

ПОЛИТИЧЕСКАЯ СОЦИОЛОГИЯ

DOI 10.35775/PSI.2025.122.10.004

УДК 32

Л.В. САДЖАЯ

магистрант факультета политологии
Московского государственного университета
имени М.В. Ломоносова,
Россия, г. Москва
E-mail: luka.sadjaya@yandex.ru

ЦЕННОСТНАЯ ПРЕДВЗЯТОСТЬ В РОССИЙСКИХ СИСТЕМАХ ИСКУССТВЕННОГО ИНТЕЛЛЕКТА: АНАЛИЗ ИДЕОЛОГИЧЕСКИХ РЕПРЕЗЕНТАЦИЙ В НЕЙРОСЕТИ ШЕДЕВРУМ

Разработка национальных систем искусственного интеллекта актуализирует проблему предвзятости в алгоритмах генерации контента. **Целью** исследования является выявление и анализ проявлений ценностной предвзятости, в частности западно-центричности, в идеологических репрезентациях, создаваемых российской нейросетью Шедеврум. Методология основана на многомесячном качественном анализе визуализаций, сгенерированных по запросам на русском и английском языках для ключевых политических идеологий (коммунизм, социализм, анархизм, национализм, фашизм, нацизм, либертарианство, либерализм, консерватизм) в категориях: сущность, будущее, общество, типичный представитель. Результаты выявили устойчивое доминирование западных культурных стереотипов. Установлено, что нейросеть систематически воспроизводит паттерны «западного взгляда»: левые идеологии ассоциируются с упадком и милитаризмом, анархизм – с хаосом, праворадикальные идеологии – с депрессивным разрушением, либеральные течения – с позитивными образами американской государственности, а национализм не получает релевантной визуализации. Это противоречит декларируемому технологическому суверенитету и детерминировано «датаколониализмом» – доминированием западных данных в обучающих выборках. Делается вывод о том, что алгоритмы генеративного ИИ выступают инструментом культурной гегемонии. Для обеспечения подлинного технологического суверенитета России критически необходимо создание репрезентативных русскоязычных датасетов, способных отражать многообразие локальных интерпретаций социально-политических феноменов.

Ключевые слова: ценностная предвзятость, западноцентричность, искусственный интеллект, нейросеть «Шедеврум», технологический суверенитет, «западный взгляд», датаколониализм, политические идеологии.

Теоретические основания исследования идеологических репрезентаций в искусственном интеллекте. Проблема предвзятости алгоритмов активно изучается международным научным сообществом. Важный вклад в понимание этого феномена внесли Улиес Мехиас и Ник Коулдри, предложившие концепцию «датаколониализма» – процесса, при котором преобладание западных данных в обучающих выборках создает форму информационной колонизации через навязывание чуждых культурных моделей [15]. Развитие этих идей представлено в исследовании Сафии Умоджи Нобл, где на основе шестилетнего анализа поисковых алгоритмов Google доказывается, что поисковые системы не нейтральны, а фаворитизируют определенные группы пользователей, усиливая стереотипы [19]. Существуют множество другие исследований, которые демонстрируют, как алгоритмические системы воспроизводят и усиливают существующие социальные предрассудки [12; 14. Р. 1-15; 17; 18. Р. 1222-1223; 13]. Совокупность этих работ формирует понимание того, как технологии искусственного интеллекта становятся инструментами культурной гегемонии, что напрямую связано с западноцентричными репрезентациями в российских нейросетях.

Следует подчеркнуть, что в работах российских авторов, опубликованных в последние годы, освещается широкий спектр вопросов близких к данной предметной области [1; 2; 3; 4; 5; 6; 7; 8; 9; 10; 11].

Однако проблеме проявлений ценностной предвзятости, в частности западноцентричности, в идеологических репрезентациях, создаваемых российской нейросетью Шедеврум нельзя назвать однозначно исчерпанной. В силу многих объективных обстоятельств изучение обозначенной темы продолжает сохранять высокий уровень актуальности.

Методология исследования идеологических репрезентаций в Шедевруме. Целью исследования является выявление и анализ проявлений ценностной предвзятости, в частности западноцентричности, в идеологических репрезентациях, создаваемых российской нейросетью Шедеврум. В ходе исследования тестировались репрезентации девяти основных идеологий (коммунизм, социализм, анархизм, национализм, фашизм, нацизм, либертарианство, либерализм и консерватизм) по четырем категориям: само понятие, образ будущего, общество и типичный представитель данной идеологии. Методология включала многократное повторение идентичных запросов на русском и английском языках в период с ноября по февраль, а для обхода блокировки некоторых идеологий (нацизма и фашизма) применялись модифицированные запросы с изменением отдельных букв. Анализ полученных изображений проводился по ключевым параметрам: цветовая гамма, присутствие специфических символов, архитектурные элементы, гендерные характеристики персонажей и общий эмоциональный тон визуальной репрезентации.

Визуальные репрезентации идеологий. Анализ сгенерированных Шедеврумом изображений выявил отсутствие четкой дифференциации между визуальными изображениями левых идеологий. Нейросеть практически

не различает социализм и коммунизм, что проявляется в идентичности образов будущего и общества при этих идеологиях – они представлены красными ржавыми массивными зданиями с минимальным присутствием людей, создавая впечатление заброшенности и упадка. При этом в социалистической символике доминируют красные розы, отсылающие к европейским социалистическим партиям, тогда как коммунизм визуализируется с ярко выраженными негативными коннотациями – разрушениями, обилием красного цвета, изображениями людей в военной форме. Особенно показательна этническая стереотипизация: представители коммунистической идеологии изображаются исключительно как азиаты (предположительно китайцы) или европейцы (предположительно русские) в красной военной форме, что отражает географическую привязку коммунизма к определенным странам.

Анархизм фактически отождествляется с анархией и постапокалиптическими сценариями. Преобладают серые тона, изображаются пустоши, разруха, одинокие люди с дубинками, напоминающие мародеров. На фоне часто присутствуют заброшенные дома европейской или американской архитектуры. Анархисты представлены либо как бродяги, либо как представляющие угрозу.

Визуализация нацизма и фашизма характеризуется изображениями поездов, людей арийской внешности в военной форме с преобладанием серых тонов, разрушениями или зданиями европейской архитектуры, общая атмосфера депрессивна и пасмурна. Будущее при фашизме представлено разрушениями, а при нацизме своеобразным киберпанком с киборгами и роботами в серых, депрессивных антиутопичных тонах. Что касается национализма, эти запросы нейросеть игнорирует, генерируя вместо релевантных образов случайные изображения.

Консерватизм представлен людьми в костюмах, напоминающими членов английского парламента или американских политиков, включая даже изображение Дональда Трампа, что особенно примечательно, учитывая цензуру нейросети на изображения реальных политических деятелей.

Либертарианство визуализируется через виды природы и зданий, напоминающих Капитолийский холм, с характерным большим желтым солнцем на заднем плане. В отличие от большинства других идеологий, атмосфера этих изображений значительно более позитивна. Либерализм также ассоциируется с американскими государственными институтами, создавая позитивную атмосферу.

Исследование визуализаций идеологий в Шедевруме наглядно демонстрирует доминирование западной перспективы и культурных стереотипов в основе алгоритмов генерации изображений. Этот феномен не уникален для российской нейросети и отражает более широкую проблему так называемого «западного взгляда» («Western gaze») в разработках генеративного искусственного интеллекта [16]. Ученые наблюдают похожие проблемы и у других нейросетей. Например, тестирование Midjourney также показывает, что нейросеть выдает сильно стереотипизированные образы [20].

Закключение. Проведенное исследование визуальных репрезентаций политических идеологий в нейросети Шедеврум выявило значительный уровень западноцентричности в работе данной системы, что указывает на более широкую проблему ценностной предвзятости в российских системах искусственного интеллекта. Решение этой проблемы требует комплексного подхода, включающего технологические меры, прежде всего – создание репрезентативных русскоязычных датасетов для разработки культурно-релевантных нейросетей, отражающих многообразие российских интерпретаций социально-политических феноменов. Важно подчеркнуть, что разработка таких систем не должна сводиться к созданию «идеологически выверенных» моделей; напротив, целью должно стать формирование алгоритмов, способных отражать все разнообразие интерпретаций и точек зрения, существующих в обществе. Дальнейшие исследования в данной области должны концентрироваться на разработке конкретных методов и инструментов для создания культурно-релевантных нейросетей, что будет иметь не только теоретическое, но и практическое значение, способствуя развитию подлинного технологического суверенитета России.

БИБЛИОГРАФИЧЕСКИЙ СПИСОК:

1. Ванеев И.М. Национальная безопасность в условиях идеологического плюрализма // Вопросы политологии. 2025. № 5.
2. Грядунов Ю.Ю. Анализ государственной политики России по укреплению традиционных идеологических ценностей // Вопросы политологии. 2024. № 7.
3. Грядунов Ю.Ю. Проблемы и перспективы интеграции традиционных ценностей в современное общество России // Вопросы национальных и федеративных отношений. 2024. № 7.
4. Кочетков А.П. О стратегии обеспечения геополитической безопасности России в условиях противостояния с коллективным Западом // Вопросы политологии. 2024. № 8.
5. Курбанов Н.И., Ушаков В.В., Цапок А.А. Формирование идеологии России после 2022 года: проблемы и перспективы // Вопросы национальных и федеративных отношений. 2025. № 3.
6. Одинокоев М.Д., Трифонов А.Д. Роль консервативных институтов гражданского общества и партийной системы в консолидации российского общества в условиях нестабильности мирового политического пространства // Евразийский Союз: вопросы международных отношений. 2023. № 7.
7. Проценко А.В., Бerviно Е.Г. Охранительная функция умеренного консерватизма в условиях геополитических трансформаций // Вопросы национальных и федеративных отношений. 2024. № 10.
8. Рузиева Г.С. Историко-политические корни взаимосвязи культуры и национальной безопасности // Вопросы политологии. 2024. № 7.

9. **Титов В.В.** Политика памяти в России: проблема институционализации и стратегического управления // Вопросы политологии. 2025. № 2.
10. **Фурман Т.Г.** Противоречия технологического общества как политико-философский конструкт национальной идеи современной России // Вопросы национальных и федеративных отношений. 2024. № 2.
11. **Щербаченко А.К., Отарашвили А.А., Григорян Д.К.** Подъем национального единства в период Специальной военной операции как формирующий фактор национальной безопасности // Евразийский Союз: вопросы международных отношений. 2025. № 2.
12. **Basu A., Babu R.V. & Pruthi D.** (2023). Inspecting the Geographical Representativeness of Images from Text-to-Image Models. Indian Institute of Science, Bangalore.
13. Biases in large image-text AI model favor wealthier, Western perspectives // University of Michigan News. 2023 // <https://news.engin.umich.edu/2023/12/biases-in-large-image-text-ai-model-favor-wealthier-western-perspectives/>.
14. **Buolamwini J., Gebru T.** (2018). Gender Shades: Intersectional Accuracy Disparities in Commercial Gender Classification. Proceedings of the 1st Conference on Fairness, Accountability and Transparency, PMLR 81.
15. **Couldry N., Mejias U.A.** (2019). The Costs of Connection: How Data Is Colonizing Human Life and Appropriating It for Capitalism. Stanford University Press.
16. **Gautam S., Venkit P.N., Ghosh S.** (2024). From Melting Pots to Misrepresentations: Exploring Harms in Generative AI.
17. **Hagerty A., Rubinov I.** (2019). Global AI Ethics: A Review of the Social Impacts and Ethical Implications of Artificial Intelligence.
18. **Kidd C., Birhane A.** (2023). How AI Can Distort Human Beliefs. Science, 380 (6651).
19. **Noble S.U.** (2018). Algorithms of Oppression: How Search Engines Reinforce Racism. New York: NYU Press, 2018.
20. Rest of World. 2023 // <https://restofworld.org/2023/ai-image-stereotypes/>.

L.V. SADZHAYA

Master student of the Faculty
of Political Science, Lomonosov Moscow
State University, Moscow, Russia

VALUE BIAS IN RUSSIAN AI SYSTEMS: ANALYSIS OF IDEOLOGICAL REPRESENTATIONS IN THE SHEDEVNUM NEURAL NETWORK

The development of national artificial intelligence systems brings to the forefront the problem of bias in content generation algorithms. The aim of the study is to identify and analyze manifestations of value bias, particularly Western-centricity, in ideological representations created by the Russian neural network Shedevrum. The methodology is based on a months-long qualitative analysis of visualizations generated via Russian and English prompts for key political ideologies (communism, socialism, anarchism, nationalism, fascism, Nazism, libertarianism, liberalism, conservatism) across categories: essence, future, society, typical representative. Results revealed a persistent dominance of Western cultural stereotypes. It was established that the neural network systematically reproduces patterns of the «Western gaze»: left-wing ideologies are associated with decline and militarism, anarchism with chaos, right-wing radical ideologies with depressive destruction, liberal movements with positive imagery of American statehood, while nationalism lacks relevant visualization. This contradicts the declared technological sovereignty and is determined by «data colonialism» – the dominance of Western data in training datasets. The conclusion states that generative AI algorithms act as tools of cultural hegemony. To ensure Russia's genuine technological sovereignty, it is critically necessary to create representative Russian-language datasets capable of reflecting the diversity of local interpretations of socio-political phenomena.

Key words: cultural bias, Western-centricity, artificial intelligence, Shedevrum, technological sovereignty, «Western gaze», data colonialism, political ideologies.