

Секция «Технологии искусственного интеллекта и нейросетевые алгоритмы в современной политике»

## Нейронные сети как инструмент политических исследований: проблемы использования

Научный руководитель – Володенков Сергей Владимирович

*Калинина Екатерина Владимировна*

*Студент (бакалавр)*

Московский государственный университет имени М.В.Ломоносова, Факультет политологии, Кафедра государственной политики, Москва, Россия

*E-mail: kalinina.ekaterina.11@mail.ru*

Прорыв в использовании нейросетей, осуществленный в последние несколько лет, характеризуется (как публицистами, так и исследователями) как революция. Один из ее результатов - меняется понимание того, что такое машина, и в связи с этим - что такое человек. Так, предложенный Дж. Серлем мысленный эксперимент «Китайская комната» фактически отрицает возможность компьютера воспроизводить интеллектуальную деятельность (демонстрируется «прохождение» теста Тьюринга без понимания семантики сообщений). На сегодняшний день генеративные нейросети (семейство ChatGPT, Midjourney и представленная в феврале Sora от OpenAI) создают контент, трудноотличимый от результатов человеческой деятельности (показательна успешная защита диплома, написанного при помощи нейросети, в 2023 г. в РГГУ) а также кроссмодальны - преобразуют текст в картинку, видео и т.п. При этом вопрос о сильном искусственном интеллекте по-прежнему открыт: действительно ли существующие (и возможные в обозримом будущем) системы генеративного ИИ интеллектуальны в человеческом смысле - или дают только имитацию интеллектуальной деятельности, не сопровождающуюся пониманием семантики и в широком смысле сознательным опытом [2]. Второй точки зрения, придерживается руководитель OpenAI С. Альтман. Ее косвенным подтверждением является тот факт, что основной «выигрыш» использования нейросетей по-прежнему состоит в автоматизации рутинной работы (в данном случае - интеллектуальной). Так, компьютер способен написать программный код быстрее человека, знающего язык программирования, однако типично человеческой деятельностью остается промпт-инжиниринг (адаптация запроса для нейросети под техническую задачу) - типичная «интерпретирующая» деятельность.

В политической науке нейросетевая революция затрагивает преимущественно следующие области: анализ общественного мнения, анализ электорального поведения, таргетирование (вплоть до абсолютной индивидуализации) пропагандистских сообщений, разоблачение fake news и политическое прогнозирование и оценка рисков. Интересно, что это множество областей обнаруживается (хотя и не в виде списка) в работе Л. Вордличека (2022 г.) [3], а также в работе В.С. Авдониной и В.Л. Силаевой, где прямо указывается, что список «составила» нейросеть [1].

В целом, исследователи предостерегают как от эйфории по поводу нейросетей, так и от сетевой фобии. Неуниверсальность нейросетей как инструмента в социальных науках сводится к следующим проблемным областям:

1. Зависимость от количества данных - нейросеть должна быть обучена на большой выборке (по данным Л. Вордличека, выборка из менее чем 1000 случаев редко дает правдоподобный результат при прогнозировании политических событий).

2. Зависимость от качества данных - выборка, на которой обучается нейросеть, должна быть предварительно обработана. Без предварительной стандартизации данных пользователь нейросети рискует столкнуться с бессмыслицей на выходе [3].

3. Невоспроизводимость результатов - со времен начала формирования научной методологии в Новое время одним из критериев научности считается воспроизводимость результата. Нейросети, обученные на неодинаковых массивах данных, дают неодинаковые предсказания одних и тех же событий. Так, сети KCore Analytics и Expert.ai успешно предсказали победу Дж. Байдена на последних выборах президента США, тогда как российская нейросеть Promobot предсказывала победу Д. Трампа. Оценки нейросетей дают в итоге не меньший «плюрализм», чем оценки политических аналитиков-людей.

4. Непрозрачность работы - нейросеть представляет собой «черный ящик», в отличие от человека неясно, каким образом она «приходит» к итоговым данным. Безусловно, основу экспертности, по М. Полани, составляют т.н. неявные знания и умения, однако принципиальное отличие состоит в том, что человек способен хотя бы частично перевести собственный багаж неявных знаний в явный вид (в то время как машина этого сделать не способна). Иной аспект непрозрачности работы нейросетей - риск получения «галлюцинаций» (явно ложных ответов, которые не обуславливаются выборкой, на которой обучалась сеть).

В работе В.С. Авдониной и В.Л. Силаевой высказываются 2 сценария нейросетевой революции в рамках политологии - ревизионистский, предполагающий обновление всего теоретического и практического инструментария, и адаптационный, предполагающий сосуществование «старых» теорий и новых инструментов[1]. Перечисленные сложности указывают, скорее, на адаптационный сценарий с необходимым участием человека в качестве «интерпретатора с правом вето». В этом контексте показательна история избирательных кампаний Х. Клинтон и Д. Трампа в США, где активно использовались нейросети. Нейросеть Ada, которая использовалась штабом Клинтон, ежедневно обрабатывала множество сценариев развития предвыборной гонки; штаб Трампа сотрудничал с британской корпорацией Cambridge Analytica, которая занималась персонализацией сообщений. Проигрыш Х. Клинтон привел к тому, что уже в кампаниях Трампа и Байдена 2020 г. нейросети использовались уже только как помощники, бравшие на себя рутинную статистическую работу.

### Источники и литература

- 1) Авдонин В.С., Силаева В.Л. Нейросети нового поколения в контексте технологий искусственного интеллекта, философии и социально-политических наук // Полит. наука. 2023. №4.
- 2) Журавлёв Д.В., Смолин В.С. Нейросетевая революция искусственного интеллекта и варианты её развития // Проектирование будущего. Проблемы цифровой реальности: труды 6-й Международной конференции (2-3 февраля 2023 г., Москва). — М.: ИПМ им. М.В.Келдыша, 2023
- 3) Wordliczek L. Neural Networks and Political Science: Testing the Methodological Frontiers // Empiria. Revista de Metodologia de Ciencias Sociales. №57, 2023