

Секция «24.14 Технологии искусственного интеллекта в современной политике»

## Политико-манипулятивные риски технологий генеративного искусственного интеллекта: обзор исследований

Научный руководитель – Федорченко Сергей Николаевич

Юдинцева А.А.<sup>1</sup>, Лещенко А.А.<sup>2</sup>

1 - Московский государственный университет имени М.В.Ломоносова, Факультет политологии, Кафедра истории и теории политики, Москва, Россия, *E-mail: agn.yudintseva@gmail.com*; 2 - Московский государственный университет имени М.В.Ломоносова, Факультет политологии, Кафедра истории и теории политики, Москва, Россия, *E-mail: alexandr.leschenko@bk.ru*

Генеративный искусственный интеллект радикально меняет характер политической коммуникации и управление общественным мнением. Потoki новостей, рекламы и агитационных сообщений сегодня корректируются, фильтруются не только редакционной политикой каналов, лидерами общественного мнения, но прежде всего системами рекомендаций, обученными на поведенческих данных граждан. Это определяет важность выявления особенностей политико-манипулятивных рисков технологий генеративного искусственного интеллекта (далее – ИИ) на основе анализа конкретных академических работ.

Современная политическая повестка для конкретного пользователя все чаще формируется не открытой публичной дискуссией, а непрозрачными вычислительными процедурами, о манипулятивных социально-политических эффектах которых российские исследователи предупреждают уже не первый год [1, с. 205; 2, с. 406; 3, с. 14].

В российской академической среде роль ИИ в политической коммуникации рассматривается двояко: с одной стороны, как нового ресурса государственно-политического управления, политического прогнозирования, моделирования конфликтов, с другой, - как источника дополнительной уязвимости цифровой публичной сферы [3, с. 14]. В ряде работ акцент делается на рисках смещения баланса между демократическими и технократическими механизмами принятия решений [1, с. 208; 2, с. 408; 4, с. 78; 5, с. 712].

Одновременно усиливается осознание конкретных политико-манипулятивных рисков. Так, С.В. Володенков, С.Н. Федорченко и др. отмечают, что технологии искусственного интеллекта уже встроены в инфраструктуру персонализации контента, анализа Big Data и автоматизированной вычислительной пропаганды [2, с. 414-416]. Это создает прямые риски манипуляции избирательным процессом и подрыва доверия к избирательным институтам. К.Г. Максимов и М.Ю. Шаповалов анализируют вызовы для национальной безопасности, связанные с автоматизацией политической коммуникации и возможностью массовой персонализированной пропаганды [4, с. 78-80].

Риски дезинформации из-за ИИ детально исследованы К.И. Мухиной и А.С. Чернавским. Авторы анализируют применение генеративных моделей Stable Diffusion и Midjourney для конструирования правдоподобного провокационного политического и аудиовизуальных материалов, способных вызывать социальные конфликты [5]. Э.Б. Уртаева указывает на риски дипфейков, влияющих на сокращение электоральной базы, а также риски информационно-безопасности [6].

Особенно заметны исследования, фиксирующие трансформацию практик дезинформации [3, с. 18; 5, с. 718]. В данном контексте знаковым видится эксперимент группы ученых, создавших 1 052 генеративных агента на основании обработки данных респондентов. Агенты продемонстрировали хорошую точность в конструировании цифровых профилей [7]. За счет такой способности имитации человеческого политического поведения генеративные агенты могут создавать иллюзию гражданской активности и конструировать «реальность социального консенсуса». Доказательством этого суждения являются результаты эксперимента, проведенного на модели GPT-3. Исследователи продемонстрировали, что можно

конструировать не просто генеративных агентов, создающих политический контент, но уже реально создавать алгоритмические копии целых социальных групп [8]. Как отмечает С.Н. Федорченко, подобные технологии создают предпосылки для «умного полемоса» - информационных войн нового типа, где искусственный интеллект выступает инструментом целенаправленного воздействия на массовое сознание [3, с. 20].

Таким образом, современные авторы в основном обращают внимание на три ключевых аспекта политико-манипулятивных рисков ИИ: (1) техническую возможность точной имитации социальных групп и персонализированного воздействия; (2) институциональные угрозы демократическим процедурам и избирательным институтам; (3) деградацию публичной сферы через дезинформацию, дипфейки и «умный полемос». Алгоритмическая генерация контента не просто дополняет традиционные каналы политической коммуникации, но и постепенно перестраивает саму логику взаимодействия между властью и аудиторией, делая её менее прозрачной и более зависимой от технологических посредников. Степень этого влияния заметно различается в зависимости от институциональной среды и характера регуляторных механизмов, что указывает на необходимость формирования национальной стратегии ответственного использования алгоритмических систем в политике.

### Источники и литература

- 1) Володенков С.В. Технологии Big Data в современных политических процессах: цифровые вызовы и угрозы // Вестник Томского государственного университета. Философия. Социология. Политология. -2018. - № 44. - С. 205–212.
- 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 // Вестник Российского университета дружбы народов. Серия: Политология. - 2024. - Т. 26, № 2. - С. 406–424.
- 3) Федорченко С.Н. «Умный полемос»: искусственный интеллект в информационных войнах // Журнал Белорусского государственного университета. Философия. Психология. - 2025. - № 2. - С. 14–22.
- 4) Максимов К.Г., Шаповалов М.Ю. Искусственный интеллект в политической коммуникации: вызовы для обеспечения национальной безопасности // Общество: политика, экономика, право. - 2025. - № 12. -С. 77–83.
- 5) Мухина К.И., Чернавский А.С. Технологии искусственного интеллекта в современной культуре и политической сфере общественной жизни (на примере Stable Diffusion и Midjourney) // Вестник науки. -2023. - Т. 3, № 11 (68). - С. 709–721.
- 6) Уртаева Э.Б. Возможности и угрозы применения искусственного интеллекта в политических коммуникациях // Общество: политика, экономика, право. - 2024. - № 2 (127). - С. 44–51.
- 7) Park J.S., O'Brien J.C., Cai C.J., Morris M.R., Liang P., Bernstein M.S. Generative Agent Simulations of 1,000 People // arXiv preprint. -2024. - URL: <https://arxiv.org/abs/2411.10109>.
- 8) Argyle L.P., Busby E.C., Fulda N., Gubler J.R., Rytting C., Wingate D. Out of One, Many: Using Language Models to Simulate Human Samples // Political Analysis. - 2023. - Vol. 31, № 3. - P. 337–351. - DOI: 10.1017/pan.2023.2.