

УДК 343.953

ШЕПИЛОВА М.В. студентка гр. ИПСПД23-03 (СГЮА)
Научный руководитель ДОМНИНА О.В., к.ю.н., доцент (СГЮА)
г. Саратов

**ДИФФЕЙКИ И ЗАЩИТА РЕПУТАЦИИ ИНТЕЛЛЕКТУАЛОВ:
ПСИХОЛОГИЧЕСКОЕ ВОЗДЕЙСТВИЕ И ЮРИДИЧЕСКИЕ СРЕДСТВА
ЗАЩИТЫ ОТ ДЕЗИНФОРМАЦИИ, СОЗДАННОЙ ИСКУССТВЕННЫМ
ИНТЕЛЛЕКТОМ**

В настоящее время происходит активное внедрение информационных технологий во все сферы жизни. Безусловно, они упрощают многие процессы в обществе, однако не всегда цифровизация используется во благо. С развитием информационной сферы появились новые виды преступлений, увеличился уровень мошенничества в сети Интернет – все это говорит о том, что процесс цифровизации как помогает обществу, так и создает серьезную угрозу. Одним из инструментов преступлений в Интернете являются диффейки.

Диффейк (англ. Deepfake, от deep learning – глубокое обучение и fake – подделка) — технология на базе искусственного интеллекта, позволяющая редактировать видео, создавая ложные изображения на основе реальных кадров [1]. Суть заключается в том, что искусственный интеллект (далее по тексту ИИ) «обучается» на большом количестве данных (видео- и аудиозаписей человека), чтобы затем генерировать новый контент. Процесс создания диффейка обычно включает следующие этапы: сбор данных, обучение модели, генерация, обработка и редактирование.

За последние десять с небольшим лет количество этических нарушений, связанных с искусственным интеллектом, выросло в 26 раз [2]. Актуальность работы также подчеркивает отчет Pindrop (платформа для многофакторной защиты от мошенничества и аутентификации в колл-центрах. Она обеспечивает безопасность и достоверность голосовых взаимодействий) о настроениях потребителей в отношении диффейков и голосовых клонов. Потребители, которые осведомлены о диффейках и клонировании голоса, выражают большую озабоченность: около 60% говорят, что они либо «очень», либо «чрезвычайно» обеспокоены, и более 90% выражают некоторую озабоченность. Наибольшие опасения потребителей связаны с областями, в которых конфиденциальная личная информация находится под угрозой и в которых ее подделка может иметь негативные последствия. В связи с этим сферы здравоохранения и страхования значительно опережают банковское дело по уровню обеспокоенности граждан. Когда дело доходит до конкретных проблем с диффейками, на первый план выходит возможность мошенничества, ассоциируемого с кражей денег.[2]

Диффейки напрямую связаны с таким понятием, как дезинформация. Под термином «дезинформация» понимается процесс манипулирования информа-

цией — введение кого-либо в заблуждение путем предоставления неполной информации, искажения контекста [3]. В психологии манипуляция определяется как особый метод психологического влияния. Выделяются осознанные и неосознанные виды манипуляции сознанием. При неосознанном виде манипуляции ко- нечная цель очень расплывчата, а сам манипулятор не задумывается о методах, к которым он прибегает для достижения результата. К осознанным видам относится осознанное воздействие человеком для достижения определенной цели. В качестве примера можно привести информационно-психологическое воздей- ствие (ИПВ) на социальные группы, отдельные лица и сферы психики человека с целью изменения их социально-психологических характеристик. Дезинформи- рование и дипфейки также являются одним из видов ИПВ. Информационно-пси- хологическое воздействие (ИПВ) рассматривается как особая форма проявления социальной активности индивидуального или группового субъекта в рамках ком- муникативного пространства, состоящая в целенаправленном переносе специ- ально подбранной информации в форме знаков (символов), образов или сигна- лов, ориентирующих воспринимаемую систему объекта воздействия относи- тельно смысла и значения данной информации, в результате чего со стороны объ- екта воздействия формируется психоэмоциональный отклик, приводящий к из- менению характера воспринимаемой данным объектом предыдущей и последу- ющей информации, что приводит к корректировке его индивидуальной картины мира в нужном субъекту воздействия ключе [4].

Интеллектуал — это человек, обладающий высоким уровнем образованно- сти, критическим мышлением и активной гражданской позицией, который не только усваивает ценности культуры и науки, но и формирует их, внося вклад в общественное развитие. Он выступает в роли аналитика и критика современных социальных, политических и культурных процессов, способствуя диалогу и рас- пространению идей, направленных на улучшение качества жизни в обществе.

Современный интеллектуал характеризуется не только глубокой эрудицией, но и способностью к самостоятельному осмыслению действительности, чему способствуют навыки междисциплинарного анализа и рефлексия над происходя- щими изменениями. Именно такая личность, как правило, активно участвует в общественных дискуссиях, выступает за защиту гражданских прав, свободы слова и культурного наследия, влияя на формирование правового государства и гражданского самосознания.

В контексте прав интеллектуалов интеллектуал — это не просто носитель знаний, а активный субъект, чьи творческие и критические возможности оказы- вают влияние на развитие правовой и социальной среды общества. Его право на свободное выражение мнений, участие в общественной жизни и формирование культурного диалога является неотъемлемой частью демократии и правового гос- ударства, что подчеркивает необходимость защиты прав данной категории граж- дан в современных условиях.

Интеллектуалы традиционно занимают ведущее место в формировании об- щественного мнения посредством свободного и объективного распространения знаний. Однако новые технологии, такие как дипфейки, могут использоваться для искажения образа интеллектуалов, создавая поддельные высказывания или

демонстрируя их в контексте, противоречащем их реальным убеждениям и деятельности. Такая манипуляция легко может подорвать доверие публичной аудитории, нарушив баланс между объективной информацией и пропагандой, что негативно сказывается на общем уровне информационной безопасности общества.

Особую опасность представляют дипфейки, способные удивительно реалистично имитировать высказывания и поведение интеллектуалов, что может быть использовано как инструмент для дискредитации и маргинализации этих гражданских активистов. Подделанные видео и аудиозаписи могут целенаправленно инсценировать провокационные события или радикальные взгляды, что, в свою очередь, приводит к поляризации общества, усилению конфликтов и подрыву доверия к институтам правопорядка. В условиях стремительного развития цифровых технологий такие манипуляции становятся серьёзной угрозой, требующей своевременного адекватного реагирования со стороны правоохранительных органов и общественных институтов.

Для интеллектуалов, чья работа основана на авторитете и доверии, дипфейки и дезинформация в целом представляют особую угрозу. Поддельное видео или аудио, в котором человек высказывает сомнительные или скандальные заявления, может нанести непоправимый ущерб его репутации, подорвать доверие аудитории и коллег, а также привести к потере профессиональных возможностей. Ущерб может быть особенно значительным, если интеллектуал работает в сферах, требующих высокого уровня доверия, например, в науке, политике или образовании. Даже если удаётся доказать поддельность дипфейка, сам факт обвинения в высказываниях, которых человек не делал, вызывает сильный стресс, тревогу и чувство бессилия. Это может привести к серьёзным психологическим последствиям. Существующие исследования психологического воздействия дезинформации и фейковых новостей показывают, что ложная информация оказывает значительное влияние на когнитивные процессы, эмоции и поведение людей, а также образует долгосрочные последствия для психологического здоровья объекта воздействия дезинформации. К таким можно отнести посттравматическое стрессовое расстройство (ПТСР). Для некоторых интеллектуалов распространение дипфейков может стать травмирующим событием, особенно если поддельные материалы содержат оскорбительный, угрожающий или компрометирующий контент. Постоянное беспокойство о репутации, необходимость публично опровергать ложную информацию и бороться с ее распространением могут привести к развитию ПТСР, характеризующегося тревожностью, бессонницей, кошмарами и избеганием ситуаций, связанных с травмирующим событием. Потеря доверия со стороны коллег, общественности и работодателей, ущерб репутации и профессиональной деятельности могут вызвать депрессию, чувство безнадежности и отчаяния. Постоянное чувство незащищенности и угрозы для репутации может привести к развитию тревожных расстройств, проявляющихся в виде панических атак, избегания публичных выступлений и социальной изоляции. Обвинения в действиях, которых человек не совершил, могут серьезно подорвать самооценку и уверенность в себе. Интеллектуал может начать сомневаться в

своих способностях и компетентности, что негативно повлияет на его профессиональную деятельность.

Интересным нововведением в сфере регулирования ИИ в России является разработка Кодекса этики в сфере ИИ. Кодекс устанавливает общие этические принципы и стандарты поведения, которыми следует руководствоваться участникам отношений в сфере ИИ, и предназначен для создания среды доверенного развития технологий искусственного интеллекта в России, он был принят на I Международном форуме «Этика искусственного интеллекта: начало доверия» в 2021 году [5]. Кодекс распространяется на отношения, связанные с этическими аспектами создания (проектирования, конструирования, пилотирования), внедрения и использования технологий ИИ на всех этапах жизненного цикла, которые в настоящее время не урегулированы законодательством Российской Федерации и/или актами технического регулирования.

В 2024 году была попытка внести изменения в уголовное законодательство: создан законопроект об уголовной ответственности за использование дипфейков. Данный проект предусматривал наложение штрафа в размере до 1,5 млн руб., также возможен арест на срок до трех месяцев или лишение свободы на срок до двух лет. Однако Правительство РФ рекомендовало доработать инициативу. В официальном отзыве кабмин отметил, что для начала нужно урегулировать вопросы использования технологий подмены личности в отраслевом законодательстве, а уже потом вносить изменения в Уголовный кодекс [6]. Кроме того, авторы отзыва рекомендовали собрать статистику, свидетельствующую об опасности дипфейков для граждан. На данный момент, согласно судебной практике, преступления, связанные с использованием дипфейков чаще всего квалифицируются по статье 159.6 Уголовного Кодекса Российской Федерации «Мошенничество в сети компьютерной информации».

В заключение проведенного исследования, посвященного влиянию дипфейков на репутацию интеллектуалов и механизмам правовой защиты от этой угрозы, можно констатировать, что технологический прогресс в сфере искусственного интеллекта создал новые вызовы для общества, требующие адекватного реагирования на всех уровнях – от индивидуального до международного. Распространение дипфейков представляет собой серьезную угрозу не только для репутации публичных персон, но и для стабильности информационного пространства и доверия к источникам информации. Исследование показало, что психологическое воздействие дипфейков на интеллектуалов может быть глубоким и долгосрочным, вызывая стресс, депрессию, снижение самооценки и повреждение профессиональной репутации. Даже после опровержения подлинности дипфейка остаточный негативный эффект может сохраняться длительное время, что требует разработки стратегий психологической поддержки и реабилитации жертв дипфейк-атак.

Анализ существующего законодательства выявил существенные пробелы в механизмах правовой защиты. Действующие нормы о защите чести, достоинства и деловой репутации, а также о клевете и диффамации, часто оказываются недостаточно эффективными для борьбы с высококачественными дипфейками. Сложности в установлении вины, доказательстве ущерба и определении юрисдикции

затрудняют привлечение виновных к ответственности. Для обеспечения устойчивости общественной безопасности и сохранения целостности правового государства необходимо разработать и внедрить специальные механизмы контроля и проверки достоверности информации в цифровом пространстве. Усиление правового регулирования в сфере синтетических медиа, а также повышение цифровой грамотности населения и систематическое взаимодействие государства с технологическими компаниями позволяют не только защитить репутацию интеллектуалов, но и минимизировать негативные последствия дипфейков для общественного порядка. Таким образом, борьба с дезинформацией становится неотъемлемой частью комплексной стратегии обеспечения безопасности и стабильности в современной информационной среде.

Список литературы:

1. Научно-образовательный портал «Большая российская энциклопедия» [Электронный ресурс] – свободная энциклопедия -/ Большая российская энциклопедия – Электронные данные. Режим доступа: URL: <https://bigenc.ru/c/dipfeik-f9f89b> , свободный (дата обращения 4.03.2025)
2. Bret Kinsella, Andrew Herndon. Deepfake and Voice Clone Consumer Sentiment Report 2023, Voicebot Research/ 2023. С.3 – 4
3. Ковалева Ю.В., Соснин В.А. Психоисторическое противостояние Запада и России в XXI веке: социокультурные и социально-психологические детерминанты // Институт психологии Российской академии наук. Социальная и экономическая психология. 2017. Т. 2. № 1. С. 119-142.
4. Долгов М.И. Классификация информационно-психологических воздействий и оценка степени их угроз для общественного сознания // Общество: политика, экономика, право. Краснодар: ХОРС, 2016. С. 47-52.
5. Искусственный интеллект Российской Федерации [Электронный ресурс] – свободный портал – Электронные данные. Режим доступа: URL: <https://ai.gov.ru/ai/regulatory/> , свободный (дата обращения 4.03.2025 г.)
6. Интернет- портал «Российской газеты»// ЭЛ № ФС 77 — 50379. [Электронный ресурс] – Электронные данные. Режим доступа: URL: <https://rg.ru/2024/09/16/v-gosdumu-vnesen-proekt-ob-ugolovnom-nakazanii-za-ispolzovanie-dipfejkov.html> , свободный (дата обращения 4.03.2025 г.)