нюансы, как интонация, ритм и тембр голоса. В комбинации с GAN синтез речи ещё больше усиливает эффект реалистичности, позволяя использовать конечный продукт в медиапространстве для имитации публичных выступлений, интервью и даже личных сообщений, что в результате приводит к возможности конструирования полноценных сцен и контекстов, которых в действительности не существовало.

Как того и следовало ожидать, неизбежно на определённом этапе научно-технического прогресса технология создания дипфейков стала доступна практически любому человеку, имеющему в своём распоряжении Интернет. Безусловно, это стало осуществимо благодаря открытым библиотекам, таким как TensorFlow и PyTorch, а также специализированным приложениям и платформам, которые упростили создание сложных моделей даже для пользователей без профессиональных навыков программирования. Массовое распространение приложений для смартфонов также в значительной мере поспособствовало повышению внимания к синтетическому контенту, серьёзно расширив круг потенциальных интересантов данной технологии. Этот процесс привёл к тому, что феномен, прежде считавшийся сугубо научным достижением, стал явным фактором риска для информационной безопасности, обостряя социокультурные и политические угрозы за счёт манипуляций общественным мнением и дестабилизации устоявшихся структур.

Здесь важно отметить, это связано ещё и с тем, что дипфейки эксплуатируют когнитивные слабости человека и играют на особенностях его восприятия, именно поэтому одним из центральных механизмов воздействия на индивида является эмоциональная суггестия (дипфейк зачастую содержит шокирующий или провокационный контент, вызывающий сильные эмоции, такие как страх, гнев или ненависть). Такой приём усиливает эффект погружения и повышает вероятность того, что зритель примет сфальсифицированный материал за подлинный. Это происходит потому, что у людей, находящихся под воздействием сильных переживаний, способность критически анализировать полученную информацию резко снижается, что делает их более восприимчивыми к манипуляциям [5]. В политическом контексте это может привести к кардинальному изменению общественного мнения по отношению к определённым личностям, институтам или событиям, открывая возможности для целенаправленного влияния, к примеру, на электоральные процессы, законодательные инициативы или даже на внешнеполитический курс.

В этом контексте нельзя обойти стороной ещё один важнейший фактор воздействия на аудиторию, коим выступают социальные сети с их масштабом и скоростью распространения контента. Современные международные платформы предоставляют практически идеальные

[1] Нейронная сеть – это модель, основанная на принципах работы человеческого мозга. В отличие от более широкого понятия искусственного интеллекта (ИИ), включающего различные методы и подходы, нейронная сеть представляет собой структуру взаимосвязанных искусственных нейронов, способных обучаться на основе вводных данных.

условия для оперативного обмена информацией, позволяя дипфейкам мгновенно охватывать миллионы пользователей. Вирусный эффект, характерный для соцсетей, способствует достаточно стремительному распространению поддельного контента, особенно если он вызывает сильную эмоциональную реакцию. В свою очередь, использование автоматизированных ботов и псевдоаккаунтов[2] для его массового тиражирования в комментариях также помогает увеличить зону охвата, создавая при этом иллюзию популярности, а вместе с тем достоверности. В результате фальшивые материалы, которые в иных условиях могли бы быть легко нейтрализованы, получают доступ к широким массам до того, как появится авторитетное опровержение. Это создаёт серьёзные препятствия для своевременного противодействия политической дезинформации и приводит к тому, что ложные факты продолжают оказывать влияние на общественное мнение даже после их разоблачения.

Пожалуй, основная сложность раскрытия дипфейков возникает по большей части из-за эффекта первичного восприятия, согласно которому первая полученная информация запоминается лучше и оказывает большее воздействие на реципиента, затрудняя последующее исправление убеждений, сформированных в ходе манипуляции [6]. Даже после предоставления доказательств подделки многие люди продолжают верить в увиденное – это связано с когнитивной предвзятостью и нежеланием пересматривать уже сложившееся мнение. Кроме того, опровержения, как правило, получают меньше внимания и реже вызывают эмоциональный отклик, чем сам дипфейк. Этот фактор особенно важен в контексте политической дезинформации, где даже кратковременное влияние фальшивой информации может иметь долгосрочные последствия. Так, должно быть, самым ярким примером, вызвавшим широкий резонанс, является скандальное видео с Б. Обамой, опубликованное в 2018 г., в котором экс-президент США якобы произносит речь, содержащую оскорбительные высказывания в адрес Д. Трампа; однако на самом деле материал был создан с применением искусственного интеллекта режиссёром Дж. Пилом совместно с изданием BuzzFeed [7]. Этот прецедент наглядно показал, насколько правдоподобным может быть синтетический контент и какие репутационные угрозы он несёт для политических лидеров, особенно в условиях, когда общество в значительной мере полагается на визуальные источники.

К слову, вмешательство в политические выборы стало, пожалуй, одной из основных сфер применения дипфейков в политике. Так, в 2020 г. в Индии во время выборов в Дели лидер оппозиционной партии М. Тивари использовал эту технологию для создания видеообращений на диалекте харьяни, которым он не владеет, с целью привлечения голосов рабочих-мигрантов. Этот случай вызвал широкий общественный и экспертный резонанс, вследствие чего возникли опасения, что такие технологии могут значительно подорвать доверие граждан к самой процедуре выборов[3]. Однако любая медаль имеет две стороны, и реверсивная ситуация произошла в Габоне, где в 2019 г. распространилось видео с президентом Али Бонго Ондимбой, который долгое время не появлялся на публике из-за лечения. Многие заподозрили, что это видео является дипфейком, что вызвало массовые протесты и попытку государственного переворота. Хотя позже было подтверждено, что видео подлинное, таким образом, первоначальное недоверие и распространение ложной информации привели к серьёзной политической нестабильности в стране[4].

Не менее значимым является использование дипфейков в контексте международного мошенничества. Показательный случай произошёл в 2020 г., когда злоумышленники с помощью аудио-дипфейка, имитирующего голос генерального директора немецкой компании, обманом по телефону заставили руководителя британского филиала перевести 220 000 евро на поддельный счет [8]. Этот инцидент наглядно продемонстрировал весь потенциал таких искусственных, но достаточно реалистичных, голосовых записей, которые, выходя за рамки финансовых махинаций, в полной мере могут быть использованы не просто в корыстных, но уже в геополитических целях.

[2] Псевдоаккаунт — это фальшивая учётная запись, созданная в социальной сети или на онлайн-платформе с целью выдачи себя за другого человека, организации либо для сокрытия реальной личности пользователя.

[3] Индийский политик записал предвыборное видео сразу на трех языках. Нет, он не полиглот, просто ему помог искусственный интеллект // RTVI. 2020. URL: https://rtvi.com/news/neural-elections-india/ (дата обращения: 24.11.2024) ; Дипфейки: дезинформация или шаг в будущее? // РБК Тренды. 2021. URL: https://trends.rbc.ru/trends/industry/605dd6979a7947afba6587a7 (дата обращения: 24.11.2024).

[4] Историк рассказал о возможности дипфейков спровоцировать госпереворот // Известия. URL: https://iz.ru/1684196/2024-04-18/istorik-rasskazal-o-vozmozhnosti-dipfeikov-sprovotcirovat-gosperevorot (дата обращения: 24.11.2024).

Стоит отметить, что специальная военная операция на Украине вывела технологии манипуляции с помощью дипфейков в совершенно новое русло. Так, в марте 2022 г. в интернете распространилось поддельное видео, на котором президент В. А. Зеленский якобы призывает украинских военных сложить оружие и сдаться[5] [9]. Это видео было впоследствии разоблачено, однако в первые часы после распространения могло серьёзно подорвать моральный дух среди определённых слоёв украинской армии и населения. В свою очередь, украинские хакеры 5 июня 2023 г. распространили дипфейк от лица Президента России В. В. Путина, где он якобы сообщал по радио и некоторым соцсетям о введении военного положения и о всеобщей мобилизации [10].

В контексте международных провокаций дипфейки также были задействованы и против США. В мае 2023 г. в социальных сетях появилась поддельная фотография, изображающая взрыв на территории Пентагона. Это изображение, созданное с помощью искусственного интеллекта, было опубликовано в ряде популярных источников [11], вследствие чего новость привела к кратковременному снижению рыночных показателей: индекс S&P 500 упал на 30 пунктов, а Dow Jones Industrial Average – на 50 пунктов[6].

Приведённые выше примеры подчёркивают существенную значимость дипфейков как инструмента информационного воздействия с широким спектром применения. Технологическая изощрённость инициаторов контента позволяет создавать материал, способный вызвать серьёзные политические, социальные и экономические последствия. Ввиду этого в условиях растущего числа инцидентов, связанных с дипфейками, становится очевидной необходимость разработки определённых мер по их выявлению и предотвращению. Ниже нами систематизированы и предложены наиболее доступные на сегодняшний день для среднестатистического пользователя способы обнаружения дипфейков (таблица).

Раскроем это наиболее подробно.

[5] Неизвестные опубликовали дипфейк Зеленского с призывом «сложить оружие» // ForkLog. 2022. URL: https://forklog.com/news/ai/neizvestnye-opublikovali-dipfejk-zelenskogo-s-prizyvom-slozhit-oruzhie (дата обращения: 24.11.2024).

[6] Сгенерированное ИИ фото атаки на Пентагон обрушило финансовые рынки // CNews. URL: https://www.cnews.ru/news/top/2023-05-23_sgenerirovannoe_ii_izobrazhenie (дата обращения: 24.11.2024).

Методика анализа материалов на предмет наличия дипфейков

Раздел методики	Ключевые технические аспекты анализа	Техники маскировки дипфейков создателем
1. Анализ текстовых материалов	Лексический и синтаксический анализ Анализ частотности слов и фраз Оценка логической связанности и последовательности	Интеграция сгенерированных данных с реальными Парафразирование
2. Анализ визуальных материалов	Анализ теней и бликов Проверка движений и микроэкспрессий Выявление несоответствий между фоном и передним планом	Легкое размытие Намеренное снижение разрешения и/или качества
3. Анализ аудиоматериалов	Анализ интонации, модуляции и монотонности Проверка на наличие и естественность фоновых шумов	Сглаживание интонаций Добавление искусственного фона Намеренное снижение качества
4. Анализ метаданных и цифровых характеристик	Проверка цифровых водяных знаков Анализ временных меток для выявления аномалий и редактирования	Изменение или удаление метаданных
5. Анализ с применением специализированного программного обеспечения	Использование готового программного обеспечения Использование настраиваемого программного обеспечения	Использование настраиваемого программного обеспечения против детекторов Микроподделки для усложнения детекции
6. Верификация и тестирование материала на подлинность	Фактчекинг и кросс-проверка с независимыми источниками Проверка логики и последовательности	Размещение на платформах с минимальной модерацией Создание ложных отзывов и комментариев для подтверждения правдоподобия

1. Анализ текстовых материалов

Для обнаружения синтетически созданных текстовых материалов применяется лексический и синтаксический анализ на предмет особенностей, характерных для автоматической генерации, таких как повторяющиеся шаблоны или некорректное построение предложений; а также статистический анализ частотности слов и фраз на предмет отклонения от нормы в виде чрезмерного использования определённых конструкций. Важно также оценивать логическую связанность и последовательность аргументации, поскольку тексты, созданные ИИ, могут страдать от несогласованности.

Для маскировки данного вида дипфейков создатели прибегают к интеграции сгенерированных данных с реальными, что затрудняет их обнаружение. Наряду с этим зачастую применяется парафразирование, дабы изменить формулировки и скрыть признаки искусственного происхождения текста.

2. Анализ визуальных материалов

Для анализа визуальных материалов (изображения и видео) большое внимание стоит уделять теням и бликам, поскольку искусственные модели, как правило, испытывают трудности с их корректной реализацией (например, тень неправильно направлена или отсутствует вовсе; блики не соответствуют источнику освещения и явно выделяются). Проверка движений и микроэкспрессий лица также позволяет распознать подделку, так как дипфейки часто не способны точно воспроизводить рефлекторные движения (моргание, малейшие изменения в мимике). Несоответствие между фоном и передним планом является важным индикатором, выдающим истинное происхождение контента (движение на заднем плане может не совпадать с передним, особенно при изменении угла камеры).

Чтобы уменьшить вероятность обнаружения этих артефактов, создатели намеренно снижают разрешение и/или качество, а также пользуются дополнительными фильтрами, оставляя лёгкое размытие.

3. Анализ аудиоматериалов

Анализ аудиоматериалов включает исследование интонации, модуляции и монотонности речи. Синтетически созданные голоса звучат подчас более монотонно, и им сложнее передать эмоциональные оттенки, что позволяет человеческому уху выявить подделку. Проверка на наличие и естественность фоновых шумов также важна, поскольку в реальных записях шумы часто присутствуют, а в дипфейках их либо нет, либо они добавлены искусственно, поэтому могут звучать не совсем естественно.

Однако здесь следует быть предельно аккуратными, ибо если речь идёт о видеозаписи, мы имеем право сопоставить два канала информации (аудиальный и визуальный), но в случае анализа только аудиосообщения ситуация значительно усложняется в силу того, что создатели могут использовать сглаживание для маскировки интонирования и добавлять шумы, что вкупе с намеренным снижением качества материала значительно усложняет процесс детекции, делая его практически неосуществимым без применения специальных средств (см. далее).

4. Анализ метаданных и цифровых характеристик

Углубленный анализ этой информации может способствовать выявлению различного рода несоответствий, например отсутствия или изменения цифровых водяных знаков; здесь также стоит обратить внимание на временные метки, анализ которых помогает выяснить, были ли изменения в файле.

Однако это может сработать только в том случае, если создатели, в целях сокрытия процесса редактирования, не прибегли к манипуляциям с метаданными или полностью не удалили нежелательные сведения.

5. Анализ с применением специализированного программного обеспечения

Упростить процесс детекции дипфейков помогают как готовые специализированные программы на основе искусственного интеллекта (Deepware Scanner, Sensity AI и т.д), так и настраиваемые (ResNet, EfficientNet и т.д.). Готовые, очевидно, проще в использовании, но могут не обладать достаточной гибкостью для решения специфических задач, в то время как настраиваемые позволяют адаптировать модели для более точного анализа, однако требуют определённых технических навыков.

В ответ на это создатели дипфейков применяют уже свои специализированные программы, заточенные для обхода детекторов. Также в их арсенале широко задействованы так называемые микроподделки, затрудняющие обучение нейросетей.

6. Верификация и тестирование материала на подлинность

Финальным этапом любого процесса детекции является верификация и тестирование материала на подлинность. Этот пункт включает

в себя фактчекинг и кросс-проверку, иными словами, сопоставление предоставленного контента с независимыми источниками. Наряду с этим важно также обращать внимание на логическую связанность материала с текущей ситуацией.

В данном случае в качестве маскировки дипфейков могут выступать ложные отзывы и комментарии, которые подтверждают правдоподобие контента, создавая иллюзию достоверности. В свою очередь, сам контент может быть размещён на платформах с минимальной модерацией, что в какой-то степени способен ввести человека в заблуждение.

Подводя итог, отметим, что информация всегда являлась ценнейшим и по праву ключевыми ресурсом, определяющим ход политических процессов и явлений. Однако в нынешних реалиях, когда дипфейки диктуют совершенно новые обстоятельства, понимание принципов их работы становится неотъемлемым фактором противодействия политической дезинформации в условиях не только локальных, но и глобальных противостояний. Проводя аналогии с XX в., современная информационная борьба с применением искусственного интеллекта носит характер своеобразной «гонки вооружений», в которой побеждает обладатель не только самого передового оружия по обману противника, но и самой передовой защиты.

Вполне очевидно, что из-за постоянной эскалации в области применения дипфейков с течением времени методы их обнаружения будут склоняться всё больше к использованию специализированного программного обеспечения. На данный момент времени, пока технологии отнюдь не совершенны, мы всё ещё можем полагаться на свои органы чувств, однако нельзя забывать о том, что в перспективе информационная реальность, благодаря использованию «глубоких фейков», будет подвергаться искажению, а значит, сомнению. Это, безусловно, приведёт к коренному изменению восприятия политической информации, а также усложнению процесса её верификации. Такое развитие событий может не только подорвать доверие к традиционным медиа и официальным источникам, но и спровоцировать появление новых механизмов контроля и фильтрации контента. В условиях, когда дипфейки станут обычной частью информационного политического пространства, ключевой задачей государства станет выработка коллективного иммунитета общества к дезинформации, основанного на медиаграмотности, технологической осведомлённости и критическом мышлении.

Список литературы

1. *Сунь-цзы.* Искусство войны : пер. с кит. М. : АСТ, 2022. 192 с.
2. *Goodfellow I J., Pouget-Abadie J., Mirza M., Xu B., Warde-Farley D., Ozair S., Courville A., Bengio Y.* Generative Adversarial Nets // Proceedings of the 27th International Conference on Neural Information Processing Systems (NIPS). Montreal, Canada. 2014. Vol. 2. P. 2672–2680. https://doi.org/10.48550/arXiv.1406.2661
3. *Сафронов Д. А., Кацер Ю. Д., Зайцев К. С.* Поиск аномалий с помощью автоэнкодеров // International Journal of Open Information Technologies. 2022. Т. 10, № 8. С. 39–45. EDN: BPCTIS
4. *Калиев А., Рыбин С. В.* Синтез речи: прошлое и настоящее // Компьютерные инструменты в образовании. 2019. № 1. С. 5–28. https://doi.org/10.32603/2071-2340-2019-1-5-28, EDN: VSZNHC
5. *Авдеенко Е. В.* Социально-психологические основы манипулятивного процесса // Каспийский регион: политика, экономика, культура. 2023. № 1 (74). С. 117–121. https://doi.org/10.54398/1818510X_2023_1_117, EDN: XJRFAY
6. *Jones E. E., Davis K. E.* From Acts to Dispositions: The Attribution Process in Person Perception // Advances in Experimental Social Psychology. 1965. Vol. 2. P. 219–266. https://doi.org/10.1016/S0065-2601(08)60107-0
7. *Каримов К.* Эксперты рассказали, кто может стать жертвой дипфейков // РИА Новости. 2021. URL: https://ria.ru/20210814/dipfeyk-1745306675.html (дата обращения: 24.11.2024).
8. *Аникин Д.* Не верь ушам своим: голосовые дипфейки // Блог Касперского. 2023. URL: https://www.kaspersky.ru/blog/audio-deepfake-technology/35694/ (дата обращения: 24.11.2024).
9. *Simonite T.* A Zelensky Deepfake Was Quickly Defeated. The Next One Might Not Be // WIRED. 2022. URL: https://www.wired.com/story/zelensky-deepfake-facebook-twitter-playbook/ (дата обращения: 24.11.2024).
10. *Лукина Ю. В.* Использование дипфейков в общественно-политической жизни // Русская политология. 2023. № 2 (27). С. 41–48. EDN: CEXDQA
11. *Котляр М.* «У Пентагона произошел взрыв». Объясняем, почему это фейк // Газета.Ru. URL: https://www.gazeta.ru/social/2023/05/22/16743344.shtml (дата обращения: 24.11.2024).

Поступила в редакцию 03.12.2024; одобрена после рецензирования 20.12.2024;
принята к публикации 20.02.2025; опубликована онлайн 30.05.2025
The article was submitted 03.12.2024; approved after reviewing 20.12.2024;
accepted for publication 20.02.2025; published online 30.05.2025