**Технологии искусственного интеллекта в социальной и политической сфере:
основные проблемы применения.**

Научный руководитель – Бочанов Максим Александрович

Ситдиков Федор Алиевич

Студент (бакалавр)

Финансовый университет, Институт онлайн-образования, Москва, Россия

E-mail: fedorolga2004@gmail.com

Технологии Искусственного Интеллекта (ИИ) в наши дни стремительно проникают в самые различные сферы общественной жизни, упрощая поиск и агрегирование информации, написание простых текстов и создание иллюстраций. Однако, повсеместное распространение систем ИИ, закрывая определённый спектр пользовательских потребностей, также создаёт и множество новых, в том числе и политических, проблем.

Так, например, большую проблему представляет собой отсутствие прозрачности нейросетевых алгоритмов. Большинство современных моделей ИИ, подобно многим иным системам, работают по принципу ввода и вывода определённых данных. В соответствии с моделью «черного ящика», при выполнении действия с системой пользователь вносит в неё определённые данные, а на выходе получает конкретный результат, в той или иной мере соответствующий его ожиданиям. В то же время, механизмы достижения данного результата остаются вне доступности пользователя. [2]

Следует отметить, что высокая структурная сложность моделей ИИ и множественность используемых переменных препятствует проведению объективного анализа действий ИИ и своевременному обнаружению ошибок в работе алгоритма. Риски некорректного поведения системы ИИ значительно возрастают, когда речь идёт не о решении повседневных рутинных задач, а о таких сложных процессах, как принятие политических решений или управление транспортным средством. [1]

Кроме того, существенной политической проблемой становится массовое применение систем ИИ при распространении дезинформации и создании фейковых новостей. Так, широкое распространение получило создание так называемых дипфейков. Данный термин происходит от английского deep learning (глубинное обучение) и fake (подделка). Под дипфейком следует понимать технологии изготовления поддельных фото и видеозаписей, которые используют компьютерный синтез изображения, при помощи искусственного интеллекта перенося черты лица человека на целевое изображение с высокой реалистичностью. [3]

Данная технология позволяет создавать графические материалы, несущие угрозу для репутации тех или иных общественных деятелей, представляя их в невыгодном свете и вкладывая в их уста фразы, которые они никогда и не думали произносить. Опасность использования дипфейков для массовых манипуляций проявляется ещё и в том, что в связи со значительной мощностью информационных потоков, сами дипфейки, как правило, содержащие резонансные и компрометирующие подробности, могут получить гораздо большее представление в медиасфере, чем их неизбежное разоблачение. [5]

Еще одной проблемой, которую хотелось бы осветить в данной работе и которая отчасти связана с предыдущими, являются особенности функционирования алгоритмических рекомендаций в сервисах популярных социальных платформ. Социальные платформы, принадлежащие крупным транснациональным корпорациям (ТНК), используя собранную информацию о предшествующей активности пользователя, формируют рекомендательные

алгоритмы, показывающие пользователям контент, которые может их заинтересовать. Рекомендательные алгоритмы связывают пользователей и создателей контента, способствуя продвижению медиаматериалов, вызывающих наибольшую вовлечённость у аудитории. [4]

Однако, следует понимать, что информация, тиражируемая при помощи алгоритмов может влиять и на политическое поведение граждан, в то время как основным интересом создателей данных платформ является максимизация собственной прибыли. В связи с этим, крайне высок риск использования рекомендательных алгоритмов для продвижения агитационных материалов определённых кандидатов, в ущерб остальным, что ведёт к росту политической субъектности ТНК, их включённости в политические процессы в самых разных государствах с целью продвижения ими собственных интересов. [4]

Следует отметить, что эти проблемы не перечёркивают всех преимуществ ИИ, обеспечивающего лёгкость и эффективность создания и распространения информационных материалов, а для минимизации их влияния на нашу социальную и политическую жизнь, на мой взгляд, могут быть уместны следующие решения.

Во-первых, после получения информации из генеративных нейросетей следует подвергать её тщательной проверке. Кроме того, можно сопоставлять ответы разных нейросетей для более глубокого понимания заданного вопроса. Во-вторых, при потреблении медиаконтента, следует обращать внимание на несостыковки и расхождения в представленных файлах с целью выявления фейковых материалов. Кроме того, можно использовать специализированные ресурсы для проверки медиаконтента на наличие искусственных модификаций. В-третьих, при использовании социальных сетей, следует отдавать предпочтение хронологической подаче материалов, изменив настройки ленты в приложении. Кроме того, можно изучить настройки конфиденциальности таких сервисов, отключив параметры сбора персональных данных, доступные для изменения.

Подводя итог, следует подчеркнуть, что системы ИИ представляют собой очень удобный и быстрый инструмент, способный существенно обогатить наши социальные практики, имеющий однако и свои недостатки.

Источники и литература

- 1) Володенков С. В., Федорченко С. Н. Риски применения алгоритмов искусственного интеллекта в социально-политической сфере: обзор современных научных работ / С. В. Володенков, С. Н. Федорченко // Дискурс-Пи. – 2024. – Т. 21, № 2. – С. 24-48. – DOI 10.17506/18179568_2024_21_2_24. – EDN DNFZOW.
- 2) Володенков С. В., Федорченко С. Н., Печенкин Н. М. Риски, угрозы и вызовы внедрения искусственного интеллекта и нейросетевых алгоритмов в современную систему социально-политических коммуникаций: по материалам экспертного исследования / С. В. Володенков, С. Н. Федорченко, Н. М. Печенкин // Вестник Российского университета дружбы народов. Серия: Политология. – 2024. – Т. 26, № 2. – С. 406-424. – DOI 10.22363/2313-1438-2024-26-2-406-424. – EDN LWSYCV.
- 3) Уртаева Э. Б. Возможности и угрозы применения искусственного интеллекта (ИИ) в политических коммуникациях / Э. Б. Уртаева // Общество: политика, экономика, право. – 2024. – № 2(127). – С. 44-51. – DOI 10.24158/per.2024.2.3. – EDN UALQLW.
- 4) Чебан А. А. Политическая субъектность транснациональных цифровых корпораций / А. А. Чебан // Вестник Московского государственного областного университета. – 2023. – № 1. – DOI 10.18384/2224-0209-2023-1-1235. – EDN KQFSOG.
- 5) Капитонов А., Линдре Ю. Дипфейк: невинная технология для развлечения или угроза современному обществу? URL: <https://russiancouncil.ru/analytics-and-comm>

ents/analytics/dipfeyk-nevinnaya-tehnologiya-dlya-razvlecheniya-ili-ugroza-sovremen
nomu-obshchestvu/]([https://russiancouncil.ru/analytics-and-comments/analytics/dipfe
yk-nevinnaya-tehnologiya-dlya-razvlecheniya-ili-ugroza-sovremennomu-obshchestvu
/](https://russiancouncil.ru/analytics-and-comments/analytics/dipfe
yk-nevinnaya-tehnologiya-dlya-razvlecheniya-ili-ugroza-sovremennomu-obshchestvu
/) (проверено 02.03.2025)