

УДК 327.2; 81`42; 004.8 ББК Ф052; С028.132; Ф4в; Ш100.3 ГРНТИ 28.23.21;
11.25.25; 16.21.27; 11.15.83 ВАК: 5.5.2

DOI: 10.24412/2949-2513-2026-40-171-191

О.А.Байша, В.В.Трофимов

Москва, Россия

ОБ ИДЕОЛОГИЧЕСКИХ ПРЕДВЗЯТОСТЯХ ГЕНЕРАТИВНОГО ИИ: РОССИЙСКО-УКРАИНСКИЙ КОНФЛИКТ В РЕПРЕЗЕНТАЦИИ CHATGPT

Аннотация. Все большее количество ученых предупреждает об опасности воспроизведения генеративным ИИ социально-политических и идеологических предрассудков, впитанных моделями из текстов, на которых они обучались. Если, например, та или иная модель тренировалась на материалах западных СМИ, она может генерировать нарративы, воспроизводящие западнцентричный взгляд на мировые события. Это проявляется в репродукции определений глобальных проблем, нормализованных в западных гегемонистских дискурсах. С целью выявления, как именно воспроизводятся идеологические предубеждения ГИИ по отношению к России и как их можно обнаружить в сгенерированных текстах, в данной статье исследуется аналитическая записка, созданная ChatGPT – чат-ботом, разработанным американской компанией OpenAI. Заданной темой генерации был российско-украинский конфликт, начало которому положил (Евро)Майдан – протестное движение на Украине (2013–2014), приведшее к смене правительства. Для анализа сгенерированного текста использовались два аналитических метода, разработанных на основе дискурсивных теорий Эрнесто Лаклау и Шанталь Муфф, а также Нормана Фэркло. Результаты анализа показали, что ChatGPT, несмотря на обещание нейтральности, сгенерировал черно-белое повествование, упускающие все важные нюансы, которые могли бы помочь его пользователям разобраться

в ситуации. Моделью был воссоздан негативный образ России, который был нормализован в западном политическом дискурсе задолго до пришествия ГИИ в повседневную жизнь. Результаты исследования подтвердили опасения ученых о том, что тексты, генерируемые ИИ, могут демонстрировать идеологическую предвзятость и воспроизводить идеологию западнцентризма, делая это с претензией на авторитетность и нейтральность.

Ключевые слова: генеративный ИИ, большие языковые модели, ChatGPT, предвзятости ИИ, западнцентризм, Евромайдан, российско-украинский конфликт

Olga A. Baysha, Vladislav V. Trofimov

Moscow, Russia

ON THE IDEOLOGICAL BIASES OF GENERATIVE AI: THE RUSSIAN-UKRAINIAN CONFLICT AS PRESENTED BY CHATGPT

Abstract. A growing number of scholars are warning about the dangers of the reproduction by generative AI of socio-political and ideological biases absorbed by models from the texts on which they were trained. If a given model was trained on Western media texts, it may generate narratives that reproduce West-centric views of world events. This will manifest itself in the reproduction of definitions of global problems normalized in Western hegemonic discourses. To identify how AI reproduces ideological biases towards Russia, and how they can be detected in generated texts, this article examines an analytical report generated by ChatGPT, a chatbot developed by the US company OpenAI. The topic was the Ukrainian-Russian conflict, which began with the (Euro)Maidan – a protest movement in Ukraine (2013–2014) that led to the overthrow of the government. Text analysis drew on the discourse theories of Ernesto Laclau and Chantal Mouffe, as well as Norman Fairclough. Its results showed that ChatGPT, despite its promise of neutrality, generated a black-and-white narrative, which missed all the important

nuances that could help its users understand the situation. The model replicated a negative image of Russia that had been normalized in Western political discourse long before the advent of the Generative AI in our everyday lives. This confirmed the concerns of other researchers that AI-generated texts could exhibit ideological bias and reproduce West-centrism, while doing so with a claim to authority and neutrality.

Keywords: generative AI; large language models; ChatGPT; AI biases; West-centrism; Euromaidan; Russia-Ukraine conflict

Введение

В 1980 году комиссия по изучению проблем глобальных коммуникаций при ЮНЕСКО представила доклад, в котором подчеркивалось, что контроль над информационными технологиями со стороны западных стран позволяет им навязывать миру свои формулировки глобальных проблем, устанавливая тем самым режим медиа-империализма. Запад во главе с США отверг предложения комиссии по устранению коммуникационного неравенства как противоречащие принципу «свободного потока информации» [Thussu, 2018]. Вслед за этим началась реализация неолиберального проекта масштабной дерегуляции глобальной телекоммуникационной индустрии, что привело к колоссальному разрыву в возможностях формулировать повестку дня между западным «цензурно-индустриальным комплексом» и остальной частью человечества [Sakwa, 2025]. Это произошло потому, что Запад сегодня контролирует львиную долю глобальных информационных ресурсов, включая генеративный искусственный интеллект (далее ГИИ), создающий новый контент, и его составляющую – большие языковые модели (далее БЯМ), генерирующие тексты.

Тексты искусственным интеллектом генерируются на основании данных, собранных из цифровой среды с помощью алгоритмов. Потенциально, любой цифровой документ может оказаться в базах данных, на которых обучаются БЯМ. Объем этих массивов данных зависит от возможностей компаний-разработчиков, принципа отбора информации,

мощности моделей, а также ряда других факторов [Ford, 2018]. Генерируются ИИ-тексты посредством статистического прогнозирования связей между отдельными словами или их частями (токенами). Статистические вероятности тех или иных комбинаций устанавливаются в ходе обучения моделей на тех же массивах данных. Сгенерированные тексты «выглядят авторитетно, поскольку они имитируют стиль множества текстов» [Di Martino, Ford, 2024, p. 354].

Внимание на этом аспекте акцентируется особо, поскольку именно псевдо-авторитетность ИИ вводит в заблуждение пользователей, верящих в точность полученной от него информации. Ученые предупреждают об опасности «галлюцинаций», когда модели авторитетно выдают ложную или бессмысленную информацию, а также их «предубеждений» [Park, Lee, 2024]. Поскольку документы, доступные для обучения БЯМ, создавались людьми, в них могут отражаться всякого рода общественные предрассудки – этнические, расовые, гендерные и т.д. Но цифровые документы могут содержать и идеологические предвзятости. Так, модели, тренированные на материалах западных СМИ, могут генерировать тексты, воспроизводящие западнцентричный взгляд на мировые события [Uzakbayev, Movkebaeva, 2025].

С целью выявления, как именно воспроизводятся идеологические предубеждения по отношению к России, в данной статье исследуется аналитическая записка, созданная ChatGPT – чат-ботом, разработанным американской компанией OpenAI. Заданной темой генерации был российско-украинский конфликт, начало которому положил (Евро)Майдан – протестное движение на Украине (2013–2014), приведшее к смене правительства.

Обзор литературы

В идеологическом плане предубеждения ГИИ чаще всего ассоциируют с его тенденцией воспроизводить левоцентристские нормативные установки – по крайней мере, этот вид идеологических предвзятостей западные ученые

изучают чаще всего. Так, например, в ходе недавнего исследования, проведенного немецкими учеными, был обнаружен серьезный идеологический крен влево, демонстрируемый БЯМ французской разработки (MISTRAL). Авторы считают, что это цифровое искажение политической реальности Европы, где популярность правых только растет, произошло непреднамеренно, а в результате обучения моделей на текстах, где преобладают левоцентристские идеи [Rettenberger, Reischl, Schutera, 2025].

Однако существует и другой взгляд на проблему. Так, группа американских экспертов, исследуя идеологические предвзятости ChatGPT, столкнулась с отказом чат-бота генерировать иллюстрации отношения американцев с правыми взглядами к расовому неравенству и положению трансгендеров в обществе. Модель объяснила свое нежелание создавать эти образы опасностью распространения стереотипов и дезинформации. Это дало ученым основание предположить, что компания-разработчик OpenAI блокирует контент, связанный с правой идеологией [Motoki, Neto, Rangel, 2025].

Эти и другие подобные исследования показывают, что идеологические искажения, генерируемые ИИ, могут быть результатом как случайных ошибок в репрезентации данных на этапе обучений моделей [Akter, McCarthy, Sajib et al., 2021], так и продуманного инжиниринга, когда компании-разработчики блокируют информацию под предлогом желания обезопасить общество от вредоносного контента (Perrigo B. *OpenAI Used Kenyan Workers on Less Than \$2 Per Hour to Make ChatGPT Less Toxic* // Time 18.01.2023 15:00 <https://time.com/6247678/openai-chatgpt-kenya-workers>). Проблема здесь в том, что диагноз общественной вредоносности ставится не уполномоченными представителями общества, а частными лицами, не имеющими на то мандата. Это открывает невиданные возможности для осуществления социального инжиниринга со стороны обладающих цифровой властью [Dean, 2024].

Но если западные ученые исследуют в основном цифровое состязание левой и правой идеологий, то их коллеги с Глобального Юга больше интересуются проблемой воспроизводства идеологических установок западнцентризма. Как отмечают казахстанские ученые Нурлан Узакбаев и Гулия Мовкебаева, обучение ИИ на преимущественно англоязычных данных, опубликованных на Западе, включая источники из американских академических кругов, СМИ и аналитических центров, воспроизводит западное либеральное мировоззрение, формирующее международную политику [*Uzakbayev, Movkebaeva, 2025, p. 317*]. Делается это, как отмечают исследователи, посредством воспроизводства определенных нарративов.

Иллюстрируя этот тезис, Чжао Хуан приводит пример из своего общения с ChatGPT, в ходе которого обсуждались антиправительственные протесты в Китае 1989 года, при подавлении которых, по официальным данным КНР, погиб 241 человек:

Вопрос: Как рассказать студентам о событиях 4 июня в Китае?

ChatGPT: «Это жестокое событие... часто называют «резней на Тяньаньмэнь» или просто «событиями на Тяньаньмэнь» [*Huang, 2024, p. 327*].

Исследователь считает, что сгенерированный ответ – это проявление западнцентризма, поскольку термин «резня», упоминаемый ChatGPT, был центральным в официальном дискурсе западных стран. Альтернативные интерпретации таким образом отсекались.

Исследователи также отмечают: воспроизводя гегемонистские нарративы, БЯМ не склонны демонстрировать сомнения и неопределенность, что делает этот информационный ресурс мало пригодным для информированного суждения (*Goldstein J.A., Sastry G. The Coming Age of AI-Powered Propaganda // Foreign Affairs 07.04.2023 <https://www.foreignaffairs.com/united-states/coming-age-ai-powered-propaganda>*). Это важное наблюдение, поскольку БЯМ все чаще используется не только студентами для написания курсовых работ [*Abbas, Jam, Khan, 2024*], но и разного рода специалистами, работающими

с информацией, например, журналистами. Особую обеспокоенность ученых в этой связи вызывает «скрытое влияние, которое ГИИ может оказывать на ценности аудитории и которое может быть менее явным по сравнению с когнитивными эффектами традиционного новостного освещения» [Shi, Sun, 2024, p. 590].

Однако ГИИ сегодня используют далеко не только журналисты. Многие работники дипломатических ведомств применяют БЯМ для проведения информационного мониторинга [Kuznetsov, Liang, 2023]. Предвзятость ГИИ может повлиять на их понимание ситуации, а также оценивание возможностей и рисков [Manor, Segev, 2025]. К тому же, помимо сбора информации, при помощи БЯМ все чаще пишутся аналитические записки и речи, что создает риск углубления межгосударственных противоречий из-за склонности моделей упрощать сложные проблемы [Di Martin, Ford, 2024]. Тенденцию использования БЯМ в самых разных сферах можно проиллюстрировать недавним инцидентом, произошедшим с главой Агентства кибербезопасности США, загрузившего конфиденциальные документы в диалог с ChatGPT, в результате чего было начато служебное расследование (Sakellariadis J. *Trump's acting cyber chief uploaded sensitive files into a public version of ChatGPT* // Politico 27.01.2026 15:30 <https://www.politico.com/news/2026/01/27/cisa-madhu-gottumukkala-chatgpt-00749361>).

Растущая популярность ГИИ вкупе с его потенциалом влиять на общественное мнение и поведенческие установки людей обуславливает необходимость изучать логику генерации ИИ-текстов. Причем, делать это, как считают ученые, нужно с «методологической диверсификацией и более широким междисциплинарным взаимодействием» [Erfiyanto, Santoso, 2025]. Одним из перспективных методов изучения ГИИ-предубеждений является дискурсивный анализ, используемый авторами для проведения данного исследования.

Описание эмпирической базы и методов исследования

Хотя разные школы дискурсивного анализа используют разные методы исследования дискурса, все они исходят из общего постулата о том, что любое

языковое представление материальной действительности – не объективная истина, а продукт дискурсивной категоризации [*Jorgensen, Phillips, 2002*]. Та или иная версия происходящего может восприниматься в одной социокультурной среде как норма; в другой – как аномалия. Иначе говоря, дискурсивная категоризация всегда культурно и исторически обусловлена. Как сформулировал это Норман Фэркло, один из разработчиков критического дискурс-анализа (КДА), сказанное в тексте (то, что явно присутствует в нем) всегда опирается на то, о чем в тексте не говорится, но подразумевается (неявные установки). Фэркло называет это «интертекстуальностью», обозначающей связь текстов с другими текстами, с «миром текстов» – с тем, «что было сказано, написано или подумано где-то ещё» [*Fairclough, 2013, p. 40*].

Анализу интертекстуальности (то есть воспроизведению «здравого смысла», закрепленного в более ранних текстах) в этом исследовании уделяется особое внимание, поскольку он дает возможность увидеть неявно выраженные предрассудки. Наряду с этим, внимание уделялось и дискурсивным средствам, позволяющим «мистифицировать и затушевывать ответственность субъекта за его действие» [*Fairclough, 2013, p. 13*]. В основном речь идет о персонификации – «оживлении» предметов и процессов – что делает субъектов действия невидимыми.

В исследовании также использовалась теория дискурса (ТД) Эрнесто Лаклау и Шанталь Муфф [*Laclau, Mouffe, 1985*], согласно которой дискурс формируется, когда несколько наиболее важных смысловых узлов – «привилегированных означающих» – объединяются в так называемую «эквивалентную цепь» [*Laclau, Mouffe, 1985, p. 98*]. Как известно из опубликованных научных работ на тему дискурсивного конструирования Майдана, про-майданный и анти-майданный дискурсы характеризовались совершенно разными эквивалентными цепочками, объединяющими разные смысловые узлы [*Baysha, 2021*]. Как видно из Рисунка 1, иллюстрирующего этот тезис, эквивалентная цепь, формирующая про-майданный дискурс, состояла из означающих «евроинтеграция», «демократизация»,

«цивилизация», «независимость от России». Эквивалентная цепь, лежащая в основе анти-майданного дискурса, объединяла в себе «национализм», «экстремизм», «нацизм», «колониальную зависимость от Запада», и т.д.



Рис. 1. Эквивалентные цепи дискурсов Майдана (источник – авторское исследование)

ChatGPT был выбран для анализа, потому что это самый популярный в мире генеративный ИИ, востребованность которого продолжает расти (*Duarte F. Number of ChatGPT Users (January 2026) // Exploding Topics 14.12.2025. <https://explodingtopics.com/blog/chatgpt-users>*). В целях проводимого исследования его модели GPT-5 (бесплатной версии) была задана следующая установка (промпт): Представьте себя в роли политического консультанта, специализирующегося на российско-украинских отношениях. Опишите предпосылки нынешнего конфликта между Россией и Украиной; роль Запада и России в Евромайдане и войне на Донбассе. Перечислите все наиболее важные факторы, приведшие к конфликту, и всех действующих лиц, способствовавших ухудшению российско-украинских отношений.

Разговор с чат-ботом велся на английском языке; длина текста не ограничивалась. Ответ ChatGPT состоял из 958 слов.

Результаты исследования

Начинался текст с сообщения, что нынешний российско-украинский конфликт «не возник внезапно» – противоречия формировались под влиянием как «внутренней украинской динамики», так и «соперничества России и Запада». Протесты на Майдане были выделены в качестве главного

внутреннего фактора ухудшения российско-украинских отношений, и объяснялись они отказом украинского президента Януковича подписать соглашение с ЕС: «Решение президента Януковича приостановить подписание соглашений под экономическим и политическим давлением вызвало протесты».

В этой формулировке примечательно использование в качестве субъекта действия «решения Януковича», которое, по версии ChatGPT, «вызвало протесты». С помощью этой конструкции украинский президент определялся в качестве главного ответственного за то, что дальше происходило на Майдане. При такой репрезентации событий складывалось впечатление, будто протесты происходили сами по себе; будто у них не было организаторов; и будто они не представляли интересы олигархов и глобалистов [Baysha, 2021]. Персонафикация «протестов» – их превращение в субъект действия – позволяла эту часть истории скрыть.

«Для многих протестующих», – говорилось в тексте – «соглашение с ЕС символизировало верховенство права, борьбу с коррупцией, разрыв с олигархическим управлением». Говоря языком дискурсивной теории Лаклау и Муффа (1985), здесь наблюдается конструирование эквивалентной цепи, в которой идентичность Майдана определяется исключительно в рамках нарратива о стремлении к законности, справедливости и сближению с ЕС. При этом альтернативная эквивалентная цепь, в которой Майдан связывался с радикализмом, экстремизмом и зависимостью от Запада – всем тем, с чем протесты ассоциировались в представлении другой половины населения страны (*The views and opinions of south-eastern regions residents of Ukraine // Kyiv International Institute of Sociology 20.04.2014. www.kiis.com.ua/?lang=eng&cat=reports&id=302&y=2014&page=9*) – в сгенерированном тексте отсутствовала.

Словосочетание «многие протестующие» не давало ответа на вопрос, какая часть активистов Майдана действительно связывала протесты с установлением социальной справедливости, а какая работала исключительно на геополитический проект разрыва отношений между Украиной и Россией.

К числу последних, несомненно, относились радикалы-националисты, сыгравшие ключевую роль в государственном перевороте [Ishchenko, 2024; Katchanovski, 2020]. Упоминания об этом в ИИ-тексте не было.

При обсуждении внешних факторов обострения ситуации в тексте совершенно по-разному характеризуются действия России и Запада. Россия, по версии чат-бота, «стремилась сохранить Украину в своей стратегической орбите» и «рассматривала Украину как буфер безопасности»; Запад же «поддерживал требования протестующих политически» и «взаимодействовал с представителями оппозиции дипломатически». Таким образом, Россия представлялась как использующая Украину для достижения своих целей; Запад же представал в роли доброжелательного и равноправного партнера украинцев, хотя соглашение с ЕС было попыткой создания «зоны комфорта» вокруг союза – «кольца друзей», которые не обязательно становились бы членами ЕС [Sakwa, 2014]. Иначе говоря, ЕС поддерживало протестующих, преследуя именно собственные интересы.

В тексте признавалось, что «смещение Януковича произошло без соблюдения всех конституционных процедур», однако не указывалось, кто был ответственен за «смещение», которое в этой грамматической конструкции было снова-таки персонифицировано и превращено в субъект. При этом эвфемизм «без соблюдения всех конституционных процедур» маскировал факт того, что переворот – «операция под ложным флагом с участием элементов олигархической и крайне правой оппозиции Майдана» [Katchanovski, 2024, p. 20] – был полностью антиконституционным. Упоминание применения оружия протестующими в повествовании чат-бота отсутствовало.

В тексте также отсутствовало и внятное изложение того, что произошло на Донбассе весной 2014 года. Признавалось, что местное население испытывало «недоверие к властям после Майдана» из-за «страха перед культурной маргинализацией, централизацией и националистической политикой». При этом, однако, не объяснялось, почему реакцией Киева на эти

опасения Донбасса стала «антитеррористическая операция» с использованием армии, целью которой, как следовало из текста, была «защита территориальной целостности». Воспроизводился западный нарратив, в котором люди Донбасса, протестующие против переворота, были дискурсивно преобразованы в «террористов» [Bausha, 2023].

При обсуждении роли России в этих событиях было отмечено, что она предоставляла повстанцам на Донбассе «военную технику, разведку, логистическую поддержку», а в ключевые моменты осуществляла «прямое военное вмешательство». Целью России, по версии чат-бота, было «предотвратить военную победу Украины» и «принудить к политическим уступкам (федерализация, нейтралитет)». При этом что в тексте речь шла о федерализации – не отделении Донбасса от Украины (что было отражено в Минских соглашениях) – воспроизводился западный нарратив о том, что повстанцы на Донбассе были «сепаратистами». Их, как указывалось в тексте, представляли «сепаратистские деятели» и «сепаратистские власти».

При этом роль Запада, по версии чат-бота, заключалась в том, что он «оказывал Украине дипломатическую и экономическую поддержку; вводил санкции против России; не обеспечивал и не гарантировал выполнение Минских соглашений; постепенно увеличивал военную помощь после 2015 года». То есть акцент делался на дипломатических усилиях Запада – военная составляющая упоминалась в последнюю очередь. Причем применительно к Западу, Минские соглашения фигурировали в непонятной формулировке (которая вполне может служить примером ИИ-галлюцинаций – что значит «не обеспечивал и не гарантировал Минские соглашения?»). При обсуждении же роли России формулировка была предельно ясна: она «использовала Минск в качестве рычага давления, а не для урегулирования спора». То есть российский вклад в обострение кризиса обсуждался исключительно в силовом ключе.

Отмечалось «нежелание» Украины предоставлять особый статус Донбассу, что предписывалась Минскими договоренностями, однако это объяснилось «давлением» – разумеется, российским. Оговорка

«под давлением», сгенерированная ИИ, выглядела как оправдание нарушению Украиной международного договора, ратифицированного ООН и ОБСЕ. Роль же Запада в срыве соглашений вообще не обсуждалась. Как было отмечено в тексте, Германия и Франция «обладали ограниченными возможностями по обеспечению соблюдения требований» договора, что подразумевало несуществующее на самом деле желание Берлина и Парижа претворить Минские договоренности в жизнь.

Наконец, при обсуждении «финальной эскалации на пути к полномасштабной войне» в тексте признавалось, что Украина «расширила военное сотрудничество с НАТО», и это было воспринято Россией как «экзистенциальная угроза», в результате чего Россия поняла, что «дипломатические инструменты потерпели неудачу». Запад же, по версии ИИ, пришел к выводу, что «сдерживание посредством санкций и военной помощи предпочтительнее уступок».

В этой конструкции, завершающей текст, обращают на себя внимание два момента. Во-первых, упоминание осознания Россией тщетности дипломатических усилий, хотя нигде в тексте ни о каких попытках России уладить конфликт с помощью дипломатии не упоминалось. Во-вторых, актуализация смыслового узла о «сдерживании» по отношению к России. Несмотря на признание того факта, что Россия воспринимала расширение НАТО как экзистенциальную угрозу, сдерживать, как следовало из текста, нужно было все же Россию – не НАТО. Это наглядный пример воспроизведения «здорового смысла», нормализованного в западном политическом дискурсе: угрозой для глобального мира при любых обстоятельствах остается Россия.

Дискуссия

Как видно из представленного анализа, несмотря на обещание чат-бота о «нейтральности» его аналитической записки, нейтральности в ней не наблюдалось. В сгенерированном тексте история Майдана и всех последующих событий, начало которым он положил, представлялась точно такой, какой она изо дня в день изображалась западными СМИ во время всех

этих событий [Boyd-Barrett, 2017]. Это был нарратив о борьбе добра и зла, свободы и тирании, демократии и автократии – черно-белое повествование, упускающие все важные нюансы, которые могли бы помочь пользователю разобраться в ситуации. Воссоздавался именно тот образ России, который задолго до пришествия ГИИ в нашу повседневную жизнь был нормализован в западном политическом дискурсе [Boyd-Barrett, Marmura (ed.), 2023; Tsygankov, 2026].

При презентации истории Майдана не были упомянуты такие важные факторы как отсутствие поддержки Майдана половиной населения Украины (*The views and opinions of south-eastern regions residents of Ukraine // Kyiv International Institute of Sociology 20.04.2014. www.kiis.com.ua/?lang=eng&cat=reports&id=302&y=2014&page=9*); организованный характер протестов, их финансирование местными олигархами и внешними игроками [Sakwa, 2014]; решающая роль националистов-радикалов в совершении переворота [Katchanovski, 2024]; и многое другое.

При обсуждении анти-майданного восстания на Донбассе в тексте не объяснялось, почему протестующих там называли «сепаратистами», если их требованием было не отделение от Украины, а ее федерализация [Ishchenko, 2023]. Не объяснялось также и то, почему новое руководство Украины направило против мятежников на Донбассе танки, в то время как само это правительство пришло к власти именно в результате вооруженного мятежа. Вопрос о том, почему протестующие на Майдане в западном дискурсе стали героями, а протестующие на Донбассе террористами, тоже не обсуждался.

В сгенерированном тексте совершенно по-разному были изображены Россия и Запад, а также их роль в развитии кризиса. Запад представлялся добросердечным партнером Украины, стремящимся помочь ее народу; Россия же – колониальной силой, использующей Украину в собственных имперских интересах. При обсуждении роли России внимание

акцентировалось на шантаже и военных угрозах; при описании роли Запада акцент делался на дипломатии.

При этом хорошо известно, что Запад относился к Украине именно как к объекту реализации своих геополитических амбиций. Как повествует Тим Вейнер в недавно вышедшей книге «Миссия», сразу же после государственного переворота, в течение одной недели, по требованию ЦРУ «было уволено свыше 90% высшего офицерского состава СБУ [Службы Безопасности Украины]» [Weiner, 2025, p. 363]. Хорошо известно также и то, что США руководили назначениями не только в новый Кабинет Министров Украины, но и в другие органы пост-майданной власти [Baysha, 2023].

Тот факт, что Минские соглашения разрабатывала Россия, которая настаивала на их выполнении, в сгенерированном тексте не упоминался. Не говорилось в нем также и о том, что западные страны – в частности Германия и Франция – не были заинтересованы в выполнении Минских договоренностей. Как в этом позже признались Ангела Меркель и Франсуа Олланд (*Prouvost T. Hollande: “There will only be a way out of the conflict when Russia fails on the ground.” // The Kyiv Independent 28.12.2022 <https://kyivindependent.com/national/hollande-there-will-only-be-a-way-out-of-the-conflict-when-russiafails-on-the-ground>*), они-то как раз и тормозили процесс с целью дать Украине восстановиться, чтобы та снова могла воевать. Начиная с 2014 года только Соединенные Штаты поставили новому украинскому режиму оружие на миллиарды долларов [Weiner, 2025].

Все наблюдения, сделанные в ходе проведенного анализа, подтвердили опасения других исследователей о том, что тексты, генерируемые БЯМ, могут демонстрировать идеологическую предвзятость [Roozafzai, 2024], воспроизводить западнцентризм [Uzakbayev, Movkebaeva, 2025], делать это с претензией на авторитетность и нейтральность [Di Martino, Ford, 2024], и что они гораздо более пригодны для пропаганды, чем для информированного суждения (*Goldstein J.A., Sastry G. The Coming Age of AI-Powered Propaganda // Foreign Affairs 07.04.2023 <https://www.foreignaffairs.com/united-states/coming-age-ai-powered-propaganda>*).

Заключение

С использованием двух методов дискурсивного анализа в данной статье была проанализирована аналитическая записка по истории ухудшения российско-украинских отношений, сгенерированная чат-ботом ChatGPT производства OpenAI, США. В результате проведенного исследования были подтверждены выводы других аналитиков о том, что ГИИ склонен воспроизводить гегемонистские идеологии посредством воспроизводства специфических нарративов. Учитывая растущую популярность БЯМ в мире, проблема выглядит серьезно: чем чаще пользователи ИИ сталкиваются с однотипными идеологическими формулировками, тем выше шанс того, что проповедуемая идеология превратится в бессознательную массовую ориентацию [Roozafzai, 2024].

Воспроизведение идеологических установок западнотризма, которыми пронизаны гегемонистские нарративы, генерируемые БЯМ, лишь укрепляет существующие отношения глобального неравенства в информационной сфере. Выводы ЮНЕСКО, сделанные 45 лет назад, оказались пророческими. Проблема, однако, крайне усугубилась; она требует всесторонних исследований и публичных обсуждений.

Библиография

Abbas M., Jam F. A., Khan T. I. Is it harmful or helpful? Examining the causes and consequences of generative AI usage among university students. *International Journal of Educational Technology in Higher Education*. 2024. Vol. 21. No. 1. P. 10–32.

Akter S., McCarthy G., Sajib S., Michael K., Dwivedi Y.K., D'Ambra J., Shen K.N. Algorithmic bias in data-driven innovation in the age of AI. *International Journal of Information Management*. 2021. Vol. 60. P. 102387.

Baysha O. Democracy, populism and neoliberalism in Ukraine: On the fringes of the virtual and the real. New York: Routledge. 2021. 141 p.

Baysha O. War, peace, and populist discourse in Ukraine. New York: Routledge. 2023. 128 p.

Boyd-Barrett J. O. Ukraine, mainstream media and conflict propaganda. *Journalism Studies*. 2017. Vol. 18. No. 8. P. 1016–1034.

Boyd-Barrett O., Marmura S. Russiagate revisited: The aftermath of a hoax. Springer Nature. 2023. 307 p.

Dean J. The Neofeudalising Tendency of communicative capitalism. TripleC: Communication, Capitalism & Critique. 2024. Vol. 22. No. 1. P. 197–207.

Di Martino L., Ford H. Navigating uncertainty: Public diplomacy vs. AI. Place Branding and Public Diplomacy. 2024. Vol. 21. No. 3. P. 350–354.

Erfiyanto R., Santoso M.P.T. Algorithmic influence and American public diplomacy: Ethical dimensions of AI-powered political advocacy on social media. Jurnal Terekam Jejak. 2025. Vol. 3. No. 1. P. 1–17.

Fairclough N. Language and power. London: Routledge. 2013. 240 p.

Ford M. Architects of intelligence. Birmingham: Packt Publishing. 2018. 554 p.

Huang Z.A. Terminology, AI bias, and the risks of current digital public diplomacy practices. Place Branding and Public Diplomacy. 2024. Vol. 21. No. 3. P. 327–333.

Ishchenko V. The Minsk Accords and the political weakness of the “Other Ukraine”. Russian Politics. 2023. Vol. 8. No. 2. P. 127–146.

Ishchenko V. Towards the abyss: Ukraine from Maidan to war. London, New York: Verso Books. 2024. 165 p.

Jorgensen M.W., Phillips L. Discourse analysis as theory and method. London: Sage. 2002. 240 p.

Katchanovski I. The far right, the Euromaidan, and the Maidan massacre in Ukraine. Journal of Labor and Society. 2020. Vol. 23. No. 1. P. 5–36.

Katchanovski I. The Maidan massacre in Ukraine: The mass killing that changed the world. New York: Palgrave Macmillan. 2024. 296 p.

Kuznetsov N. M., Liang F. Digital diplomacy of USA and China in the era of datalization. Vestnik of Saint Petersburg University. International Relations. 2023. Vol. 16. No. 2. P. 191–200.

Laclau E., Mouffe C. Hegemony and socialist strategy. New York: Verso. 1985. 197 p.

Manor I., Segev E. What ChatGPT “thinks” about your country? Sentiments and frames of AI geographies. Policy & Internet. 2025. Vol. 17. No. 3. P. e70013.

Motoki F. Y., Neto V. P., Rangel V. Assessing political bias and value misalignment in generative artificial intelligence. Journal of Economic Behavior & Organization. Vol. 234. P. 106904.

Park D. M., Lee H. J. Literature review of AI hallucination research since the advent of ChatGPT: Focusing on papers from arXiv. Informatization Policy. 2024. Vol. 31. No. 2. P. 3–38.

Rettenberger L., Reischl M., Schutera M. Assessing political bias in large language models. Journal of Computational Social Science. 2025. Vol. 8. P. 42.

Roozafzai Z. S. Unveiling power and ideologies in the age of algorithms: Exploring the intersection of critical discourse analysis and artificial intelligence. *Qeios*. 2024. Vol. 10. No. 1. P. 51–63.

Sakwa R. *Frontline Ukraine: Crisis in the borderlands*. London: Tauris. 2014. 347 p.

Sakwa R. *The culture of the second cold war*. London: Anthem Press. 2025. 198 p.

Shi Y., Sun L. How generative AI is transforming journalism: Development, application and ethics. *Journalism and Media*. 2024. Vol. 5. No. 2. P. 582–594.

Thussu D. K. *International communication: Continuity and change*. Bloomsbury Publishing. 2018. 392 p.

Tsygankov A. P. “Canceling” Russia: The Ukraine war and the rise of the Western hawks. Cham, Switzerland: Palgrave Macmillan. 2026. 251 p.

Uzakbayev N., Movkebaeva G. Teaching international relations in the age of AI-Western bias embedded in ChatGPT. *Pedagogical Sciences*. 2025. Vol. 77. No. 2. P. 317–329.

References

Abbas M., Jam F. A., Khan T. I. (2024). Is it harmful or helpful? Examining the causes and consequences of generative AI usage among university students. *International journal of educational technology in higher education* 21(1): 10–32. <https://doi.org/10.1186/s41239-024-00444-7>

Akter S., McCarthy G., Sajib S., Michael K., Dwivedi Y.K., D’Ambra J., Shen K.N. (2021). Algorithmic bias in data-driven innovation in the age of AI. *International Journal of Information Management* 60: 102387. <https://doi.org/10.1016/j.ijinfomgt.2021.102387>

Baysha O. (2021). *Democracy, populism and neoliberalism in Ukraine: On the fringes of the virtual and the real*. New York: Routledge. <https://doi.org/10.4324/9781003228493>

Baysha O. (2023). *War, peace, and populist discourse in Ukraine*. New York: Routledge. <https://doi.org/10.4324/9781003379164>

Boyd-Barrett J. O. (2017). Ukraine, mainstream media and conflict propaganda. *Journalism Studies* 18(8): 1016–1034. <https://doi.org/10.1080/1461670x.2015.1099461>

Boyd-Barrett O., Marmura S. (eds.). (2023). *Russiagate revisited: The aftermath of a hoax*. Springer Nature. <https://doi.org/10.1007/978-3-031-30940-3>

Dean J. (2024). The neofeudalising tendency of communicative capitalism. *TripleC: Communication, Capitalism & Critique* 22(1): 197–207. <https://doi.org/10.31269/triplec.v22i1.1460>

Di Martino L., Ford H. (2024). Navigating uncertainty: Public diplomacy vs. AI. *Place Branding and Public Diplomacy* 21(3): 350–354. <https://doi.org/10.1057/s41254-024-00330-z>.

Erfiyanto R., Santoso M.P.T. (2025). Algorithmic influence and American public diplomacy: Ethical dimensions of AI-powered political advocacy on social media. *Jurnal Terekam Jejak* 3(1): 1–17.

Fairclough N. (2013). *Language and power*. London: Routledge. <https://doi.org/10.4324/9781315838250>

Ford M. (2018). *Architects of intelligence*. Birmingham: PacktPublishing.

Huang Z.A. (2024). Terminology, AI bias, and the risks of current digital public diplomacy practices. *Place Branding and Public Diplomacy* 21(3): 327–333. <https://doi.org/10.1057/s41254-024-00324-x>

Ishchenko V. (2023). The Minsk Accords and the political weakness of the “Other Ukraine”. *Russian Politics* 8(2): 127–146. <https://doi.org/10.30965/24518921-00802002>

Ishchenko V. (2024). *Towards the abyss: Ukraine from Maidan to war*. London, New York: Verso.

Jorgensen M.W., Phillips L. (2002). *Discourse analysis as theory and method*. London: Sage. <https://doi.org/10.4135/9781849208871>

Katchanovski I. (2020). The far right, the Euromaidan, and the Maidan massacre in Ukraine. *Journal of Labor and Society* 23(1): 5–36. <https://doi.org/10.1111/wusa.12457>

Katchanovski I. (2024). *The Maidan massacre in Ukraine: The mass killing that changed the world*. New York: Palgrave Macmillan.

Kuznetsov N. M., Liang F. (2023). Digital diplomacy of USA and China in the era of datalization. Vestnik of Saint Petersburg University. *International Relations* 16(2): 191–200. <https://doi.org/10.21638/spbu06.2023.206>

Laclau E., Mouffe C. (1985). *Hegemony and socialist strategy*. New York: Verso.

Manor I., Segev E. (2025). What ChatGPT “thinks” about your country? Sentiments and frames of AI geographies. *Policy & Internet* 17(3): e70013. <https://doi.org/10.1002/poi3.70013>

Motoki F. Y., Neto V. P., Rangel V. (2025). Assessing political bias and value misalignment in generative artificial intelligence. *Journal of Economic Behavior & Organization* 234: 106904. <https://doi.org/10.1016/j.jebo.2025.106904>

Park D. M., Lee H. J. (2024). Literature review of AI hallucination research since the advent of ChatGPT: Focusing on papers from arXiv. *Informatization Policy* 31(2): 3–38. <https://doi.org/10.22693/NIAIP.2024.31.2.003>

Rettenberger L., Reischl M., Schutera M. (2025). Assessing political bias in large language models. *Journal of Computational Social Science* (2025) 8: 42. <https://doi.org/10.1007/s42001-025-00376-w>

Roozafzai Z. S. (2024). Unveiling Power and Ideologies in the Age of Algorithms: Exploring the Intersection of Critical Discourse Analysis and Artificial Intelligence. *Qeios*. 10(1): 51–63. <https://doi.org/10.32388/60ye02>

Sakwa R. (2014). *Frontline Ukraine: Crisis in the borderlands*. London: Tauris. <https://doi.org/10.5040/9780755603756>

Sakwa R. (2025). *The culture of the second cold war*. London: Anthem Press. <https://doi.org/10.2307/jj.24751871>

Shi Y., Sun L. (2024). How generative AI is transforming journalism: Development, application and ethics. *Journalism and Media* 5(2): 582–594. <https://doi.org/10.3390/journalmedia5020039>

Thussu D. K. (2018). *International communication: Continuity and change*. Bloomsbury.

Tsygankov A. P. (2026). “Canceling” Russia: The Ukraine war and the rise of the Western hawks. Cham, Switzerland: Palgrave Macmillan. <https://doi.org/10.1007/978-3-032-11320-7>

Uzakbayev N., Movkebaeva G. (2025). Teaching international relations in the age of AI-Western bias embedded in ChatGPT. *Pedagogical Sciences* 77(2): 317–329. <https://doi.org/10.48371/PEDS.2025.77.2.019>

Сведения об авторах

Байша Ольга Анатольевна – Ph.D., доцент факультета мировой экономики и мировой политики Национального исследовательского университета «Высшая школа экономики» (Москва) e-mail: obaysha@hse.ru

Трофимов Владислав Викторович – студент магистерской программы «Доказательное развитие образования» (НИУ ВШЭ, Москва), стажёр-исследователь Института образования НИУ ВШЭ (Москва), e-mail: yv.trofimov@hse.ru

Baysha Olga A. – Ph.D., Associate Professor of Faculty of World Economy and International Affairs, National Research University–Higher School of Economics, Moscow, Russia e-mail: obaysha@hse.ru.

Trofimov Vladislav V. – Student of the Master’s program “Evidence-Based Educational Development” (National Research University Higher School of Economics, Moscow), Research Assistant in Institute of Education, National Research University–Higher School of Economics, Moscow, Russia, e-mail: vv.trofimov@hse.ru.