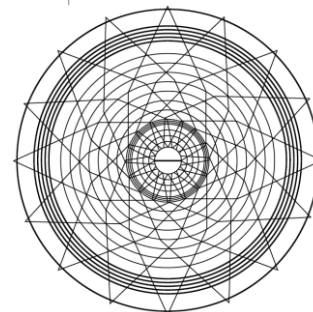


[Научные статьи]

Шомова С. А.

«Ах, обмануть меня не трудно...»: феномен политического дипфейка в коммуникативных практиках Рунета



<https://doi.org/10.17323/cmd.2025.30456>

«АХ, ОБМАНУТЬ МЕНЯ НЕ ТРУДНО...»: ФЕНОМЕН ПОЛИТИЧЕСКОГО ДИПФЕЙКА В КОММУНИКАТИВНЫХ ПРАКТИКАХ РУНЕТА

Шомова С. А.

доктор политических наук, профессор
Национальный исследовательский университет
«Высшая школа экономики»

(Москва, Россия)

sshomova@hse.ru



Аннотация:

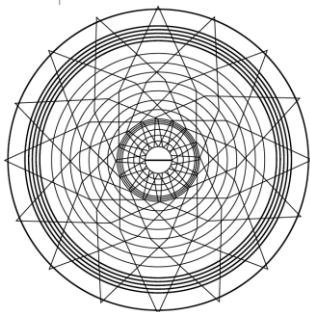
Политический дипфейк — явление относительно новое в повседневных коммуникативных практиках, однако активно завоевывающее современное медиапространство и связанное как с немалыми возможностями, так и серьезными рисками. В статье фиксируется узел академических проблем и дискуссий, связанных с изучением данного феномена, рассматриваются основные теоретические подходы к осмыслению дипфейка в целом (тезаурус, классификации, методы верификации), предлагается рабочее определение политического дипфейка, вычленяются его ведущие разновидности в рунете.

Ключевые слова: искусственный интеллект, дипфейк, политический дипфейк

Введение

«Старая максима “все не всегда так, как кажется” кажется более верной, чем когда-либо, в эпоху дипфейков. Дипфейк — это изображение, видео или аудиозапись, отредактированные с помощью алгоритма для замены человека в оригинале кем-то другим (особенно публичной фигурой) таким образом, чтобы это выглядело подлинным. Смысл части “фейк” (подделка) в слове “дипфейк” прозрачен: дипфейки не являются реальными. Суть компонента “дип” (глубина) менее очевидна: эта половина термина находится под особым влиянием *глубокого обучения*, то есть машинного обучения с использованием искусственных нейронных сетей с несколькими слоями алгоритмов», — так словарь Мерриам-Вебстера¹, сделавший слово «дипфейк» частью своего тезауруса еще в 2020 г., описывает концептуальные особенности феномена.

¹ Deepfake. (n.d.). In Merriam-Webster.com dictionary. <https://www.merriam-webster.com/dictionary/deepfake#dictionary-entry-1>



[Научные статьи]

Шомова С. А.

«Ах, обмануть меня не трудно...»: феномен политического дипфейка в коммуникативных практиках Рунета

Любопытно, что если в процитированном определении словарь фиксирует внимание на «иллюзии подлинности» (правдоподобию) такого контента, то чуть выше, в другой части этой же словарной статьи, акценты расставлены несколько иначе, с упором на принципиально дезинформационную сущность явления: дипфейк — это «изображение или запись, которые были убедительно изменены и обработаны с целью представить кого-то в ложном свете, как делающего или говорящего что-то, что на самом деле не было сделано или сказано» (там же).

Ворвавшись в академические словари и в индустриальные дискуссии несколько лет назад, концепт дипфейка быстро оброс собственной «мифологией» и вызвал обсуждение прецедентов последствий, споры о возможностях правоприменительных практик по отношению к такого рода контенту, попытки сформировать списки потенциальных маркеров и пр. Настоящая статья аккумулирует в себе результаты трехлетнего наблюдения за одним из самых распространенных типов дипфейков — политическим — в отечественном и (для сравнения) зарубежном медийном пространстве.

Внимание именно к этому типу контента связано не только с личным интересом автора к современным политическим технологиям, но и с высокой вирусогенностью «политического творчества» нейросетей, масштабируемостью и быстротой его распространения в социальных сетях, способностью вызывать некритичное доверие пользователей. Актуализирует тему исследования и то, что именно политические дипфейки, как говорится в исследовании DeepMind от Google, — самый популярный способ злоупотребления ИИ: наиболее распространенной целью подобных генераций было формирование общественного мнения или влияние на него. При этом аудитория вполне может не распознавать подобный контент как фейковый, а распространение его способно изменить мнение избирателей².

Исследованиям зарубежных технологических платформ вторят и наблюдения российских специалистов: по их мнению, в последние годы Россия столкнулась с резким ростом числа дипфейков, распознавать контент которых становится все сложнее³. По подсчетам аналитиков, «большинство русскоязычных дипфейков

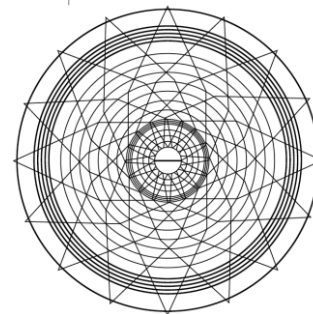
² Criddle, C. (2024, June 25). Political deepfakes are the most popular way to misuse AI. Arstecnica. <https://arstecnica.com/ai/2024/06/political-deepfakes-are-the-most-popular-way-to-misuse-ai/>

³ Чернов, А. (2025, 5 июня). На Глобальном цифровом форуме обсудили борьбу с дипфейками. Gazeta.ru. <https://www.gazeta.ru/social/news/2025/06/05/25968524.shtml>

[Научные статьи]

Шомова С. А.

«Ах, обмануть меня не трудно...»: феномен политического дипфейка в коммуникативных практиках Рунета



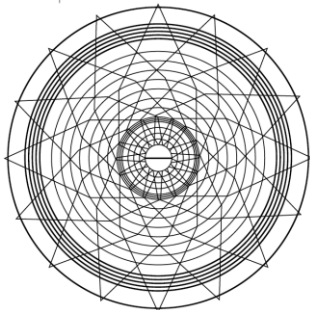
(89%) в 2025 г. были связаны с политической сферой, силовыми структурами или государственной безопасностью»⁴.

Следуя безымянной, но известной в гуманитарной науке формуле «прежде чем объяснить явление, следует его описать», дизайн исследования был выстроен по логике «от постановки теоретических рамок изучения политического дипфейка — через фиксацию отечественного и зарубежного опыта их создания и применения — к описанию и анализу конкретных кейсов и коммуникативных практик их использования». Основными исследовательскими вопросами стали вопросы о сущности и разновидностях политического дипфейка в рунете, а также о том, каковы его ведущие семантические и нарративные особенности. Соответственно, среди главных методов исследования следует назвать методы классификации, семантического и нарративного анализа, а также элементы дискурс-анализа. Целью работы стала фиксация современного состояния феномена политического дипфейка в рунете: его определение, классификация, выявление семантики и нарративов.

Дипфейк как коммуникативный феномен (определение, классификации, методы верификации)

Несмотря на то, что сам феномен дипфейка (как и в целом доступность генеративного ИИ для творчества пользователей) — относительно новое явление в Сети, он в некоторой степени осмыслен и в академическом, и в индустриальном поле. Исследователи подходят к изучению дипфейка как с чисто технологической, так и с философской, психологической, социологической, юридической, этической точек зрения. Они обозначают эпистемические угрозы, которые несут дипфейк-видео (Fallis, 2021), исследуют возможности алгоритмов в генерации текста и видео, преимущества дипфейков и связанные с ними риски, их влияние на общество (Giansiracusa, 2021). В ряде работ прослеживается, как создаются и обнаруживаются дипфейки, включая выявление недостатков защитных решений (Mirsky, Lee, 2021), анализируется восприятие дипфейков в профессиональной медиасреде — в том числе возможные преувеличения их опасности ради позиционирования института журналистики в качестве главного оплота против фальсификаций (Wahl-Jorgensen et al., 2021). Исследователи также предупреждают о распространяющейся практике «ложных ярлыков» — попытках делегитимизировать информацию путем объявления ее «дипфейком» (Hameleers

⁴ Лейба, Г. (2025, 7 апреля). Изображение хуже губернаторского. Коммерсант. <https://www.kommersant.ru/doc/7638958>



[Научные статьи]

Шомова С. А.

«Ах, обмануть меня не трудно...»: феномен политического дипфейка в коммуникативных практиках Рунета

& Marquart, 2023), рефлексиируют законодательные механизмы работы с дипфейковым контентом в разных странах (Ray, 2021) и т. д. Обозначая свежие тенденции и тренды в развитии литературы по вопросу, авторы статьи Deepfakes: evolution and trends резюмируют, что в целом исследования быстро продвигаются в направлении создания обобщенных решений для обнаружения подобного контента в реальном времени, понимания методов создания дипфейков и оценки воздействия на общество (Gil et al., 2023).

Однако при всех локальных, «частных» ракурсах, при всех множественных и разнообразных «подходах к снаряду», общая исследовательская оптика все еще заметно страдает от «размытости» взгляда на дипфейк и отсутствия единой внятной парадигмы его изучения. Хотя интерес к дипфейку весьма высок, в современной науке отмечается крайне фрагментарный характер существующей теоретической литературы, «серьезные пробелы в нарративах» и «тематическая напряженность и дисбалансы» в работах об этой разновидности фейкового контента (Vasist, Krishnan, 2022). Тем не менее можно констатировать, что основных «проблемных узлов», которым так или иначе уделяет внимание множество исследователей, в осмыслении дипфейка существует три: а) проблематичность самого сущностного определения феномена; б) подходы к классификации дипфейков и в) вопросы фиксации маркеров, потенциально указывающих на дипфейковый характер контента, и недостаток инструментов его распознавания.

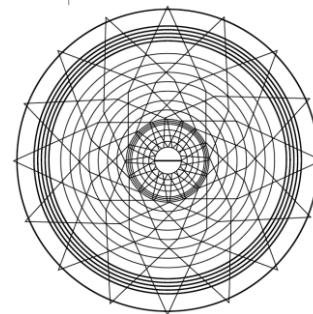
Первый вопрос носит не просто важный, но принципиальный характер, задающий теоретическую рамку восприятия термина: от того, понимаем ли мы под дипфейком злонамеренное искажение реальности, целенаправленную дезинформацию, или видим в этом термине прежде всего обозначение некой технологии (вне зависимости от добросовестности или недобросовестности ее применения), зависит и общественное восприятие термина, и мотивация решений крупных платформ по удалению дипфейк-контента, и механизмы правоприменения по отношению к тем или иным высказываниям в современной медиасреде. Определяя дипфейк как разновидность синтетического контента, сконструированного с помощью глубоких нейронных сетей⁵, одни авторы делают акцент на основах его производства, нейтрально трактуя интересующий нас феномен как «совокупность технологических трансформаций изображения и видео, созданных с использованием искусственного интеллекта» (Красовская и Гуляев, 2020), как методику компьютерного синтеза изображения, основанную на искусственном интеллекте, которая используется для соединения и наложения

⁵ Минцифры России, ФРИИ. (2024). Дипфейки: риски и возможности для бизнеса. <https://www.zircon.ru/upload/iblock/e0d/dipfejki-riski-i-vozmozhnosti-dlya-biznesa1.pdf>

[Научные статьи]

Шомова С. А.

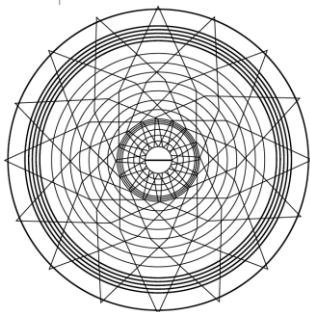
«Ах, обмануть меня не трудно...»: феномен политического дипфейка в коммуникативных практиках Рунета



существующих изображений и видео на исходные изображения или видеоролики (Иванов и Игнатовский, 2020), как контент (видеоролики, аудио или изображения людей), «созданный полностью или частично с помощью передовых алгоритмов искусственного интеллекта» (Graber-Mitchell, 2021), а другие — на том, что эта технология создает «гиперреалистичные видео, обработанные с помощью цифровых манипуляций, чтобы изображать людей, говорящих и делающих вещи, которые на самом деле никогда не происходили» (Westerlund, 2019), уточняя при этом, что обычно речь идет о продуцировании «поддельных видео» за счет замены лица человека на лицо другого человека с целью добиться высокой реалистичности и «невидимости» для зрителя (Tolosana et al., 2021).

Иными словами, уже на этом этапе осмысления интересующего нас феномена очевидна дихотомия: для кого-то самым важным в определении сущности дипфейка является его алгоритмическая природа, механизм создания контента искусственным интеллектом, а кто-то фактически ставит знак равенства между понятиями «дипфейк» и «дезинформация», подчеркивая поддельность, фальшивость созданных таким образом сообщений. Нам представляется весьма удачной, обобщающей разные подходы и операциональной для наших целей дефиниция, данная И. А. Ворониным и Д. П. Гавра, которые, выделяя три базовых подхода к созданию соответствующей формулировки (техноцентричный, нормативный и синтетический), предлагают определять дипфейк как «опирающийся на технологии искусственного интеллекта целенаправленно созданный коммуникационный продукт (видео, фотографии, аудиозаписи и т. д.), направленный на формирование у целевых аудиторий фальсифицированного образа социального субъекта с целью получения выгоды или нанесения ущерба» (Воронин и Гавра, 2024).

В вышеназванной работе предпринята попытка собрать воедино различные (и существенно отличающиеся друг от друга параметрами и степенью проработанности) классификации дипфейков, существующие в академическом поле, — по каналу восприятия (подробно рассмотрены всего два типа, соотносящиеся с аудиальными и визуальными дипфейками), сфере использования и цели создания, где авторы выделяют коммуникационные, дискредитирующие, мошеннические, развлекательные дипфейки (там же, с. 37–43). В данной работе хотелось бы также обратить внимание на классификацию Я. Кицмана и его соавторов (Kietzmann et al., 2019), в которой дипфейки систематизированы по формату. Исследователи выделяют фотодипфейки (например, сгенерированные изображения, показывающие, как человек может выглядеть спустя десятилетия), аудиодипфейки (голосовые сообщения, изменяющие речь, добавляющие что-то к



[Научные статьи]

Шомова С. А.

«Ах, обмануть меня не трудно...»: феномен политического дипфейка в коммуникативных практиках Рунета

ней или имитирующие голос), видеодипфейки, в свою очередь, подразделяющиеся на несколько типов (Face-swapping — подмена лиц с одного на другое, Face-morphing — плавное превращение лица и т. д.) и аудиовидеодипфейки (например, наиболее привычная нам «говорящая голова»). Н. Джансиракуса делает к этой классификации важное дополнение, добавляя в нее текстовые дипфейки — например, генерацию фейковых новостей, сообщений или комментариев с помощью ИИ (Giansiracusa, 2021). Исходя из преподавательского опыта, мы бы упомянули в качестве примера текстовую генерацию несуществующих литературных источников или вымышленных цитат для исследовательских работ.

М. Вестерлунд предлагает классифицировать дипфейки по их субъекту, то есть по типу их производителей: 1) сообщества любителей дипфейков, 2) политические игроки, такие как правительства и различные активисты, 3) злонамеренные субъекты, мошенники и 4) легитимные акторы, как, например, телевизионные компании (Westerlund, 2019). У разных авторов можно встретить также классификации дипфейков по назначению (затрагивающие сферы политики, шоу-бизнеса и жизни знаменитостей, юмора, досуга и др.), по области криминального применения (порнографические подделки, банковские мошенничества) и пр. Система является открытой: с развитием технологий создания дипфейков в медиапространстве появляются все новые их разновидности, все более изощренные варианты, порой не укладывающиеся в уже существующие классификации.

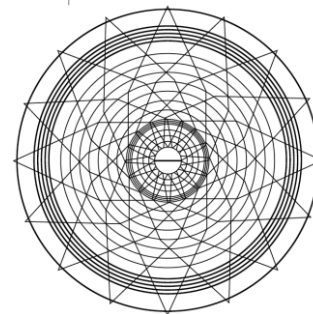
Третий проблемный узел из обозначенных нами выше, связанный с признаками, характеристиками дипфейков и инструментами их распознавания, вызывает особенно глубокий интерес как теоретиков, так и практиков индустрии: именно он определяет возможные правоприменительные решения по маркировке и удалению подобного контента, именно с ним зачастую сопряжены экспертные дискуссии и развитие методик алгоритмического предупреждения распространения недостоверной информации в Сети. Заметим, что авторы научных работы в этой области рассматривают проблему как с чисто технической, основанной на ИИ-возможностях, точки зрения, так и с позиций общего развития медиаграмотности, критического мышления, навыков осознанного потребления информации (Matern et al., 2019; Sandotra& Arora, 2024; Tolosana et al., 2021; Yang et al., 2019).

Один из авторов последней в упомянутом списке работы, представляющей собой обзор технологий создания и обнаружения дипфейков с использованием подходов глубокого обучения, — С. Лю, — в более популярной статье для издания Conversation писал: «Видеоролики формата Deepfake трудно обнаружить

[Научные статьи]

Шомова С. А.

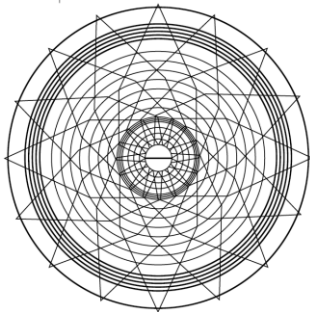
«Ах, обмануть меня не трудно...»: феномен политического дипфейка в коммуникативных практиках Рунета



неподготовленному глазу, поскольку они могут быть довольно реалистичными. Независимо от того, используются ли они в качестве личного оружия мести, для манипулирования финансовыми рынками или для дестабилизации международных отношений, видео, изображающие людей, делающих и говорящих то, чего они никогда не делали и не говорили, они представляют собой фундаментальную угрозу давней идее «увидеть — значит поверить». Больше нет... Наука обнаружения дипфейков, по сути, представляет собой гонку вооружений: мошенники совершенствуют свои вымыслы, поэтому наши исследования всегда должны стараться не отставать и даже немного опережать их... Мы нашли способ добавлять к цифровым фотографиям или видео специально разработанный шум, который не виден человеческому глазу, но может обмануть алгоритмы распознавания лиц»⁶. Таково одно из технических решений проблемы обнаружения дипфейков, предлагаемых специалистами, хотя с его помощью не удалось решить ее целиком. Другие алгоритмические инструменты также совершенствуются по мере повышения изобретательности дипфейков, однако до стопроцентного обнаружения недостоверного контента, по-видимому, все еще далеко.

От IT-специалистов в обозначенной выше «гонке вооружений» не отстают фактчекерские, просветительские, медийные платформы и ресурсы, делающие ставку на повышение медийной и ИИ-грамотности пользователя социальных сетей. Так, российская мультимедийная энциклопедия «Сделано.медиа» предлагает следующий перечень характеристик, которые могут указывать на дипфейковый характер контента: *моргание* (как правило герои дипфейков моргают реже, чем обычные люди, и закрывают глаза лишь наполовину, в то время как реальный человек так бессознательно сделать не может); *несоответствие движений губ и голоса* (современные алгоритмы не в состоянии полностью воспроизводить мимику рта при разговоре, из-за чего герои дипфейков в большинстве своем либо совсем не говорят, либо говорят крайне мало); *морщины, мимические изменения во внешности* (лицо героя дипфейка, как правило, более совершенно, чем лицо реального человека — в пример приводится реклама оператора связи «Мегафон» с легальным дипфейком Брюса Уиллиса, в которой артист выглядит намного моложе, имеет идеальную кожу и блестящие глаза, что значительно отличается от его образа в ролике для банка «Траст», хотя эта реклама была снята на девять лет раньше); *ошибки нейросети* (при высоких бюджетах их удастся избегать, однако при менее качественном продакшене можно заметить, как, например, открывается

⁶ Lyu, S. (2019, June 25). Detecting deep fakes by looking closely reveals a way to protect against them. The Conversation. <https://theconversation.com/detecting-deepfakes-by-looking-closely-reveals-a-way-to-protect-against-them-119218>



[Научные статьи]

Шомова С. А.

«Ах, обмануть меня не трудно...»: феномен политического дипфейка в коммуникативных практиках Рунета

маска актера, нанесенная на другого человека)⁷. В других научных, популярных, образовательно-просветительских материалах, помимо только что перечисленных, можно также обнаружить такие признаки, как ошибки с тенями и светом, размытые детали внешности (уши, губы, волосы), роботизированный тон речи.

Все списки маркеров, отличающих дипфейк, носят не только инструментально-прагматический, но и, наш взгляд, онтологический характер: вопрос о том, что именно считать дипфейком (только ли завуалированный, мимикрирующий под правду, пытающийся выдать себя за подлинный контент или же любое творчество ИИ с подменой лиц, даже не пытающееся спрятать признаки нейросетевого присутствия), важен и для теории, и для практики работы с недостоверной информацией. Мы придерживаемся взгляда, что дипфейк — это прежде всего технология, и лишь во вторую очередь — попытка дезинформации. Поэтому и откровенно сконструированные искусственным интеллектом политические ролики, не старающиеся замаскировать свою «поддельную» природу, но имеющие зримый коммуникативный посыл, мы склонны также квалифицировать как дипфейк-контент.

Подобного контента, в том числе создаваемого «легально», в Сети немало. В качестве примера дипфейка, не претендующего на то, чтобы ввести зрителей в заблуждение, можно привести «рождественское» видео с королевой Елизаветой II, созданное для британского канала Channel 4 оscarоносной студией VFX Framestore. В нем «королева» без обиняков высказывает мнение о прошедшем 2020 г., откровенничает с аудиторией, а под конец даже танцует на столе. По словам технического директора Channel 4 Яна Катца (Ian Katz), «видео показали аудитории, чтобы привлечь внимание к опасностям технологии дипфейков. Он описывает ролик как мощное напоминание о том, что мы больше не можем доверять своим глазам»⁸. Хотя данное видео и вызвало неоднозначное (скорее, негативное) отношение британцев, оно снова доказывает, что дипфейк может иметь просветительские цели и быть не только попыткой злонамеренного введения в заблуждение, но и шуткой, социально-ориентированным контентом, «предупреждающей» рекламой или акцией, нацеленной на повышение цифровой грамотности. Все эти форматы объединяет лишь факт создания с помощью ИИ, что

⁷ Черных, А. (2024, 13 ноября). Три фейка недели: про «школьную диктатуру», «добровольную вакцинацию» и «слежку для турбулентности». Сделано.медиа.

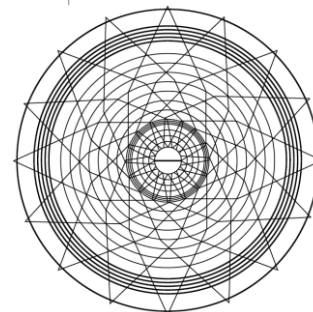
https://sdelano.media/3fakes/?ysclid=migdmhiluw_514971621

⁸ Парфенов, В. (2020, 25 декабря). Дипфейк королевы Елизаветы II выступил на британском телеканале и вызвал бурю негодования. Naked Science. <https://naked-science.ru/article/media/dipfejk-korolevy-elizavety-ii-vystupil-na-britanskom-telekanale-i-vyzval-buryu-negodovaniya>

[Научные статьи]

Шомова С. А.

«Ах, обмануть меня не трудно...»: феномен политического дипфейка в коммуникативных практиках Рунета



снова подводит к мысли о первичности *технологии* в определении дипфейка и вторичности его *целевой направленности*.

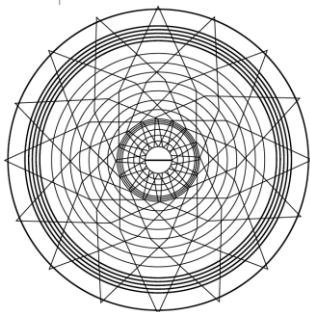
В то же время, каковы бы ни были целевые установки дипфейка, на повестке дня внимательного пользователя по-прежнему остаются задачи верификации интернет-контента, а также обнаружения и разоблачения недостоверной информации. По мнению М. Вестерлунда, существует четыре способа борьбы с дипфейками: 1) законодательное регулирование, 2) корпоративная политика и добровольные действия, 3) образование и обучение, 4) использование технологических инструментов, которые позволяют обнаруживать дипфейки, аутентифицировать контент и предотвращать распространение дипфейков (Westerlund, 2019). Все эти способы берут на вооружение государственные структуры, технологические платформы-гиганты, фактчекинговые ресурсы и научное сообщество, формируя пул инструментов атрибутирования и маркировки недостоверного ИИ-контента. Особенно актуально это в тех случаях, когда речь идет о политических дипфейках, способных критическим образом «исказить наше коллективное понимание социально-политической реальности»⁹.

Политический дипфейк: к уточнению тезауруса

Ни в англоязычной, ни в русскоязычной литературе пока не существует единого устоявшегося мнения по поводу того, какой именно контент можно четко атрибутировать как политический дипфейк. Общепринятым является взгляд на этот феномен как на созданный с помощью нейросетей недостоверный медиапродукт (видео, аудио, фотоизображение и т. д.), направленный на занижение имиджа политических лидеров, дискредитацию акторов политики, в том числе партий, влияние на ход электоральных процессов и т. д. Так, например, Е. А. Виноградова в своих работах определяет политический дипфейк как «специальную кампанию с применением технологии искусственного интеллекта для подрыва репутации политических лидеров с целью изменения хода избирательной борьбы или для дискредитации действующего политика» (Виноградова, 2023).

Однако реальность гораздо сложнее, нежели академические дефиниции. Приведенное выше определение выглядит не до конца корректным хотя бы потому, что на сегодняшний день известно немало дипфейков, используемых в политическом дискурсе с обратной целью — для создания «позитивной» электоральной повестки, то есть не занижения, а, напротив, повышения имиджевых характеристик лидера. В качестве примера можно привести достаточно

⁹ Criddle, C. (2024, June 25). Political deepfakes are the most popular way to misuse AI. Arstecnica. <https://arstecnica.com/ai/2024/06/political-deepfakes-are-the-most-popular-way-to-misuse-ai/>



[Научные статьи]

Шомова С. А.

«Ах, обмануть меня не трудно...»: феномен политического дипфейка в коммуникативных практиках Рунета

старый кейс, относящийся к 2020 г. и позволивший комментаторам заговорить о политических дипфейках как о «джинне, выпущенном из бутылки»¹⁰. Накануне выборов в Законодательное собрание индийской столицы Дели в WhatsApp распространили два видео, быстро ставшие вирусными. В них Манодж Тивари, президент Партии БДП, критиковал действующее правительство Дели и говорил на разных языках: в одном из роликов по-английски, во втором — на хинди, на диалекте хариани. «“Мы использовали алгоритм “фальшивой синхронизации по губам” и обучили его по речам Маноджа Тивари”, — говорит Сагар Вишной, главный стратег в The Ideaz Factory... Реакция на эти видео была обнадеживающей. “Домохозяйки в группе сказали, что им было радостно наблюдать, как наш лидер говорит на их языке”, — сказал он. После первого ролика партия выпустила второе видео, на котором Тивари говорит уже по-английски, предназначенное для “городских избирателей Дели”»¹¹. В данном случае дипфейк был призван не подорвать, а, напротив, укрепить в глазах различных электоральных страт репутацию лидера, создавая его образ как образа политика, «способного говорить с аудиторией на разных языках», — хотя на самом деле ролик был сгенерирован нейросетью.

Существуют и иные примеры «позитивного» применения дипфейков в политическом дискурсе: подобные кейсы (например, использование такого контента боливийским каналом коренных народов Voces Digitales Indígenas для привлечения молодежи или проекты цифровой культурной дипломатии в социальных сетях Мексики) обсуждаются, в частности, в работах российских исследователей (Цветкова и Гришанина, 2023). Основопологающей установкой в таких случаях становится не обрушение репутации, а продвижение той или иной политической идеи, политического бренда, политической ценности или идентичности. Это, на наш взгляд, коренным образом меняет представление о политическом дипфейке исключительно как дискредитирующем феномене.

Опираясь на общую дефиницию И. Воронина и Д. Гавра, приведенную выше, и учитывая, что дипфейк политической направленности не всегда нацелен на злоумышленное введение в заблуждение и не всегда дискредитирует своих персонажей (но признавая при этом, что в подавляющем большинстве случаев такой контент действительно «склонен» обманывать и подрывать репутацию),

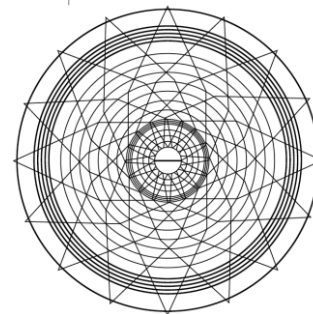
¹⁰ Nileshe, C. (2020, February, 18). We've Just Seen the First Use of Deepfakes in an Indian Election Campaign. Vice.com. https://translated.turbopages.org/proxy_u/en-ru.ru.ea60a722-68a5a810-6267aedef-74722d776562/https/www.vice.com/en/article/the-first-use-of-deepfakes-in-indian-election-by-bjp/

¹¹ В Индии дипфейк использовали для «позитивной» избирательной кампании (2020, 19 февраля). Хабр. <https://habr.com/ru/news/489006/>

[Научные статьи]

Шомова С. А.

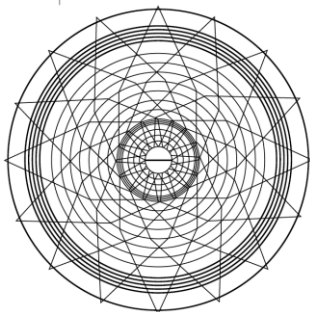
«Ах, обмануть меня не трудно...»: феномен политического дипфейка в коммуникативных практиках Рунета



хотелось бы предложить следующее рабочее определение интересующего нас феномена. Под политическим дипфейком мы предлагаем понимать целенаправленно созданный медийный продукт (видео, фотографии, аудиозаписи и т. д.), сконструированный на основе технологий искусственного интеллекта, эксплуатирующий актуальный новостной повод и связанный с политическим дискурсом напрямую или опосредованно (за счет центрального персонажа, семантики сообщения, отсылки к политическим событиям, датам, идеям, ценностям и т. д.); будучи направлен на формирование у целевых аудиторий сфальсифицированного с той или иной степенью открытости образа социального субъекта, такой медийный продукт может ставить своей целью как нанесение репутационного ущерба, так и продвижение идеи или личности, и во всех случаях рассчитан на получение определенной информационной выгоды.

Для нас важно отметить, что авторы академических исследований последних лет несколько расходятся в мнениях о том, насколько реально политические дипфейки способны воздействовать на аудиторию: действительно ли они так сильно влияют на мнение людей о политиках или партиях, что представляют собой угрозу для демократии, и в самом ли деле почти неотличимы от реальных новостных видео или подлинных записей выступлений политиков? В работе «Обманут дважды: люди не умеют распознавать дипфейки, но думают, что умеют» исследователи обращают внимание на когнитивную сложность распознавания такого контента и на то, что люди существенно переоценивают свою способность к обнаружению дипфейков, становясь из-за этого весьма уязвимыми перед ними (Köbis et al., 2021). Т. Доббер с соавторами на основе проведенного эксперимента утверждают, что определенные технологии (а именно используемые ими методы микротаргетинга) усиливают воздействие дипфейков, позволяя злоумышленникам адаптировать данный контент к восприимчивости получателя (Dobber et al., 2020). В другом экспериментальном исследовании авторский коллектив, частично пересекающийся с предыдущим, приходит к выводам, что дипфейки все же менее достоверны, нежели подлинные новости по той же теме; «Хотя мы и обнаружили, что дезинформация влияет на восприятие достоверности и оценку источников людьми, которые склонны соглашаться с позицией, изложенной в аргументах дезинформации, наши выводы свидетельствуют о том, что сильная общественная обеспокоенность по поводу дестабилизирующего воздействия дипфейков на демократию все же не полностью оправдана» (Hameleers et al., 2022).

В определенной степени обнадеживающими для всех, кто верит в усилия цифровой и медиаграмотности, являются выводы статьи «Обнаружение



[Научные статьи]

Шомова С. А.

«Ах, обмануть меня не трудно...»: феномен политического дипфейка в коммуникативных практиках Рунета

политических дипфейков», авторы которой, используя дипфейки с участием Б. Джонсона и Б. Обамы (в которых оба политика говорили то, чего никогда не говорили в реальной жизни), провели три последовательных эксперимента и обнаружили, что, чем выше были оценки участников по аналитическому мышлению (исследования 1–2) и политическим интересам (исследование 1), тем лучше они распознавали дезинформацию. Более того, люди с развитыми аналитическим мышлением и политическими интересами лучше выявляли неточность фейковой новостной статьи — независимо от того, была ли она сопровождена дипфейк-видео (Appel, Prietzel, 2022). К схожим выводам приходят в своей работе и С. Барари с коллегами (Barari, 2021), исследование которых демонстрирует, что, хотя дипфейки могут быть весьма эффективными в манипуляции информацией, они не превосходят по степени влияния на аудиторию традиционные форматы новостей. При этом решение проблемы требует серьезных усилий на уровне образования и повышения грамотности населения в области медиатехнологий (развитие критического мышления, уроки политической, цифровой, медиаграмотности и т. д.).

Однако воспринимать дипфейк как безобидный контент все же, безусловно, не стоит. Скандальные истории, многие из которых заканчиваются не просто разочарованием обманутых потребителей информации, как в случае с «примирением» принцев Уильяма и Гарри на коронации их отца или с ложными «звонками» Байдена избирателям в 2024 г., а даже судебными исками (например, в апреле 2024 г. в Италии были заведены уголовные дела по статье «Клевета» против двух подозреваемых, которые создавали дипфейки порнографического характера с участием премьер-министра Италии Дж. Мелони, меняя лица порноактрис на ее лицо)¹², не позволяют относиться к политическому дипфейк-контенту легкомысленно. Это соображение актуализируется и постоянным прогрессом в изобретении новых, более изощренных типов и форматов дипфейк-контента: если несколько лет назад вершиной фальсификации казалось видео с Н. Пелоси, где она плохо выговаривает слова, а речь кажется заторможенной и бессвязной (при этом воздействие на реальную запись сводилось «всего лишь» к искусственному замедлению и искажению речи политика¹³), то сегодня мы имеем дело с высококвалифицированными, сгенерированными на безупречном техническом уровне подделками внешности, речи, манеры поведения, паттернов реакций политиков. Особенно много подобных фальшивок стало появляться в

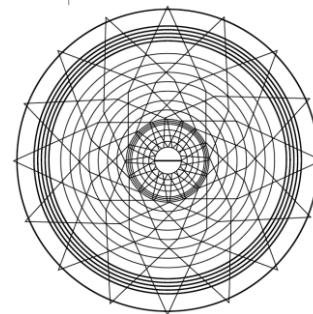
¹² См.: Скандалы с дипфейками: зарубежный опыт (2025, 7 апреля). Коммерсант. <https://www.kommersant.ru/doc/7638961>

¹³ См. об этом, например: Clayton, T. (2023, December 5). 11 Examples Of Deepfakes In Politics. Rigorous. <https://rigorousthemes.com/blog/examples-of-deepfakes-in-politics/>

[Научные статьи]

Шомова С. А.

«Ах, обмануть меня не трудно...»: феномен политического дипфейка в коммуникативных практиках Рунета



последние полтора-два года. Они связаны, в частности, с выборами президента США в 2024 г.¹⁴

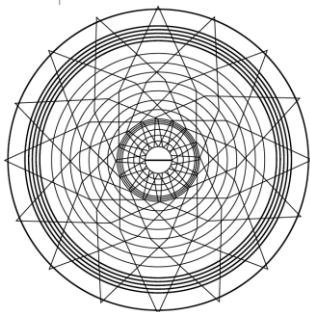
Впрочем, о разновидностях современного политического дипфейка и возможных подходах к его классификации имеет смысл поговорить отдельно.

Дипфейк в политическом пространстве рунета: персоналии и семантика

Не только в зарубежных работах, о которых было упомянуто выше, но и в русскоязычной академической традиции существуют попытки классифицировать дипфейки в целом и политический дипфейк-контент в частности. Они предпринимаются разными авторами, однако имеют, на наш взгляд, один общий недостаток (вполне объяснимый из-за малоизученности феномена): большинство этих классификаций опираются на самые разные параметры, несводимые к общему предметному полю, и это делает предложенные «списки» похожими на разношерстное лоскутное одеяло. Так, Е. А. Виноградова, изучая когнитивное влияние политического дипфейка на аудиторию, перечисляет в общем ряду контент, выделяемый по принципу персоналии («дипфейк действующего политического лидера» или «дипфейк умершего политика»), целенаправленно мошеннические сообщения («дипфейк-фишинг»), дипфейки, рассчитанные на определенный тип аудитории, — например, склонной к виртуальному взаимодействию («дипфейк-аватар»), шутливо-пародийные дипфейки («дипфейк-карикатура») и т. д. (Виноградова, 2024). Н. А. Цветкова и Т. А. Гришанина идут другим путем и, не пытаясь создать общую типологию, делят политические дипфейки по широте географии распространения (высокий уровень, средний и низкий), стратегии влияния (помимо информационно-психологического влияния, к которому авторы относят деструктивные дипфейки, связанные с дезинформацией, это также культурная цифровая дипломатия, персональный бренд инфлюэнсера и др.) и техническим характеристикам — частичный и полный дипфейк (Цветкова, Гришанина, 2023). Выше уже было сказано, что система жанров и видов дипфейк-контента, в том числе в политическом дискурсе, сегодня носит «открытый» характер, а потому, несмотря на то, что с большинством из существующих классификаций хотелось бы подискутировать, каждая подобная попытка видится нам важным вкладом в осмысление данного медийного феномена.

Наблюдая за политическим дипфейком в социальных сетях и мессенджерах рунета в течение нескольких лет (общее количество проанализированных сообщений — более 100), хотелось бы предложить свою, локальную и узкую,

¹⁴ Разбор конкретных кейсов с примерами фактчека см. подробнее: Farid, H. (2024, October 28). The 2024 election: The growing threat of deepfakes. The 2024 Election: Deepfakes. <https://farid.berkeley.edu/deepfakes2024election/>



[Научные статьи]

Шомова С. А.

«Ах, обмануть меня не трудно...»: феномен политического дипфейка в коммуникативных практиках Рунета

классификацию (одну из многих возможных, не претендующих при этом на полноту охвата явления — это сегодня попросту невозможно), которая рассматривает дипфейк-контент с точки зрения его семантики, учитывая одновременно ведущее целеполагание создания коммуникативного меседжа и характер задействованных в нем персоналий. Отметим, что в эмпирическую базу вошли политические дипфейки, имеющие визуальную природу (главным образом видеоролики, хотя иллюстративный и фото-контент тоже рассматривался автором) и циркулировавшие в рунете начиная с 2020 г., однако ниже мы постарались привести наиболее свежие примеры, относящиеся к 2024–2025 г.

Первый тип политических дипфейков, который стоит рассмотреть, наиболее распространен и может быть условно обозначен как *agenda-дипфейк* — он тесно связан с актуальной новостной повесткой, эксплуатирует острое любопытство россиян к происходящему вокруг и зачастую затрагивает важнейшие, витальные интересы массовой аудитории, поскольку касается вопросов безопасности, политики, экономики, повседневной жизни регионов. Типичным примером можно считать, например, дипфейк-контент, использующий образ экс-нардепа Украины, блогера О. Царева (например, недавний ролик, где «Царев» рассказывает, что землетрясение на Камчатке — «результат применения сейсмического оружия со стороны США, поражения базы подводных лодок и удара по российской ядерной триаде», с количеством жертв, «о которых мы узнаем десятилетия спустя или никогда»). Сам Царев в своем телеграм-канале прокомментировал ролик так: «Нейросети становятся все искуснее, но все равно видно, что это нейросеть. И ведь все равно находятся идиоты, которые воспринимают такие ролики за правду»¹⁵. Другие примеры — дипфейк с «участием» мэра Белгорода Валентина Демидова о якобы готовящейся полной эвакуации города (также разоблаченный самим чиновником¹⁶); видео, на котором премьер-министр И. Боложан якобы анонсирует участие Румынии в военном конфликте на территории Украины (опровергнутое МВД страны)¹⁷ и т. д.

Характерный признак такого рода дипфейков — участие в них официальных лиц или лидеров общественного мнения и быстрое опровержение с их стороны. С точки зрения семантики подобные ролики почти всегда являются не только остроактуальным откликом на news agenda, но и затрагивают темы, вызывающие

¹⁵ Телеграм-канал Олега Царева (2025, 1 августа). <https://t.me/olegtsarov/31843>

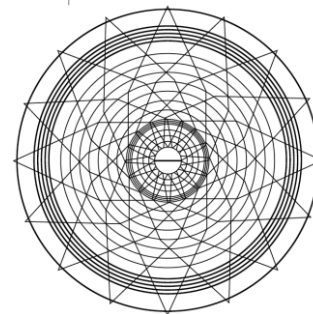
¹⁶ Телеграм-канал Демидов (2025, 14 августа). https://t.me/v_v_demidov/8585

¹⁷ Москаленко В. МВД Румынии опровергло данные о решении участвовать в конфликте на Украине (2025, 29 июля). АиФ.ру. <https://aif.ru/politics/mvd-rumynii-oproverglo-dannye-o-reshenii-uchastvovat-v-konflikte-na-ukraine>

[Научные статьи]

Шомова С. А.

«Ах, обмануть меня не трудно...»: феномен политического дипфейка в коммуникативных практиках Рунета



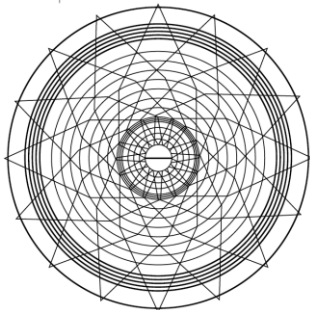
страх, панику, гнев, отчаяние, то есть могут быть атрибутированы как «фейки-пугала» (Шомова, 2023). Нарративно дипфейк-видео, о которых идет речь, практически всегда содержат одного повествователя (характерный формат ролика — «говорящая голова») и представляют собой краткое обращение, репрезентирующее в недостоверном виде и манипуляционном ключе актуальные новостные нарративы СМИ. Возможный подход фактчекинга — поиск исходного видео (которое чаще всего видоизменено с помощью наложения иного аудиального ряда), обнаружение признаков нейросетевой генерации (например, визуальные несоответствия, о которых шла речь в первом разделе данной статьи), отслеживание официальных реакций, задействованных в дипфейке лиц.

Второй тип политического дипфейка, на который хотелось бы обратить внимание, — это *агитационно-пропагандистский дипфейк*, меньше предыдущего связанный с актуальной новостной повесткой, но рассчитанный на продвижение определенных политических установок, ценностей, идеологем. Среди известных примеров — дипфейк-ролик «с участием» губернатора белгородской области В. Гладкова, в котором он рассказывает «хорошую новость»: во всех школах Белгородчины вводятся уроки украинского языка и украинской истории, потому что это «наиболее востребованный язык в нашем регионе с большой перспективой трудоустройства». За основу дипфейка было взято реальное обращение Гладкова, опубликованное в официальном телеграм-канале, а на фальшивость ролика указывает неестественность мимики и моргания в сочетании с плоской роботизированной интонацией¹⁸. Стоит отметить, что, по мнению российских экспертов, «излюбленные персонажи изготовителей такого контента — главы российских регионов, а наиболее часто затрагиваемая тема — спецоперация и все, что с ней связано»¹⁹.

Подобные манипулятивные приемы отрабатываются с разных сторон, в разных странах: так, к дипфейкам были отнесены, например, видеообращение В. Зеленского 16 марта 2022 г. к народу и вооруженным силам Украины с призывом сложить оружие и сдаться российской армии, поздравления российских спикеров с Днем независимости России — 2025 с упоминанием о некоем «преемнике» В. Путина или несколько роликов, завирусившихся в июне 2025 г. и использующих образы казахстанских чиновников, журналистов, лидеров общественного мнения

¹⁸ Глава Белгородской области заявил о необходимости изучения украинского языка (2024, 20 августа). Лапша Медиа. <https://lapsha.media/feiky/gubernator-belgorodskoj-oblasti-zajavil-o-neobходимosti-objazatel'nogo-izuchenija-ukrainskogo-jazyka-i-istorii/>

¹⁹ Лейба, Г. (2025, 7 апреля). Изображение хуже губернаторского. Коммерсант. <https://www.kommersant.ru/doc/7638958>



[Научные статьи]

Шомова С. А.

«Ах, обмануть меня не трудно...»: феномен политического дипфейка в коммуникативных практиках Рунета

с призывами к участию в СВО (хотя наемничество в этой стране запрещено законом). «Активизация дипфейков на постсоветском пространстве рассматривается экспертами как часть более широкой стратегии информационных манипуляций. Цель — вызвать общественное недоверие к публичным фигурам, посеять хаос и создать фальшивую повестку в социальных сетях», — отмечали по этому поводу СМИ центральноазиатских стран²⁰.

С точки зрения нарративных особенностей данный тип дипфейков весьма схож с предыдущим, однако семантика его гораздо шире: концептуально он может затрагивать самые разнообразные политические и социальные темы, продвигая политическую повестку и делая попытки сформировать общественное мнение в нужном создателям ключе. С технической точки зрения подобный контент может быть достаточно качественным, что делает его «визуально достоверным для большинства пользователей»²¹, однако внимательный взгляд может зафиксировать неестественную мимику говорящего или то, что изображение несколько «рябит». Подобная внимательность в век нейросетей становится must-have навыком для интернет-пользователей, не говоря уже о привычке перепроверять информацию, поданную в визуальном формате.

В отличие от двух предыдущих типов, *третий* интересующий нас феномен — *предвыборные дипфейки*, связанные с электоральным циклом, — далеко не всегда носит строго завуалированный характер и не всегда пытается выглядеть достоверными и реальными съемками. Разумеется, существует дипфейк-контент, активно пытающийся ввести избирателя в заблуждение — особенно щедрыми на такие ИИ-подделки были выборы президента в США в 2024 г. В качестве примеров можно привести серию фейковых фотографий, изображающих Д. Трампа в компании с чернокожими людьми, возможно, в попытке привлечь чернокожих избирателей; серию снимков с изображением президента Дж. Байдена в военной форме, якобы готового дать разрешение на нанесение военных ударов; видео Камалы Харрис с «нетрезвой» и бессвязной речью. В большинстве подобных случаев заметны рассинхронизация аудио и видео, неестественность изображения и/или признаки визуальной ИИ-генерации: фотомонтаж, дезинформация объектов на снимках и пр.²². При этом

²⁰ Дипфейки и дезинформация: в Казахстане фиксируют волну фальшивых видео с призывами к участию в военных действиях (2025, 27 июня). Maili.uz. <https://maili.uz/2025/06/27/deepfakes-and-disinformation-kazakhstan-records-wave-of-fake-videos/>

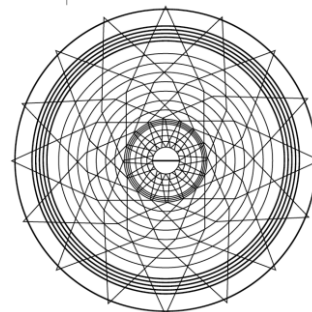
²¹ Там же.

²² Farid, H. (2024, October 28). The 2024 election: The growing threat of deepfakes. The 2024 Election: Deepfakes. <https://farid.berkeley.edu/deepfakes2024election/>

[Научные статьи]

Шомова С. А.

«Ах, обмануть меня не трудно...»: феномен политического дипфейка в коммуникативных практиках Рунета



распространение сгенерированных с помощью ИИ дипфейков настолько настораживает американскую общественность и власти, что там в настоящее время быстро растет количество штатов, ограничивающих использование ИИ в ходе избирательных кампаний (даже при создании обычных рекламных роликов, постеров и т. д.).

Впрочем, нередко в Сети обнаруживаются и политические предвыборные дипфейки, не пытающиеся скрыть свой ИИ-характер. Так, в ходе кампании выборов в Госдуму 2021 г. пользователи рунета активно обсуждали видеоролик, героями которого стали глава МИД России С. Лавров, министр обороны С. Шойгу и главный врач больницы №40 в Коммунарке Д. Проценко – по сюжету ролика, они «появлялись» дома у молодого избирателя и тем самым стимулировали того проявить гражданскую ответственность и пойти на выборы²³. В данном случае мы имеем дело не с намеренной дезинформацией пользователя, а с относительно нестандартным агитационным ходом, рекламным креативом (другой вопрос — удачным или нет).

В августе 2025 г. политолог А. Кынев рассказал в своем телеграм-канале о видео, сгенерированном ИИ для избирательной кампании выборов главы Республики Коми²⁴. Не слишком изобретательно сделанный рекламный ролик показывает нам нынешнего руководителя региона, шагающего по заснеженным просторам, под северным сиянием, в компании бурого медведя (слоган: «Небо, под которым хочется жить»). В следующем ролике этой серии политик вместе со своим медведем появляется уже на Красной площади, а затем и в Кремле, в кабинете Владимира Путина (слоган, сопровождающий сгенерированные кадры их «беседы»: «Большая страна говорит на одном языке!»)²⁵.

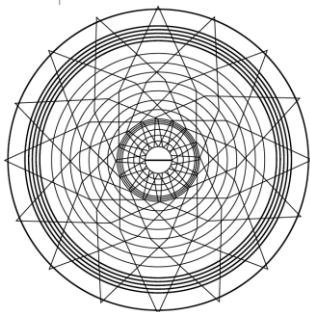
Об эффективности подобной предвыборной коммуникации можно дискутировать, однако для нас важно зафиксировать специфические черты данного вида политического контента. Он видится нам более разнообразным по семантике и нарративам, нежели предыдущие рассмотренные нами варианты дипфейков. Нейросетевая генерация порой показана откровенно, без попыток дезориентировать или обмануть зрителя (а иногда контент содержит и прямое указание на то, что сообщение сгенерировано искусственным интеллектом).

²³ См. об этом, например: Алфимов И. (2021, 27 августа). Лавров и Шойгу у вас дома. В Сети появился ролик о выборах с дипфейком. АиФ.ру.

https://aif.ru/politics/russia/lavrov_i_shoygu_u_vas_doma_v_seti_poyavilsya_rolik_o_vyborah_s_deepfake

²⁴ См: Телеграм-канал Кынев (2025, 17 августа). https://t.me/alexander_kynev/15180

²⁵ См.: Телеграм-канал Кынев (2025, 30 августа). https://t.me/alexander_kynev/15481



[Научные статьи]

Шомова С. А.

«Ах, обмануть меня не трудно...»: феномен политического дипфейка в коммуникативных практиках Рунета

Вместо единственного повествователя, излагающего свое монотонное сообщение соло — «говорящей головы», — в кадре фото- или видеообъектива появляются многофигурные композиции (Трамп с темнокожими избирателями, Лавров с коллегами или глава Республики Коми с медведем), а в визуальном сюжете просматриваются завязка, кульминация и развязка. Предвыборный дипфейк разноформатен и по коммуникационному жанру: это может быть рекламный ролик, информационный месседж, маркетинговая коммуникация, а может быть и дезинформационный контент, манипулирующий избирателем в тех или иных целях. Однако именно потому, что предвыборный дипфейк столь разнообразен и многопланов, он способен содержать и много маркеров, указывающих на искусственную природу даже очень реалистичного на первый взгляд видео или фото — от элементов галлюцинаций нейросетей (визуальных артефактов) и признаков генерации аудио до искажения пропорций, несоответствия движений губ звуку, неестественности мимики и пр.

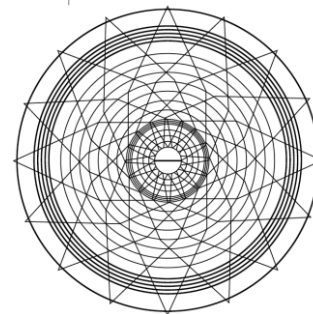
На этом перечисление основных типов политического дипфейка в рунете можно было бы считать завершенным, однако не хотелось бы обойти вниманием еще одну разновидность данного контента. Этот тип, наиболее «безобидный» из всех, можно условно обозначить как досуговый, *рекреационный дипфейк*. Он носит развлекательный характер и может выглядеть как сатира или пародия, а может — как социальная реклама, культурно-просветительское видео, музыкальный клип, послание-поздравление к праздничной дате и т. д. В качестве примера приведем распространившийся в российских социальных сетях весной 2025 г. ролик, в котором романтизированные образы улыбающихся политиков — В. Путина, Д. Трампа, А. Лукашенко, Си Цзиньпина, Н. Моди, Ж. Болсонару и других — под песню «Сегодня праздник у девчат», шествуют навстречу зрителю с цветами в руках ²⁶. Синтезированная сущность и нарочитая «красивость» этого поздравления не пытаются мимикрировать под подлинную съемку: политический месседж (ролик использовал образы правителей «дружественных» государств) и использование дипфейк-технологий здесь очевидны, но в то же время нет грубого намерения ввести зрителя в заблуждение — мы имеем дело с явным дипфейком без дезинформационных целей, что не исключает, конечно, возможности его принятия на веру определенной частью аудитории. При этом его активно распространяли в социальных сетях, пересылали друг другу как своего рода «поздравительную открытку», то есть, выполняя рекреационную функцию («для

²⁶ См.: Президенты поздравляют женщин в клипе ИИ. (2025). VK видео. https://vk.com/video-213439268_456253284

[Научные статьи]

Шомова С. А.

«Ах, обмануть меня не трудно...»: феномен политического дипфейка в коммуникативных практиках Рунета



души»), ролик успешно решал и политическую задачу, репрезентируя пользователям Сети «политику с человеческим лицом». Можно резюмировать, что дипфейк-контент сегодня выходит практически на безграничный семантический и нарративный простор: мы видим ролики и клипы, сгенерированные нейросетями, на самые разные темы — связанные с «мягкой силой» и культурным продвижением, государственной образовательной политикой, социальной работой и многими иными вопросами. Творчество искусственного интеллекта на пересечении медиа и политики (безусловно, способное обернуться большими рисками и нуждающееся во внимательном правовом и этическом регулировании) продолжается и еще способно преподнести немало сюрпризов.

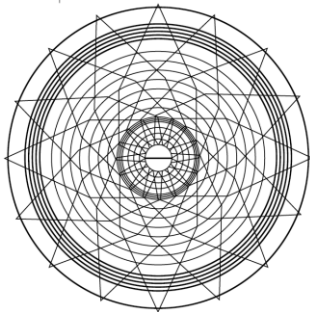
В заключение еще раз напомним, что пространство современного политического дипфейка — открытая система; в наши планы не входила нереалистичная задача обрисовать все возможные варианты интересующего нас контента: важно было лишь зафиксировать рабочее определение данного феномена и предложить новые подходы к его классификации. Разумеется, описанным нами кругом проблем академическая работа по осмыслению политического дипфейка не ограничивается; весьма перспективным направлением развития такого знания (не нашедшим отражения в настоящей статье, но естественным образом продолжающим ее) могло бы стать, например, изучение феномена «дивиденда лжеца» — ситуации, когда в эпоху дипфейков людям (в том числе политикам, чиновникам, лидерам общественного мнения) становится все проще утверждать, что то или иное видео, фото, аудиозапись являются недостоверными. В эпоху постправды это размывает границы лжи и истины и становится новым вызовом в информационном пространстве.

Но эта тема, безусловно, требует уже отдельного анализа и представляет собой предмет для дальнейших исследований.

БИБЛИОГРАФИЯ

Виноградова, Е. А. (2023). Злонамеренное использование политических дипфейков и попытки их нейтрализации в странах Латинской Америки. *Латинская Америка*, 5, 35–48. <https://latamerica-journal.ru/s0044748x0025404-3-1/>

Виноградова, Е. А. (2024). Потенциальные угрозы несанкционированного использования политических дипфейков в период политических выборов: международный опыт. *Мировая политика*, 3, 44–60. https://enotabene.ru/wi/article_71519.html



[Научные статьи]

Шомова С. А.

«Ах, обмануть меня не трудно...»: феномен политического дипфейка в коммуникативных практиках Рунета

Воронин, И. А., Гавра, Д. П. (2024). Дипфейки: современное понимание, подходы к определению, характеристики, проблемы и перспективы. Российская школа связей с общественностью, 33. https://ruscoms.ru/images/archive/vipusk_33.pdf

Иванов, В. Г., Игнатовский, Я. Р. (2020). Deepfakes: перспективы применения в политике и угрозы для личности и национальной безопасности. Вестник РУДН. Серия: Государственное и муниципальное управление, 7(4), 379–386. <https://doi.org/10.22363/2312-8313-2020-7-4-379-386>

Красовская, Н. Р., Гуляев, А. А. (2020). Технологии манипуляции сознанием при использовании дипфейков как инструмента информационной войны в политической сфере. Власть, 28(4), 93–98. <https://doi.org/10.31171/vlast.v28i4.7440>

Цветкова, Н. А., Гришанина, Т. А. (2023). «Цифровая сила» в странах Латинской Америки: хэштеги и дипфейки как политические инструменты. Латинская Америка, 3, 21–33. <https://latamerica-journal.ru/s0044748x0024414-4-1/>

Шомова, С. А. (2023). От конспирологии до розыгрыша: проблема мис- и дезинформации в пандемийном дискурсе. Коммуникации. Медиа. Дизайн, 8(1), 5–23. <https://doi.org/10.17323/cmd.2023.16694>

Appel, M., & Prietzel, F. (2022). The detection of political deepfakes. Journal of Computer-Mediated Communication, 27(4), zmac008. <https://doi.org/10.1093/jcmc/zmac008>

Barari, S., Lucas, C., & Munger, K. (2021). Political Deepfakes Are As Credible As Other FakeMedia And (Sometimes) Real Media. [Preprint]. OSF. <https://doi.org/10.31219/osf.io/cdfh3>

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, 623–643. <https://doi.org/10.1007/s13347-020-00419-2>

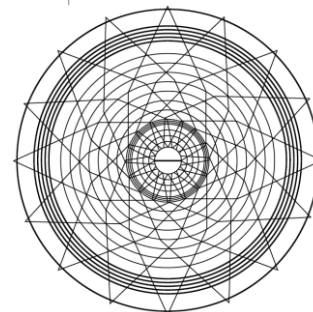
Giansiracusa, N. (2021). How algorithms create and prevent fake news: Exploring the impacts of social media, deepfakes, GPT-3, and more. Apress. <https://doi.org/10.1007/978-1-4842-7155-2>

Gil, R., Virgili-Gomà, J., López, D., & García, R. (2023). Deepfakes: evolution and trends. Soft Computing, 27, 11295–11318. <https://doi.org/10.1007/s00500-023-08605-y>

[Научные статьи]

Шомова С. А.

«Ах, обмануть меня не трудно...»: феномен политического дипфейка в коммуникативных практиках Рунета



Graber-Mitchell, N. (2021). Artificial illusions: Deepfakes as speech. *Intersect: The Stanford Journal of Science, Technology, and Society*, 14(3).

<https://ojs.stanford.edu/ojs/index.php/intersect/article/view/1925>

Hameleers, M., & Marquart, F. (2023). It's nothing but a deepfake! The effects of misinformation and deepfake labels delegitimizing an authentic political speech. *International Journal of Communication*, 17, 6291–6311.

Hameleers, M., van der Meer, T. G. L. A., & Dobber, T. (2022). You won't believe what they just said! The effects of political deepfakes embedded as vox populi on social media. *Social Media + Society*, 8(3). <https://doi.org/10.1177/20563051221116346>

Kietzmann, J., Lee, L. W., McCarthy, I. P., & Kietzmann, T. C. (2020). Deepfakes: Trick or treat? *Business Horizons*, 63(2), 135–146. <https://doi.org/10.1016/j.bushor.2019.11.006>

Köbis, N., Dolezalova, B., & Soraperra, I. (2021). Fooled twice: People cannot detect deepfakes but think they can. *iScience*, 24(11), 103364. <https://doi.org/10.1016/j.isci.2021.103364>

Matern, F., Riess, C., & Stamminger, M. (2019). Exploiting visual artifacts to expose deepfakes and face manipulations. 2019 IEEE Winter Applications of Computer Vision Workshops (WACVW), 83–92. <https://doi.org/10.1109/WACVW.2019.00020>

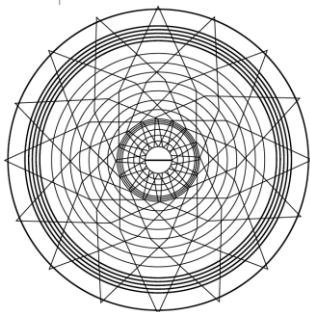
Mirsky, Y., & Lee, W. (2021). The creation and detection of deepfakes: A survey. *ACM Computing Surveys*, 54(1), 1–41. <https://doi.org/10.1145/3425780>

Ray, A. (2021). Disinformation, deepfakes and democracies: The need for legislative reform. *University of New South Wales Law Journal*, 44(3), 983–1016. <https://doi.org/10.53637/DELS2700>

Sandotra, N., & Arora, B. (2024). A comprehensive evaluation of feature-based AI techniques for deepfake detection. *Neural Computing and Applications*, 36, 3859–3887. <https://doi.org/10.1007/s00521-023-09288-0>

Tolosana, R., Romero-Tapiador, S., Fierrez, J., & Vera-Rodriguez, R. (2021). DeepFakes evolution: Analysis of facial regions and fake detection performance. In A. Del Bimbo et al. (Eds.), *Pattern recognition. ICPR international workshops and challenges (Lecture Notes in Computer Science, Vol. 12665, pp. 442–456)*. Springer. https://doi.org/10.1007/978-3-030-68821-9_38

Vasist, P. N., & Krishnan, S. (2022). Deepfakes: An integrative review of the literature and an agenda for future research. *Communications of the Association for Information Systems*, 51, 590–636. <https://doi.org/10.17705/1CAIS.05126>



[Научные статьи]

Шомова С. А.

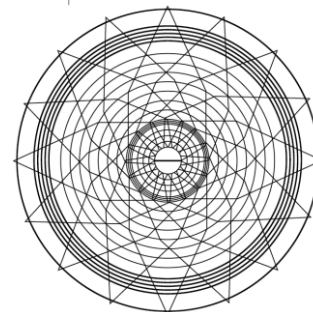
«Ах, обмануть меня не трудно...»: феномен политического дипфейка в коммуникативных практиках Рунета

Wahl-Jorgensen, K., & Carlson, M. (2021). Conjecturing fearful futures: Journalistic discourses on deepfakes. *Journalism Practice*, 15(6), 803–820.

<https://doi.org/10.1080/17512786.2021.1908838>

Westerlund, M. (2019). The emergence of deepfake technology: A review. *Technology Innovation Management Review*, 9(11), 39–52. <https://doi.org/10.22215/timreview/1282>

Yang, X., Li, Y., & Lyu, S. (2019). Exposing deep fakes using inconsistent head poses. *ICASSP 2019 - 2019 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 8261–8265. <https://doi.org/10.1109/ICASSP.2019.8683164>



"AH, IT'S SO EASY TO DECEIVE ME...": THE PHENOMENON OF POLITICAL DEEPPAKE IN RUNET COMMUNICATION PRACTICES

Shomova S. A.

Doctor of Political Sciences, Professor

HSE University

(Moscow, Russia)

sshomova@hse.ru

Abstract:

Political deepfakes are a recent addition to our everyday communication landscape, rapidly expanding into the modern media sphere. They offer significant opportunities but also pose serious risks. This article delves into the academic discourse surrounding this phenomenon, exploring key theoretical frameworks for understanding deepfakes (such as taxonomies, classification systems, and verification methods). It proposes a working definition of political deepfakes and identifies their primary forms within the Russian internet space.

Keywords: artificial intelligence, deepfake, political deepfake

REFERENCES

Appel, M., & Prietzel, F. (2022). The detection of political deepfakes. *Journal of Computer-Mediated Communication*, 27(4), zmac008.

<https://doi.org/10.1093/jcmc/zmac008>

Barari, S., Lucas, C., & Munger, K. (2021). Political Deepfakes Are As Credible As Other FakeMedia And (Sometimes) Real Media. [Preprint]. OSF.

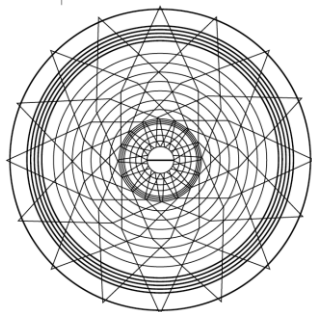
<https://doi.org/10.31219/osf.io/cdfh3>

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, 623–643. <https://doi.org/10.1007/s13347-020-00419-2>

Giansiracusa, N. (2021). How algorithms create and prevent fake news: Exploring the impacts of social media, deepfakes, GPT-3, and more. Apress.

<https://doi.org/10.1007/978-1-4842-7155-1>



Gil, R., Virgili-Gomà, J., López, D., & García, R. (2023). Deepfakes: evolution and trends. *Soft Computing*, 27, 11295–11318. <https://doi.org/10.1007/s00500-023-08605-y>

Graber-Mitchell, N. (2021). Artificial illusions: Deepfakes as speech. *Intersect: The Stanford Journal of Science, Technology, and Society*, 14(3). <https://ojs.stanford.edu/ojs/index.php/intersect/article/view/1925>

Hameleers, M., & Marquart, F. (2023). It's nothing but a deepfake! The effects of misinformation and deepfake labels delegitimizing an authentic political speech. *International Journal of Communication*, 17, 6291–6311.

Hameleers, M., van der Meer, T. G. L. A., & Dobber, T. (2022). You won't believe what they just said! The effects of political deepfakes embedded as vox populi on social media. *Social Media + Society*, 8(3). <https://doi.org/10.1177/20563051221116346>

Ivanov, V. G., & Ignatovskiy, Ya. R. (2020). Deepfakes: perspektivy primeneniya v politike i ugrozy dlya lichnosti i natsionalnoy bezopasnosti. *Vestnik RUDN. Seriya: Gosudarstvennoe i municipalnoe upravlenie*, 7(4), 379–386. <https://doi.org/10.22363/2312-8313-2020-7-4-379-386>

Kietzmann, J., Lee, L. W., McCarthy, I. P., & Kietzmann, T. C. (2020). Deepfakes: Trick or treat? *Business Horizons*, 63(2), 135–146. <https://doi.org/10.1016/j.bushor.2019.11.006>

Köbis, N., Dolezalova, B., & Soraperra, I. (2021). Fooled twice: People cannot detect deepfakes but think they can. *iScience*, 24(11), 103364. <https://doi.org/10.1016/j.isci.2021.103364>

Krasovskaya, N. R., & Gulyaev, A. A. (2020). Tekhnologii manipulatsii soznaniem pri ispolzovanii dipfeykov kak instrumenta informatsionnoy voyny v politicheskoy sfere. *Vlast*, 28(4), 93–98. <https://doi.org/10.31171/vlast.v28i4.7440>

Matern, F., Riess, C., & Stamminger, M. (2019). Exploiting visual artifacts to expose deepfakes and face manipulations. *2019 IEEE Winter Applications of Computer Vision Workshops (WACVW)*, 83–92. <https://doi.org/10.1109/WACVW.2019.00020>

Mirsky, Y., & Lee, W. (2021). The creation and detection of deepfakes: A survey. *ACM Computing Surveys*, 54(1), 1–41. <https://doi.org/10.1145/3425780>

Ray, A. (2021). Disinformation, deepfakes and democracies: The need for legislative reform. *University of New South Wales Law Journal*, 44(3), 983–1016. <https://doi.org/10.53637/DELS2700>

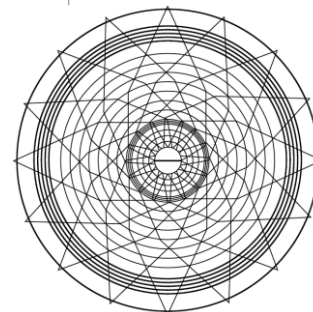
Sandotra, N., & Arora, B. (2024). A comprehensive evaluation of feature-based AI techniques for deepfake detection. *Neural Computing and Applications*, 36, 3859–3887. <https://doi.org/10.1007/s00521-023-09288-0>

Shomova, S. A. (2023). Ot konspirologii do rozygrysha: problema mis- i dezinformatsii v pandemicheskoy diskurse. *Kommunikatsii. Media. Dizayn*, 8(1), 5–23. <https://doi.org/10.17323/cmd.2023.16694>

[Научные статьи]

Шомова С. А.

«Ах, обмануть меня не трудно...»: феномен политического дипфейка в коммуникативных практиках Рунета



Tolosana, R., Romero-Tapiador, S., Fierrez, J., & Vera-Rodriguez, R. (2021). DeepFakes evolution: Analysis of facial regions and fake detection performance. In A. Del Bimbo et al. (Eds.), Pattern recognition. ICPR international workshops and challenges (Lecture Notes in Computer Science, Vol. 12665, pp. 442–456). Springer.
https://doi.org/10.1007/978-3-030-68821-9_38

Tsvetkova, N. A., & Grishanina, T. A. (2023). «Tsifrovaya sila» v stranakh Latinskoy Ameriki: kheshtegi i dipfeyki kak politicheskie instrumenty. Latinskaya Amerika, 3, 21–33. <https://latamerica-journal.ru/s0044748x0024414-4-1/>

Vasist, P. N., & Krishnan, S. (2022). Deepfakes: An integrative review of the literature and an agenda for future research. Communications of the Association for Information Systems, 51, 590–636. <https://doi.org/10.17705/1CAIS.05126>

Vinogradova, E. A. (2023). Zlonamerennoe ispolzovanie politicheskikh dipfeykov i popytki ikh neytralizatsii v stranakh Latinskoy Ameriki. Latinskaya Amerika, 5, 35–48. <https://latamerica-journal.ru/s0044748x0025404-3-1/>

Vinogradova, E. A. (2024). Potentsialnye ugrozy neszaktsionirovannogo ispolzovaniya politicheskikh dipfeykov v period politicheskikh vyborov: mezhdunarodnyy opyt. Mirovaya politika, 3, 44–60. https://e-notabene.ru/wi/article_71519.html

Voronin, I. A., & Gavra, D. P. (2024). Dipfeyki: sovremennoe ponimanie, podkhody k opredeleniyu, kharakteristiki, problemy i perspektivy. Rossiyskaya shkola svyazey s obshchestvom, 33. https://ruscoms.ru/images/archive/vipusk_33.pdf

Wahl-Jorgensen, K., & Carlson, M. (2021). Conjecturing fearful futures: Journalistic discourses on deepfakes. Journalism Practice, 15(6), 803–820.
<https://doi.org/10.1080/17512786.2021.1908838>

Westerlund, M. (2019). The emergence of deepfake technology: A review. Technology Innovation Management Review, 9(11), 39–52. <https://doi.org/10.22215/timreview/1282>

Yang, X., Li, Y., & Lyu, S. (2019). Exposing deep fakes using inconsistent head poses. ICASSP 2019 - 2019 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), 8261–8265. <https://doi.org/10.1109/ICASSP.2019.8683164>