

МЕЖДУНАРОДНЫЕ, ГЛОБАЛЬНЫЕ И РЕГИОНАЛЬНЫЕ ПРОЦЕССЫ
Международные отношения, глобальные
и региональные исследования
INTERNATIONAL, GLOBAL AND REGIONAL PROCESSES
International Relations, Global and Regional Studies

Научная статья

Политические науки

УДК 327

[https://doi.org/10.53658/RW2026-4-1\(19\)-43-62](https://doi.org/10.53658/RW2026-4-1(19)-43-62)

Возможности и ограничения научного использования LLM: первичный анализ данных, проблемы предвзятости и валидация

Юрий Юрьевич Колотаев^{1a}✉, Мария Михайловна
Базлуцкая^{2b}✉

¹ Санкт-Петербургский государственный университет, Санкт-Петербург, Россия

² Автономная некоммерческая научно-исследовательская организация «Координационная лаборатория» (АНО «Колаборатория»), Санкт-Петербург, Россия

^a yury.kolotaev@mail.ru, <https://orcid.org/0000-0001-8372-1193>

^b m.bazlutskaya@gmail.com, <https://orcid.org/0000-0002-7118-8632>

Аннотация. Большие языковые модели (LLM) становятся популярными среди научного сообщества, а их применение все чаще можно встретить в социогуманитарных исследованиях. Представленная обзорная статья обобщает раскрытые возможности внедрения LLM в текстовый анализ данных и систематизирует ограничения, с которыми приходится сталкиваться ученым на этом пути. Авторы обозначают наилучшие «точки входа» в исследования с помощью LLM, но наибольшее внимание уделяется проблемам предвзятости моделей, валидации (проверке) и репликации

(воспроизводимости) результатов исследований, в которых использовался ИИ. В работе предлагаются несколько возможных стратегий улучшения качества работы с генеративными моделями в соответствии с процедурой триангуляции. Они включают проверку альтернативных запросов, тестирование разных выборок данных и использование связки моделей. Практика применения LLM в гуманитарных науках показывает, что при грамотной настройке они имеют потенциал к снижению временных издержек, расширяют аналитические возможности ученых и могут содействовать в выявлении скрытых закономерностей в текстовых массивах. Однако эффективность научного применения LLM напрямую зависит от исследовательской осмотрительности, выраженной в понимании границ применимости инструмента, корректности постановки задач, качества исходных данных, а также умения нормализовать неструктурированные данные. Без этих условий использование моделей рискует превратиться в симуляцию научного исследования. Статья призвана стать стартовой точкой для политологов и исследователей международных отношений, заинтересованных в качественном внедрении LLM в свою аналитическую работу.

Ключевые слова: искусственный интеллект (ИИ), социальные науки, большая языковая модель (LLM), предвзятость, валидация, триангуляция

Для цитирования: Колотаев Ю.Ю., Базлутцкая М.М. Возможности и ограничения научного использования LLM: первичный анализ данных, проблемы предвзятости и валидация // Россия и мир: научный диалог. 2026. № 1(19). С. 43-62, [https://doi.org/10.53658/RW2026-4-1\(19\)-43-62](https://doi.org/10.53658/RW2026-4-1(19)-43-62)

Original article

Political Sciences

[https://doi.org/10.53658/RW2026-4-1\(19\)-43-62](https://doi.org/10.53658/RW2026-4-1(19)-43-62)

Capabilities and Limitations in Scientific Application of LLMs: Preliminary Data Analysis, Bias, and Validation

Yury Yu. Kolotaev^{1a}, Mariya M. Bazlutckaya^{2b}¹ Saint-Petersburg State University, Saint-Petersburg, Russia² Autonomous Non-Commercial Research Organisation "Coordination Lab" (ANO Colaboratoria), Saint-Petersburg, Russia^a yury.kolotaev@mail.ru, <https://orcid.org/0000-0001-8372-1193>^b m.bazlutckaya@gmail.com, <https://orcid.org/0000-0002-7118-8632>

Abstract. Large language models (LLMs) are becoming increasingly popular within the academic community, and their use is now more frequently observed in social science. This article summarizes the emerging opportunities for integrating LLMs into textual data analysis and systematizes the limitations that scholars encounter along the way. The authors identify the most effective "points of entry" for incorporating LLMs into research, while devoting particular attention to issues of model bias, validation, and the replication of results produced with the AI-assistance. The paper proposes several possible strategies for enhancing the quality of work with generative models in accordance with the triangulation procedure. These include examining alternative prompts, testing various data samples, and employing a combination of models. Existing studies show that, when

properly configured, LLMs can reduce time costs, expand researchers' analytical capacities, and help uncover hidden patterns within large textual corpora. However, the effectiveness of scientific applications of LLMs directly depends on scholarly diligence, including a clear understanding of the tool's scope, careful problem formulation, high-quality input data, and the ability to normalize unstructured data. Without these conditions, the use of such models risks devolving into a simulation of scientific inquiry. The article is intended to serve as a starting point for political scientists and international relations researchers interested in integrating of LLMs into their analytical work.

Keywords: artificial intelligence (AI), social science, large language model (LLM), bias, validation, triangulation

For citation: Kolotaev Yu.Yu., Bazlutckaya M.M. Capabilities and Limitations in Scientific Application of LLMs: Preliminary Data Analysis, Bias, and Validation. Russia & World: Scientific Dialogue. 2026; 1(19): 43-62, [https://doi.org/10.53658/RW2026-4-1\(19\)-43-62](https://doi.org/10.53658/RW2026-4-1(19)-43-62)

Введение

Всплеск общественного и научного интереса к искусственному интеллекту (ИИ) и большим языковым моделям (large language models, LLM) вызвал обширную дискуссию о границах и этике их применения в исследовательском поле. Технические возможности, позволяющие открыто работать в интерактивной форме с любой текстовой информацией, создали основу для делегирования части научной работы ИИ-ассистентам. В перечень этой работы уже входит аннотирование, предобработка, кластеризация, а также моделирование научных данных.

Вместе с тем уровень развития генеративных моделей, а также принципы работы с ними показывают, что некритическое отношение к ИИ может приводить лишь к имитации научного процесса. Ведь легкость получения ответов на комплексные вопросы остается во многом иллюзией. Это, однако, не снижает эвристического потенциала комбинированных методов научного анализа, включающих генеративный ИИ в рабочий процесс.

Данная работа стремится отразить актуальную дискуссию о возможностях научного использования генеративных моделей. Ключевое внимание уделено социально-гуманитарному направлению исследований с особым акцентом на текстовых методах анализа. Статья рассматривает наилучшие «точки входа» для применения LLM, где вероятность научных искажений маловероятна или незначительна. В частности, рассматривается возможность делегирования LLM задач по очистке и стандартизации текстовых данных, а также моделированию и предварительной проверке гипотез.

Основной фокус работы направлен на осмысление опыта научного применения открытых генеративных моделей (ChatGPT, DeepSeek, Gemini, Anthropic и т.д.) к исследованию массивов данных и выявлению в них закономерностей.

Ключевое внимание в работе уделяется фундаментальным предвзятостям LLM, которые выступают существенным ограничением в рамках применения генеративных моделей к анализу текстов. Авторы статьи также рассматривают применимость принципов валидации и репликации к комбинированным с

ИИ методам научного анализа. Дискуссии об этичности использования ИИ при написании научных текстов остаются за рамками данного исследования.

Материалы и методы

Обзор применения LLM в социогуманитарных исследованиях

Долгое время ключевой массив работ, затрагивающих принципы работы LLM, носил технический характер. Однако вместе с массовым распространением потребительских LLM увеличилось и количество гуманитарных исследований, посвященных ИИ как социальному феномену. Ученые стали философски рассматривать работу с LLM [7], разрабатывать вопросы ИИ в образовании [6], правовую основу использования ИИ [5], способность машин решать моральные дилеммы [3] и этику применения ИИ к исследованиям [12]. Одновременно с этим появляются обзорные работы, цель которых – обозначить потенциал применения LLM в гуманитарной науке [32]. Так, Тао и Шен отмечают, что из 1 227 проанализированных ими аннотаций около 83% положительно отзывались о научно-исследовательском потенциале и педагогической применимости LLM и только около 10% работ были негативно настроены в отношении ИИ [34].

Параллельно возник пласт научных текстов, продвигающий инструментализацию ИИ для усиления результатов проводимых исследований. С конца 2022 г. отношение специалистов смещается от настороженного наблюдения за использованием моделей к допущению, что LLM могут выступать в качестве полноценных помощников при проведении социального и политического анализа. При этом если ранее для комплексного анализа «больших данных» чаще всего возникала необходимость в сторонней помощи ИТ-сообщества по настройке компьютерных программ [4], то новейшие труды показывают потенциал самостоятельной работы специалистов по социально-гуманитарным дисциплинам за счет взаимодействия с ИИ (prompt engineering). В период с 2023 г. начинается активное сравнение возможностей ученых и LLM в исследованиях текстовой информации [35] (контент-анализ, лингвистический анализ, сентимент-анализ и тематическое моделирование) и количественных данных [1]. Первые эксперименты с ChatGPT версии 3.5 подчеркивают сокращение сроков проведения исследований и снижение предвзятости самого ученого за счет взаимодействия с большими языковыми моделями [2]. Особое внимание исследователи уделяют потенциалу конкретных видов LLM, а именно – предобученных трансформеров (general pre-trained transformers, GPT). Эти модели позволяют автоматизировать ряд рутинных, но важных этапов – от первичного поиска и фильтрации источников до тематического кодирования и выявления ключевых концептов. В отличие от машинного обучения (ML) и обработки естественного языка (NLP) большие языковые модели способны улавливать семантические взаимосвязи между текстами, что

упрощает анализ научного дискурса на уровне понятий и аргументативных структур [19].

Появление моделей, имитирующих логические рассуждения (reasoning-LLM), таких как ChatGPT 4, DeepSeek R1, Claude 3.7 Sonnet и других, раскрывает новые возможности в части проведения социогуманитарных исследований. Например, в рамках аннотации текста такие модели показывают более высокую точность (93,4%) и согласованность (0,982 по альфе Криппендорфа [22]), нежели предыдущие версии LLM, предрасположенные к более поверхностным результатам из-за недостаточно глубокого контекстного погружения [35]. Исследователи также отмечают, что результаты анализа, проведенного новейшими моделями, могут превосходить таковые экспертов-кодировщиков. Особую ценность подключение LLM приобретает в случае, если модель помогает формулировать и тестировать гипотезы [23], а также интерпретировать входные данные [21]. Это может высвободить дополнительное время исследователей для более глубокого погружения в проблематику, тем самым способствуя развитию научного знания.

Текстовый анализ с помощью LLM в политических исследованиях и международных отношениях

В политической науке и международных отношениях используется широкий спектр исследовательских подходов, которые условно подразделяются на качественные, количественные и смешанные методы. Существенное место среди них занимает работа с текстами, такими как политические речи, официальные документы, интервью, публикации в СМИ и сообщения в социальных сетях. В условиях стремительной цифровизации и роста объемов информации возрастает потребность в анализе больших текстовых массивов, часть из которых крайне неструктурирована. Это обстоятельство стимулирует внедрение вычислительных и автоматизированных инструментов, включая методы NLP и LLM в исследования международных отношений и политологии. Особенно заметным оказывается потенциал LLM в области дискурс- и контент-анализа, поскольку именно эти методы объединяют количественную обработку текста с интерпретацией политических смыслов и нарративов.

В первую очередь LLM хорошо справляются с нормализацией текста, исправлением орфографических ошибок, приведением к единому формату, что является важнейшим этапом перед началом анализа [10]. Например, если текст содержит устаревшие формы слов, LLM могут адаптировать текст для понимания кодировщиками, тем самым создавая корпуса данных высокого качества [27]. Подсчет частотности употребляемых слов и тем также выступает для моделей типовой задачей [14]. Одновременно с этим LLM могут помочь протестировать аналитический потенциал гипотезы, а также провести их предварительную проверку [37]. Вышеперечисленные операции остаются на сегодня наилучшими «точками входа» для применения LLM в гуманитарных исследованиях.

Благодаря LLM для научного сообщества одновременно расширяются и становятся более доступными возможности автоматизированного дискурс-анализа, то есть анализа неявных смысловых структур, включая подтекст. Так, К.Гао и Д.Фенг применяют LLM на открытом исходном коде Llama2 (13b), чтобы оценить скрытые смыслы, прослеживающиеся в одной из крупнейших газет Гонконга «Oriental Daily News». Авторы приходят к выводу, что модели с 80% вероятностью (относительно результата профессиональных кодировщиков) определяют как явные выражения отношения (attitude), так и менее выраженные [17], что оставляет пространство для улучшения, но уже говорит о значимой степени точности машинного дискурс-анализа.

В рамках контент-анализа применение LLM распространено более отчетливо. В частности, благодаря большим языковым моделям развиваются анализ тональности текста (сентимента), тематическое моделирование, анализ нарративов, фреймов и иных форм предвзятости. LLM используются и для совершенствования справочников кодировщиков (codebooks) [18].

Среди всего спектра работ по анализу тональности текста и политических предпочтений выделяется коллективный труд под руководством Л.Бойич [9]. В нем предлагается системное сопоставление когнитивных характеристик работы человека и современных языковых моделей при решении задач анализа латентного содержания текста. Авторы указали, что западные LLM демонстрируют высокую степень согласованности (intercoder reliability) при определении тональности текста и политической направленности, в ряде случаев превосходя показатели согласованности ответов кодировщиков. При этом выявлено, что при автоматизированном анализе модели склонны недооценивать степень выраженности эмоций в тексте.

Исследования данных из социальных сетей, форумов и даже даркнета формируют среди научного сообщества спрос на тематическое моделирование – «критически важный инструмент для выявления и анализа тематических структур неструктурированных текстовых данных» [15, р. 67159]. Авторы отмечают, что LLM, в отличие от ранее применявшихся компьютерных методов, способны показывать более точные результаты при определении тематик коротких неформальных сообщений, которые часто содержат мало слов, используют нестандартные выражения и сильно зависят от контекста. К тому же, LLM выделяют более осмысленные и контекстно точные двухсловные темы, нежели такие традиционные методы, как векторизация (Term Frequency – Inverse Document Frequency, TF-IDF) и латентное распределение Дирихле (Latent Dirichlet Allocation, LDA), что повышает точность последующих выводов.

Таким образом, изучение работы LLM в привязке к анализу нарративов и предвзятостей показывает, что большие языковые модели могут применяться в качестве инструмента совершенствования результата, а не просто автоматизации. Как подчеркивают С.Дженнер и соавторы, LLM способны расширять нарративный анализ, выполненный вручную, через обнаружение дополнительных интерпретаций смысла в тексте [20]. Коллектив российских исследователей, в свою очередь, видит

потенциал LLM в определении политической предвзятости в цифровой дипломатии, что позволяет эффективнее определять акценты и смещенность (bias) при анализе медиакоммуникаций [8].

Однако задействование LLM в рамках исследований связано с рядом затруднений (не в последнюю очередь из-за новизны самой технологии). Во-первых, ученые обращают особое внимание на внутренние ограничения моделей, которые непреднамеренно или целенаправленно могут исказить результаты исследования. Это проблематизирует использование LLM в науке. Так, группа американских исследователей подчеркивает, что хотя LLM имитируют способности к рассуждению, они зачастую опираются на распознавание шаблонов, а не на подлинное логическое рассуждение, что приводит к непоследовательности при решении сложных задач [31]. Кроме того, в отношении LLM системно поднимается вопрос о характере и составе данных, на которых обучена модель. Последние выступают в качестве основы для появления предвзятости получаемых результатов, так как в «обучающих» модель датасетах присутствуют определенные культурные или идеологические установки [13]. Наконец, чувствительность LLM к вводным данным и отсутствие инструментов, которые могли бы предсказывать уязвимость промпта, способствуют искажению результатов проводимого исследования [30]. Отсюда вытекают и наиболее насущные в работе с LLM проблемы валидации (проверки достоверности) и репликации (воспроизведения) результатов, которые будут раскрыты ниже.

Результаты исследования

Проблема предвзятости моделей

Несмотря на сильные стороны научного применения пользовательских LLM, сохраняются множественные ограничения, с которыми приходится сталкиваться на этапах очистки, стандартизации и предобработки данных. Исходя из прикладных характеристик, их можно объединить в 5 основных подгрупп. Каждая подгруппа основывается на типовых предвзятостях (bias) моделей: 1) предвзятость формулировки; 2) ошибки стандартизации данных; 3) избыточный / недостаточный контекст; 4) пользовательская ориентированность; 5) потеря контекста. При рядовом использовании LLM все эти предвзятости или ошибки ведут к иллюзии, что модель «врет» или, выражаясь технически, «галлюцинирует».

В действительности же можно наблюдать переплетение предубеждений, формирующееся из-за стохастичности¹ самого процесса генерации текста, различных

1 В контексте ИИ «стохастичность» означает, что генерация ответа или решения происходит не строго по одной определенной схеме, а с элементом случайности, базируясь на вероятностных моделях с внутренней системой веса входящих данных. Стохастичность особенно высока на ранних этапах обучения модели, когда веса задаются случайными числами.

смещений в процессе обучения моделей, а также пользовательских ошибок при обращении с ИИ-ассистентами. Для их минимизации следует учитывать природу базовых предубеждений на каждом этапе работы с данными.

Первый из них связан с формулировкой промптов. «Предвзятость формулировки» проявляется в качественных характеристиках запроса к LLM. С научной точки зрения каждый промпт, адресованный модели, является шагом тестирования гипотезы, иначе говоря, «актом исследования». Это означает, что неверно заданный вопрос, наводящие формулировки или чрезмерная детализация способны сузить поле анализа и склонить модель к подтверждению пользовательских предположений. Любой промпт, построенный как утверждение («да-нет» вопрос, указание на ожидаемый вывод), вызывает эффект подтверждения. LLM начинает искать не общие закономерности, а подтверждение предположения.

Эта тенденция многократно подтверждается экспериментально [11; 33]. Решением выступает практика «безопасных» промптов: формулировка нейтральных, конкретных и структурированных задач, избегание оценочной лексики, указание строгих ограничений и требования ссылаться на фрагменты данных, подтверждающих вывод.

Второй источник предвзятости в работе с LLM характерен для широкого поля количественных исследований. Ошибка в стандартизации данных (или ее отсутствии), проявляющаяся на этапе предварительной обработки и нормализации выбранных для обработки через языковую модель сведений. Хотя LLM готова работать даже с «грязными» и нерегулярными данными, у такого подхода всегда будут наблюдаться издержки, связанные с низкой воспроизводимостью исследования. Модель может структурировать материал по собственным латентным признакам, не всегда соответствующим исследовательским задачам. Переход от нерегулярных текстовых данных к JSON-структуре или другим формализованным методам отражения данных качественно меняет принципы работы с данными.

Следует учитывать, что преобразование данных также возможно с применением LLM, но только если исследователь последовательно подходит к процессу установления параметров фильтрации. Пропущенные категории или чрезмерная фрагментация текстовых данных ведут к потере смысловой связности и искажению выборки. В частности, нормализация текстовых данных имеет риск редуцирования контекста [28], что также влияет на достоверность выводов в социальных исследованиях. В качестве одного из способов улучшения качества структурирования можно рассматривать включение в процесс нормализации данных нескольких образцовых примеров данных со стандартизированной разметкой.

Третья фундаментальная предвзятость связана с избыточным или недостаточным контекстом. Как уже отмечалось выше, при структурировании данных неизбежна фильтрация, а следовательно, часть информации выпадает. К такой информации могут относиться как изначально избыточные метаданные, так и внутритекстовые «выбросы» (повторения, визуальные символы и т.д.). Слишком узкий контекст делает ответы модели ограниченными, тогда как избыточный

Рисунок 1. Пример структуризации данных для синтетического информационного сообщения в JSON-разметке

Figure 1. Example of Data Structuring for a Synthetic Information Message in JSON Format

```
{
  «raw_text»: «Сегодня состоялось рабочее совещание по вопросам развития.
  Приняты решения по оптимизации процессов. Планируется дальнейшее
  обсуждение.»,
  «structured_data»: {
    «event»: {
      «type»: «meeting»,
      «date»: «2023-11-15»,
      «topic»: «process_optimization»
    },
    «decisions»: [
      «process_optimization_approved»
    ],
    «next_steps»: [
      {
        «action»: «further_discussion»,
        «status»: «planned»
      }
    ],
    «metadata»: {
      «source_type»: «official_message»,
      «processing_date»: «2023-11-16»,
      «version»: «1.0»
    }
  }
}
```

Источник: составлено авторами с применением Deepseek
Source: compiled by the authors using Deepseek

контекст (слишком длинные или перегруженные блоки для анализа) ведет к размыванию смысла и потере фокуса [24; 25]. Несмотря на различную природу избыточного и недостаточного контекста, природа самого искажения остается одной – ложная интерпретация системных связей единицы данных по отношению ко всему массиву.

Потенциальным решением может быть поэтапный анализ, прописанный в стандартизирующем или аналитическом промпте:

- сначала извлечение сущностей (предварительная кластеризация текстовых единиц по темам / временным рамкам);

- затем фреймов (описание в аналитическом промпте разбивки каждой интересующей исследователя категории);

- в конце следует интерпретация (введение в промпт маркеров интерпретации).

Структурирование анализа позволяет контролировать объем и управляемость контекста. Возможен также дизайн исследований, подразумевающий многоэтапность и разбивку анализа на несколько промежуточных промптов для вычленения этапов и их перепроверки на предмет промежуточных ошибок. Стоит, однако, учитывать, что определение оптимальных объемов контекста и их влияние на искажение модели пока что остается открытым для дискуссии вопросом.

Четвертым искажением анализа данных на базе открытых LLM выступает «пользовательская ориентированность» публичных языковых моделей. Большинство из них демонстрирует адаптивное поведение, проявляющееся в подстраивании под роль, стиль и предположения пользователя. В результате обусловленная коммерческим дизайном особенность платформ становится серьезной издержкой в научной работе.

ИИ-ассистенты, будучи исследовательскими помощниками, также воспроизводят эпистемологические установки пользователя. Так как исследователь, формулируя параметры анализа (в т.ч. через *role prompting*²), фактически настраивает модель, он обеспечивает и трансляцию исследовательской позиции в алгоритмическую логику. Данная особенность работы LLM становится проблемой, если в работу модели не включаются корректирующие механизмы, которые будут также рассмотрены ниже. Отсюда возникает необходимость постоянной перепроверки выданных языковой модели и пересмотра дизайна исследования в сторону принципов, при котором ИИ – это не просто помощник, но «коллега, с которым ты спиришь».

Пятое искажение находится на стыке между проблемами контекста и дизайном языковых моделей. Его содержание можно выразить как потерю контекста из-за ограничений длины контекста (количества токенов). Любая пользовательская LLM в зависимости от формата и системы предоставления доступа неизбежно имеет как длину единичного запроса, так и общего содержания «диалога» [16; 36].

Базовым принципом борьбы с утратой контекста служит разбивка больших корпусов данных на части с поэтапным анализом. Избыточность объема также можно компенсировать формированием репрезентативной выборки по временным или тематическим параметрам, устанавливаемым на этапе предобработки или вычищения данных. Это позволит минимизировать искажения, вызванные усечением текста или «забыванием» модели раннее вводимых данных и переменных. Вместе с тем научное сообщество на текущий момент испытывает фундаментальные трудности с подбором конкретных методик калибровки длины контекста и оценки

² Стилизация содержания промпта, при которой LLM делегируется выполнение определенной роли (аналитик, копирайтер, дизайнер) или имитация работы какого-либо специалиста (по конкретным отраслям знания) при генерации ответов. Такой подход предопределяет стиль, тон и содержание ответов.

потери информации. В частности, это вызвано периодическим обновлением пользовательских условий и возможностей с каждой новой итерацией и сборкой общедоступных языковых моделей.

Таким образом, системная природа предвзятостей в работе LLM возникает на стыке человеческого выбора (постановка задачи, подбор корпуса) и машинных ограничений (архитектура модели, стохастичность вывода, лимит контекста).

Помимо представленных выше методов снижения рисков, для их преодоления требуются и более комплексные стратегии, опирающиеся на общенаучные принципы валидации и репликации. Их рассмотрение далее укажет на системные, а не точечные пути повышения достоверности выводов. Тем не менее рассмотренные вопросы предвзятости остаются открытыми и для дальнейшего теоретического осмысления. Попытки отыскать технические способы улучшить результат оставляют неотвеченным более философский вопрос о принципиальной соотносимости предвзятостей ИИ-агентов с человеческими искажениями.

Проблема валидации и репликации результатов

Фундаментальные проблемы и предвзятости текстового анализа с использованием открытых LLM оставляют во многом не решенной проблему соответствия подобных методов базовым научным принципам, связанным с возможностью оценки (валидации) и перепроверки (репликации) результатов. Ключевым показателем этой проблемы становится слабая воспроизводимость результатов промптов на разных моделях (а часто и тех же самых). Разногласия в интерпретациях, расхождения в извлекаемых связях, а также разная чувствительность к нюансам языка становятся их следствием [30]. В частности, невозможна полная репликация результатов при работе с открытыми LLM, так как каждая версия модели содержит разнородные весовые параметры, определяющие вероятностное распределение ответов [29].

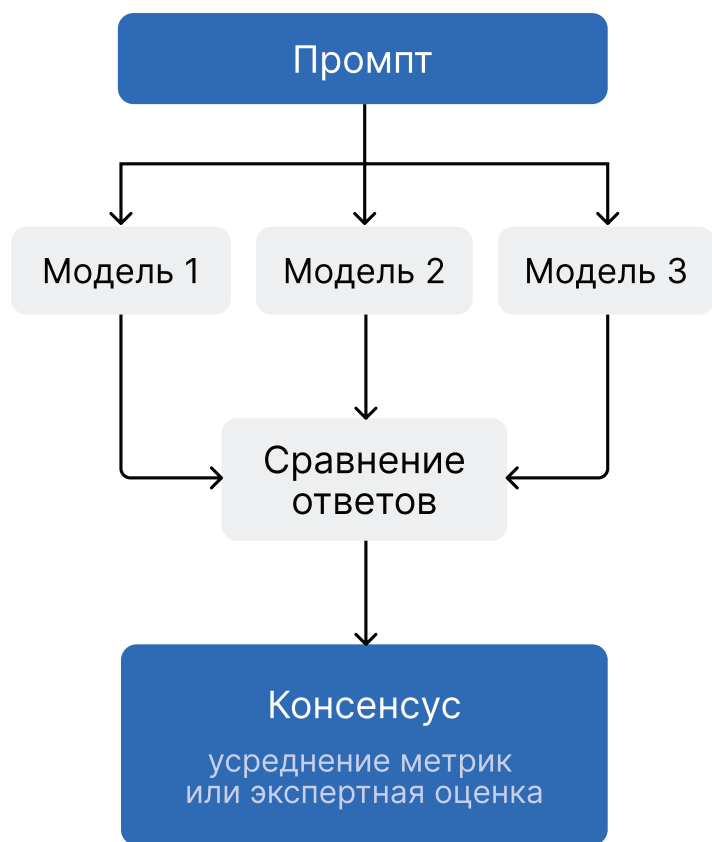
Тем не менее, учитывая потребность в обработке больших объемов данных при ограниченном количестве специалистов по большинству научных направлений, использование языковых моделей сохраняет свой потенциал – если не как базовый инструмент анализа, то как способ предпроверки и моделирования гипотез. В этом смысле любая проверка и перепроверка результатов, полученных с помощью LLM, должна рассматриваться в категориях приближенной, а не строгой воспроизводимости.

Вместе с тем возможна и общая минимизация проблем предвзятости. Ключевым способом является проверка устойчивости и согласованности выводов, полученных от разных моделей и при разных формулировках промптов, через многоуровневую процедуру триангуляции. Она включает в себя проверку альтернативных объяснений, тестирование на контрпримерах и использование связки моделей. Такая «триангуляция» позволяет частично компенсировать встроенные предубеждения и случайные ошибки вывода.

Триангуляция в работе с LLM предполагает согласование нескольких независимых источников вывода для получения более устойчивого результата. На практике возможно применение трех основных схем триангуляции:

Триангуляция моделей. Один и тот же промпт подается нескольким моделям. Их ответы сравниваются по содержанию, выявленным структурам данных или интерпретационным маркерам. Консенсус формируется либо через усреднение метрик, либо через внешнюю экспертную оценку совпадений и расхождений. Такой подход минимизирует влияние архитектурных и обучающих различий моделей.

Рисунок 2. Принципы триангуляции моделей
Figure 2. Principles of Model Triangulation



Источник: составлено авторами
Source: compiled by the authors

Триангуляция промптов. Используются несколько вариантов формулировок одного и того же исследовательского промпта. Модель выполняет серию параллельных рассуждений, а исследователь анализирует устойчивость получаемых выводов.

Совпадение ключевых элементов и среднего результата при разных формулировках указывает на вероятную устойчивость вывода, тогда как расхождения фиксируют проблемные зоны.

Рисунок 3. Принципы триангуляции промптов
Figure 3. Principles of Prompt Triangulation



Источник: составлено авторами
Source: compiled by the authors

Триангуляция данных. Проверяется устойчивость выводов при изменении корпуса данных. К примеру, происходит изменение временных срезов или категорий подмножеств в датасете. Такой вид триангуляции позволяет проверить, сохраняются ли выявляемые моделью закономерности при варьировании контекста. Иначе говоря, устойчивые корреляции отделяются от случайных совпадений в рамках конкретного корпуса.

Рисунок 4. Принципы триангуляции данных
Figure 4. Principles of Data Triangulation



Источник: составлено авторами
Source: compiled by the authors

Все три разновидности триангуляции позволяют повысить контроль вариативности результатов, с фокусом на выявлении границ колебаний выдач при поиске общих тенденций. Однако основным вкладом триангуляции является повышение валидности выводов без реального вклада в их реплицируемость. Для повышения репликации требуется рассмотрение дополнительных инструментов «вмешательства» в процесс анализа.

Среди подобных подходов особое внимание следует уделить контекстному обучению (in-context learning) как способу калибровки модели в пределах конкретной исследовательской задачи. Формулирование промпта и предоставление контекста становятся аналогом обучения на малом наборе данных в традиционном машинном обучении [26]. Здесь важна спецификация инструкций, форматов вывода и ограничений, что уменьшает интерпретационную вариативность.

В контексте архитектуры LLM контекстное обучение может обеспечиваться через внедрение «модулей индукции / внимания» (deduction heads), т.е. специализированных «слоев», отвечающих за воспроизведение закономерностей в последовательностях данных. Их использование требует критической оценки степени надежности модели, т.к. внутренние логические механизмы моделей не полностью прозрачны, а значит, прямые последствия внедрения подобных модулей не всегда можно отследить.

Укрепление эффекта модулей преимущественно обеспечивается через использование малого набора размеченных данных, из которых система должна вывести закономерность. При использовании пользовательской LLM этот принцип реализуется через контекстное обучение внутри промпта без изменения параметров модели. Подобный эффект возможно достигнуть в рамках нескольких стратегий:

В промпт включаются несколько пар «пример → ответ», иллюстрирующих требуемую логику решения (контекстная демонстрация). Модель фиксирует эту схему и применяет ее к новому примеру. Таким образом, заданный контекст выступает аналогом «временной» обучающей выборки.

Контекстная информация подается в виде правил или критериев оценки («оцени по шкале от 1 до 5», «выделяй только явные упоминания»). Модель учится следовать этим ограничениям (инструкциям) внутри текущего контекста, формируя шаблон поведения без дополнительного обучения.

Результаты первой выдачи модели анализируются исследователем, после чего промпт корректируется: уточняются термины, формат, исключаются неудачные параметры. Каждый последующий запрос выступает как дополнительный шаг (итеративного) контекстного обучения, повышающего воспроизводимость результатов.

Модель вводится в определенную исследовательскую роль («ты анализируешь политические нарративы»). Это создает устойчивый (ролевой) контекст, усиливающий согласованность ответов и снижающий произвольность интерпретаций.

С практической точки зрения контекстное обучение с малым числом примеров можно рассматривать как реализацию принципа, при котором модель не переобучается на уровне параметров, а временно перестраивает систему вероятностных связей внутри контекста задачи. Это делает процесс гибким, но обеспечивает реплицируемость лишь в пределах текущей контекстной конфигурации модели. Тем не менее внедрение модулей внимания позволяет обеспечить частичную воспроизводимость через сопоставление, контекстуализацию и постепенное уточнение.

Отдельно стоит упомянуть, что каждая стратегия вводит несколько формальных требований к принципам работы с моделью. К ним относятся: необходимость фиксировать версии моделей, точные промпты, даты запусков и примеры исходных выводов. В совокупности это создает основу для научной репликации в рамках вероятностных ограничений.

Триангуляция и контекстное обучение потенциально способны повысить показатели научности. Для установления широты применимости предложенных стратегий требуется множественная проверка каждой из стратегий на практике. Однако уже можно отметить, что преимуществом стратегий является изменение процедуры анализа без увеличения вычислительной мощности модели. Альтернативой становятся формально-логические принципы проверки. Таким образом, даже при высокой стохастичности выдач триангуляция и модули внимания создают некоторый методологический каркас для научного применения пользовательских LLM. Эти стратегии позволяют интегрировать вероятностные ответы алгоритма в исследовательские подходы, где важен не абсолютный результат, а возможность проверки степени его устойчивости.

Выводы

Современные LLM демонстрируют значительный потенциал в области социогуманитарных исследований, особенно на этапах предобработки или структурирования данных, а также при первичной проверке гипотез. При грамотной настройке их использование способно прежде всего снизить временные издержки для исследователя. LLM также расширяют аналитические возможности, позволяя выявлять скрытые закономерности в текстовых массивах. Однако эффективность научного применения LLM напрямую зависит от исследовательской осмотрительности, выраженной в понимании границ применимости инструмента, корректности постановки задач и качества исходных данных. Без этих условий использование моделей рискует превратиться в симуляцию научного исследования.

Анализ выявленных ограничений показывает, что фундаментальной проблемой применения LLM в научных целях остается техническая и эпистемологическая предвзятость. К первой категории относятся ошибки стандартизации, потеря контекста, стохастическая изменчивость выводов, проистекающие из структуры самих LLM. К эпистемологическим предвзятостям и ошибкам следует отнести влияние исследовательских ожиданий, обостряющееся в процессе применения LLM из-за пользовательской ориентированности моделей. Эти искажения, возникающие на стыке человеческого и машинного факторов, подрывают воспроизводимость и достоверность научных результатов. Поэтому ключевым направлением развития ориентированной на ИИ методологии становится разработка процедур валидации и репликации, адаптированных к вероятностной природе генеративных выдач.

Именно здесь значимую роль играют стратегии триангуляции и контекстного обучения, которые позволяют проверять устойчивость результатов модели. Переход от интуитивного использования LLM к системно обоснованной интеграции их в научный процесс требует выстраивания нового методологического каркаса, сочетающего технические, логические и общенаучные принципы воспроизводимости и верификации. LLM могут стать инструментом усиления научного анализа при

условии их критического применения как вероятностных, контекстно зависимых систем. В перспективе именно комбинированные подходы, объединяющие машинную обработку, человеческую интерпретацию и процедуры проверки достоверности, способны сформировать устойчивый стандарт научного применения LLM.

Список литературы

1. Ашихмин Е.Г., Левченко В.В., Селеткова Г.И. Опыт применения больших языковых моделей для анализа количественных социологических данных [Experience in Applying Large Language Models to Analyse Quantitative Sociological Data] // Вестник университета. 2024. № 11. С. 205–215. <https://doi.org/10.26425/1816-4277-2024-11-205-215>.
2. Базлутцкая М.М., Сытник А.Н. Трансмедийное вовлечение: фрейм-анализ цифровой дипломатии США в России при помощи искусственного интеллекта [Transmedia Engagement: AI-Driven Frame Analysis of the U.S. Digital Diplomacy in Russia] // Россия и мир: научный диалог. 2024. № 4(14). С. 63–85. [https://doi.org/10.53658/RW2024-4-4\(14\)-63-85](https://doi.org/10.53658/RW2024-4-4(14)-63-85).
3. Игнатъев А.Г. Этико-философские проблемы проектирования искусственного морального агента [Ethical and Philosophical Problems of Designing Artificial Moral Agent] // Этическая мысль. 2024. Т. 24. № 1. С. 87–100.
4. Коршунов А., Белобородов И., Бузун Н., Аванесов В., Пастухов Р., Чихрадзе К., Козлов И., Гомзин А., Андрианов И., Сысоев А., Ипатов С., Филоненко И., Чуприна К., Турдаков Д., Кузнецов С. Анализ социальных сетей: методы и приложения [Social Network Analysis: Methods and Applications] // Труды Института системного программирования РАН. 2014. Т. 26. № 1. С. 439–456. [https://doi.org/10.15514/ISPRAS-2014-26\(1\)-19](https://doi.org/10.15514/ISPRAS-2014-26(1)-19).
5. Соменков С.А. Искусственный интеллект: от объекта к субъекту? [Artificial Intelligence: from Object to Subject?] // Вестник университета имени О.Е.Кутафина. 2019. № 2(54). С. 75–85. <https://doi.org/10.17803/2311-5998.2019.54.2.075-085>.
6. Сысоев П.В., Филатов Е.М. ChatGPT в исследовательской работе студентов: запрещать или обучать? [ChatGPT in Students' Research Work: to Forbid or to Teach?] // Вестник Тамбовского университета. Серия: Гуманитарные науки. 2023. Т. 28. № 2. С. 276–301. <https://doi.org/10.20310/1810-0201-2023-28-2-276-301>.
7. Baiburin A., Berezkin Yu., Gromov A., Kovalenko K., Sokolov E., Kovalyova N., Anna Moskvitina A., Shirobokov I., Stanulevich N., Utekhin I., Boitsova O. Artificial Intelligence in the Social Sciences and Humanities // Forum for Anthropology and Culture. 2024. № 20. P. 11–60. <https://doi.org/10.31250/1815-8870-2024-20-20-11-60>.
8. Bazlutckaia M., Sytnik A., Tsvetkov T., Punchenko P. AI-Assisted Bias Detection of US Digital Diplomacy in Russia (2009–2023): A ChatGPT Approach // International Conference on Human-Computer Interaction. Cham: Springer Nature Switzerland, 2025. P. 189–209. https://doi.org/10.1007/978-3-031-93536-7_14.
9. Bojić L., Zagovora O., Zelenkauskaitė A., Vuković V., Čabarkapa M., Jerković S.V., Jovančević A. Comparing Large Language Models and Human Annotators in Latent Content Analysis of Sentiment, Political Leaning, Emotional Intensity and Sarcasm // Scientific Reports. 2025. Vol. 15(1). P. 11477. <https://doi.org/10.1038/s41598-025-96508-3>.
10. Braga M., Milanese G.C., Pasi G. Investigating Large Language Models' Linguistic Abilities for Text Preprocessing // arXiv preprint arXiv:2510.11482. 13 Oct. 2025. <https://doi.org/10.48550/arXiv.2510.11482>.
11. Brucks M., Toubia O. Prompt Architecture Induces Methodological Artifacts in Large Language Models // PLoS one. 2025. Vol. 20(4). P. e0319159. <https://doi.org/10.1371/journal.pone.0319159>.
12. Calderon R., Herrera F. And Plato Met ChatGPT: An Ethical Reflection on the Use of Chatbots in Scientific Research Writing, with a Particular Focus on the Social Sciences // Humanities and Social Sciences Communications. 2025. Vol. 12(1). P. 1–13. <https://doi.org/10.1057/s41599-025-04650-0>.
13. Chen K., He Z., Yan J., Shi T., Lerman K. How Susceptible are Large Language Models to Ideological Manipulation? // arXiv preprint arXiv:2402.11725. 18 Jun. 2024. <https://doi.org/10.48550/arXiv.2402.11725>.
14. Colonel J.T., Lin B. Word Clouds as Common Voices: LLM-Assisted Visualization of Participant-Weighted Themes in Qualitative Interviews // Proceedings of the Fourth Workshop on Bridging Human-Computer Interaction and Natural Language Processing (HCI+ NLP). 2025. P. 169–177. <https://doi.org/10.18653/v1/2025.hcinlp-1.14>.
15. De-Marcos L., Domínguez-Díaz A. LLM-Based Topic Modeling for Dark Web Q&A forums: A Comparative Analysis with Traditional Methods // IEEE Access. 2025. Vol. 13. P. 67159–67169. <https://doi.org/10.1109/ACCESS.2025.3560543>.

16. Fei W., Niu X., Zhou P., Hou L., Bai B., Deng L., Han W. Extending Context Window of Large Language Models Via Semantic Compression // Findings of the Association for Computational Linguistics: ACL. 2024. P. 5169–5181. <https://doi.org/10.18653/v1/2024.findings-acl.306>
17. Gao Q., Feng D. Deploying large language models for discourse studies: An exploration of automated analysis of media attitudes // PloS one. 2025. Vol. 20(1). P. e0313932. <https://doi.org/10.1371/journal.pone.0313932>.
18. Halterman A., Keith K. A. Codebook LLMs: Evaluating LLMs as Measurement Tools for Political Science Concepts // arXiv preprint arXiv:2407.10747. 9 Jan. 2025. <https://doi.org/10.48550/arXiv.2407.10747>.
19. Herbst P., Baars H. Accelerating Literature Screening for Systematic Literature Reviews with Large Language Models-Development, Application, and First Evaluation of a Solution // LWDA: Learning, Knowledge, Data, Analysis. 2023. P. 41–51.
20. Jenner S.E., Raidos D., Anderson E., Fleetwood S., Ainsworth B., Fox K., Kreppner J., Barker M. Using Large Language Models for Narrative Analysis: A Novel Application of Generative AI // Methods in Psychology. 2025. Vol. 12. P. 100183. <https://doi.org/10.1016/j.metip.2025.100183>.
21. Karjus A. Machine-Assisted Quantitizing Designs: Augmenting Humanities and Social Sciences with Artificial Intelligence // arXiv preprint arXiv:2309.14379. 20 Oct. 2024. <https://doi.org/10.48550/arXiv.2309.14379>.
22. Krippendorff, K. Computing Krippendorff's Alpha-Reliability. Philadelphia: University of Pennsylvania, 2011.
23. Kulkarni A., Alotaibi F., Zeng X., Wu L., Zeng T., Yao B.M., Liu M., Zhang Sh., Huang L., Zhou D. Scientific Hypothesis Generation and Validation: Methods, Datasets, and Future Directions // arXiv preprint arXiv:2505.04651. 6 May 2025. <https://doi.org/10.48550/arXiv.2505.04651>.
24. Kuribayashi T., Oseki Yo., Brassard A., Inui K. Scientific Context limitations make neural language models more human-like // arXiv preprint arXiv:2205.11463. 1 Nov. 2022. <https://doi.org/10.48550/arXiv.2205.11463>.
25. Li X., Tang H., Chen S., Wang Z., Chen R., Abramet M. Why Does In-Context Learning Fail Sometimes? Evaluating in-Context Learning on Open and Closed Questions // arXiv preprint arXiv:2407.02028. 2 Jul. 2024. <https://doi.org/10.48550/arXiv.2407.02028>.
26. Li Y. A Practical Survey on Zero-Shot Prompt Design for In-Context Learning // arXiv preprint arXiv:2309.13205. 22 Sep. 2023. https://doi.org/10.26615/978-954-452-092-2_069.
27. Liu J., Yang Ch., Yan Zh., Ma X., Peiet L. Leveraging Generative AI through Prompt Engineering for Corpus Construction and In-Depth Intelligent Interpretation of Ancient Texts // Digital Scholarship in the Humanities. 2025. Vol. 40, Issue 3. P. 846–862. <https://doi.org/10.1093/lc/fqaf043>.
28. Manohar K., Pillai L. G. What is Lost in Normalization? Exploring Pitfalls in Multilingual ASR Model Evaluations // Proceedings of the 2024 Conference on Empirical Methods in Natural Language Processing. 2024. P. 10864–10869. <https://doi.org/10.18653/v1/2024.emnlp-main.607>
29. Piper A., Bagga S. Using Large Language Models for Understanding Narrative Discourse // Proceedings of the The 6th Workshop on Narrative Understanding. 2024. P. 37–46. <https://doi.org/10.18653/v1/2024.wnu-1.4>.
30. Razavi A., Soltangheis M., Arabzadeh N., Salamat S., Zihayat M., Bagheri E. Benchmarking Prompt Sensitivity in Large Language Models // European Conference on Information Retrieval. Lecture Notes in Computer Science. Cham: Springer Nature Switzerland, 2025. P. 303–313. https://doi.org/10.1007/978-3-031-88714-7_29.
31. Rueda A., Hassan M.S., Perivolaris A., Teferra B.G., Samavi R., Rambhatla S., Wu Y., Zhang Y., Cao B., Sharma D., Krishnan S., Bhatet V. Understanding LLM Scientific Reasoning through Promptings and Model's Explanation on the Answers // arXiv preprint arXiv:2505.01482. 25 Jul. 2025. <https://doi.org/10.48550/arXiv.2505.01482>.
32. Sebastian R., Kottekkadan N.N., Thomas T.K., Niyas M. Generative AI Tools (ChatGPT*) in Social Science Research // Journal of Information, Communication and Ethics in Society. 2025. Vol. 23(2). P. 284–290. <https://doi.org/10.1108/JICES-10-2024-0145>.
33. Sun Y., Kok S. Investigating the Effects of Cognitive Biases in Prompts on Large Language Model Outputs // arXiv preprint arXiv:2506.12338. 14 Jun. 2025. <https://doi.org/10.48550/arXiv.2506.12338>.
34. Tao Y., Shen Q. Academic Discourse on ChatGPT in Social Sciences: A Topic Modeling and Sentiment Analysis of Research Article Abstracts // PloS one. 2025. Vol. 20(10). P. e0334331. <https://doi.org/10.1371/journal.pone.0334331>.
35. Törnberg P. Large Language Models Outperform Expert Coders and Supervised Classifiers at Annotating Political Social Media Messages // Social Science Computer Review. 2024. Vol. 43, Issue 6. <https://doi.org/10.1177/08944393241286471>.
36. Wang X., Salmani M., Omid P., Ren X., Rezagholizadeh M., Eshaghiet A. Beyond the Limits: A Survey of Techniques to Extend the Context Length in Large Language Models // arXiv preprint arXiv:2402.02244. 29 May 2024. <https://doi.org/10.48550/arXiv.2402.02244>.

37. Zabaleta M., Lehman J. Simulating Tabular Datasets through LLMs to Rapidly Explore Hypotheses about Real-World Entities // arXiv preprint arXiv:2411.18071. 27 Nov. 2024. <https://doi.org/10.48550/arXiv.2411.18071>.

Информация об авторах

КОЛОТАЕВ Юрий Юрьевич. Кандидат политических наук. Старший преподаватель кафедры европейских исследований Санкт-Петербургского государственного университета. <https://orcid.org/0000-0001-8372-1193>. Адрес: Российская Федерация, 199034, г. Санкт-Петербург, Университетская наб., 7-9. yury.kolotaev@mail.ru

БАЗЛУЦКАЯ Мария Михайловна. Кандидат политических наук. Исполнительный директор автономной некоммерческой научно-исследовательской организации «Координационная лаборатория» (АНО «Коллаборатория»). <https://orcid.org/0000-0002-7118-8632>. Адрес: Российская Федерация, 194064, г. Санкт-Петербург, пр-кт Раевского, д. 16, литера А, помещ. 5-н. m.bazlutckaya@gmail.com

Раскрытие информации о конфликте интересов

Авторы заявляют об отсутствии конфликта интересов.

Информация о статье

Поступила в редакцию: 16 ноября 2025 г. Одобрена после рецензирования: 10 февраля 2026 г. Принята к публикации: 20 февраля 2026 г. Опубликована: 1 марта 2026 г.

Авторы прочитали и одобрили окончательный вариант рукописи.

Информация о рецензировании

«Россия и мир: научный диалог» благодарит анонимных рецензентов за их вклад в рецензирование этой работы.

References

1. Ashikhmin E.G., Levchenko V.V., Seletkova G.I. Experience in Applying Large Language Models to Analyse Quantitative Sociological Data. Vestnik universiteta [Vestnik Universiteta]. 2024; 11:205–215 [In Russian]. <https://doi.org/10.26425/1816-4277-2024-11-205-215>.
2. Bazlutckaia M.M., Sytnik A.N. Transmedia Engagement: AI-Driven Frame Analysis of the U.S. Digital Diplomacy in Russia. Rossiya i mir: nauchnyj dialog [Russia & World: Sc. Dialogue]. 2024; (4):63–85 [In Russian]. [https://doi.org/10.53658/RW2024-4-4\(14\)-63-85](https://doi.org/10.53658/RW2024-4-4(14)-63-85).
3. Ignatev A.G. Ethical and Philosophical Problems of Designing Artificial Moral Agent. Eticheskaya mysl' [Ethical thought]. 2024; 24(1):87–100 [In Russian].
4. Korshunov A., Beloborodov I., Buzun N., Avanesov V., Pastukhov R., Chykhhradze K., Kozlov I., Gomzin A., Andrianov I., Sysoev A., Ipatov S., Filonenko I., Chuprina Ch., Turdakov D., Kuznetsov S. Social Network Analysis: Methods and Applications. Trudy Instituta sistemnogo programmirovaniya RAN [Proceedings of the Institute for System Programming of the RAS]. 2014; 26(1):439–456 [In Russian]. [https://doi.org/10.15514/ISPRAS-2014-26\(1\)-19](https://doi.org/10.15514/ISPRAS-2014-26(1)-19).
5. Somenkov S.A. Artificial Intelligence: from Object to Subject? Vestnik universiteta imeni O.E.Kutafina [Courier of Kutafin Moscow State Law University (MSAL)]. 2019; (2):75–85 [In Russian]. <https://doi.org/10.17803/2311-5998.2019.54.2.075-085>.
6. Sysoev, P.V., Filatov, E.M. ChatGPT in Students' Research Work: to Forbid or to Teach? Vestnik Tambovskogo universiteta. Seriya: Gumanitarnye nauki [Tambov University Review. Series: Humanities]. 2023; 28(2):276–301 [In Russian]. <https://doi.org/10.20310/1810-0201-2023-28-2-276-301>.
7. Baiburin A., Berezkin Yu., Gromov A., Kovalenko K., Sokolov E., Kovalyova N., Anna Moskvitina A., Shirobokov I., Stanulevich N., Utekhin I., Boitsova O. Artificial Intelligence in the Social Sciences and Humanities. Forum for Anthropology and Culture. 2024; 20:11–60 [In English]. <https://doi.org/10.31250/1815-8870-2024-20-20-11-60>.

8. Bazlutckaia M., Sytnik A., Tsvetkov T., Punchenko P. AI-Assisted Bias Detection of US Digital Diplomacy in Russia (2009–2023): A ChatGPT Approach. In International Conference on Human-Computer Interaction. Cham: Springer Nature Switzerland, 2025:189–209 [In English]. https://doi.org/10.1007/978-3-031-93536-7_14.
9. Bojić L., Zagovora O., Zelenkauskaitė A., Vuković V., Čabarkapa M., Jerković S.V., Jovančević A. Comparing Large Language Models and Human Annotators in Latent Content Analysis of Sentiment, Political Leaning, Emotional Intensity and Sarcasm. Scientific Reports. 2025; 15(1):11477 [In English]. <https://doi.org/10.1038/s41598-025-96508-3>.
10. Braga M., Milanese G.C., Pasi G. Investigating Large Language Models' Linguistic Abilities for Text Preprocessing. arXiv preprint arXiv:2510.11482. 13 Oct. 2025 [In English]. <https://doi.org/10.48550/arXiv.2510.11482>.
11. Brucks M., Toubia O. Prompt Architecture Induces Methodological Artifacts in Large Language Models. PLoS one. 2025; 20(4):e0319159 [In English]. <https://doi.org/10.1371/journal.pone.0319159>.
12. Calderon R., Herrera F. And Plato Met ChatGPT: An Ethical Reflection on the Use of Chatbots in Scientific Research Writing, with a Particular Focus on the Social Sciences. Humanities and Social Sciences Communications. 2025; 12(1):1–13 [In English]. <https://doi.org/10.1057/s41599-025-04650-0>.
13. Chen K., He Z., Yan J., Shi T., Lerman K. How Susceptible are Large Language Models to Ideological Manipulation? arXiv preprint arXiv:2402.11725. 18 Jun. 2024 [In English]. <https://doi.org/10.48550/arXiv.2402.11725>.
14. Colonel J.T., Lin B. Word Clouds as Common Voices: LLM-Assisted Visualization of Participant-Weighted Themes in Qualitative Interviews, In Proceedings of the Fourth Workshop on Bridging Human-Computer Interaction and Natural Language Processing (HCI+ NLP). 2025:169–177 [In English]. <https://doi.org/10.18653/v1/2025.hcinlp-1.14>.
15. De-Marcos L., Domínguez-Díaz A. LLM-Based Topic Modeling for Dark Web Q&A forums: A Comparative Analysis with Traditional Methods. IEEE Access. 2025; 13:67159–67169 [In English]. <https://doi.org/10.1109/ACCESS.2025.3560543>.
16. Fei W., Niu X., Zhou P., Hou L., Bai B., Deng L., Han W. Extending Context Window of Large Language Models Via Semantic Compression. Findings of the Association for Computational Linguistics: ACL. 2024:5169–5181 [In English]. <https://doi.org/10.18653/v1/2024.findings-acl.306>
17. Gao Q., Feng D. Deploying large language models for discourse studies: An exploration of automated analysis of media attitudes. PLoS one. 2025; 20(1):e0313932 [In English]. <https://doi.org/10.1371/journal.pone.0313932>.
18. Halterman A., Keith K.A. Codebook LLMs: Evaluating LLMs as Measurement Tools for Political Science Concepts. arXiv preprint arXiv:2407.10747. 9 Jan. 2025 [In English]. <https://doi.org/10.48550/arXiv.2407.10747>.
19. Herbst P., Baars H. Accelerating Literature Screening for Systematic Literature Reviews with Large Language Models-Development, Application, and First Evaluation of a Solution. LWDA: Learning, Knowledge, Data, Analysis. 2023:41–51 [In English].
20. Jenner S.E., Raidos D., Anderson E., Fleetwood S., Ainsworth B., Fox K., Kreppner J., Barker M. Using Large Language Models for Narrative Analysis: A Novel Application of Generative AI. Methods in Psychology. 2025; 12:100183 [In English]. <https://doi.org/10.1016/j.metip.2025.100183>.
21. Karjus A. Machine-Assisted Quantizing Designs: Augmenting Humanities and Social Sciences with Artificial Intelligence. arXiv preprint arXiv:2309.14379. 20 Oct. 2024 [In English]. <https://doi.org/10.48550/arXiv.2309.14379>.
22. Krippendorff K. Computing Krippendorff's Alpha-Reliability. Philadelphia: University of Pennsylvania, 2011 [In English].
23. Kulkarni A., Alotaibi F., Zeng X., Wu L., Zeng T., Yao B.M., Liu M., Zhang Sh., Huang L., Zhou D. Scientific Hypothesis Generation and Validation: Methods, Datasets, and Future Directions. arXiv preprint arXiv:2505.04651. 6 May 2025 [In English]. <https://doi.org/10.48550/arXiv.2505.04651>.
24. Kuribayashi T., Oseki Yo., Brassard A., Inui K. Scientific Context limitations make neural language models more human-like. arXiv preprint arXiv:2205.11463. 1 Nov. 2022 [In English]. <https://doi.org/10.48550/arXiv.2205.11463>.
25. Li X., Tang H., Chen S., Wang Z., Chen R., Abramet M. Why Does In-Context Learning Fail Sometimes? Evaluating in-Context Learning on Open and Closed Questions. arXiv preprint arXiv:2407.02028. 2 Jul. 2024 [In English]. <https://doi.org/10.48550/arXiv.2407.02028>.
26. Li Y. A Practical Survey on Zero-Shot Prompt Design for In-Context Learning. arXiv preprint arXiv:2309.13205. 22 Sep. 2023 [In English]. https://doi.org/10.26615/978-954-452-092-2_069.
27. Liu J., Yang Ch., Yan Zh., Ma X., Peiet L. Leveraging Generative AI through Prompt Engineering for Corpus Construction and In-Depth Intelligent Interpretation of Ancient Texts. Digital Scholarship in the Humanities. 2025; 40(3):846–862 [In English]. <https://doi.org/10.1093/llc/fqaf043>.
28. Manohar K., Pillai L. G. What is Lost in Normalization? Exploring Pitfalls in Multilingual ASR Model Evaluations // Proceedings of the 2024 Conference on Empirical Methods in Natural Language Processing. 2024. P. 10864–10869. <https://doi.org/10.18653/v1/2024.emnlp-main.607>
29. Piper A., Bagga S. Using Large Language Models for Understanding Narrative Discourse/ In Proceedings of the The 6th Workshop on Narrative Understanding. 2024:37–46 [In English]. <https://doi.org/10.18653/v1/2024.wnu-1.4>.
30. Razavi A., Soltangheis M., Arabzadeh N., Salamat S., Zihayat M., Bagheri E. Benchmarking Prompt Sensitivity in Large Language Models. In European Conference on Information Retrieval. Lecture Notes in Computer Science. Cham: Springer Nature Switzerland, 2025:303–313 [In English]. https://doi.org/10.1007/978-3-031-88714-7_29.
31. Rueda A., Hassan M.S., Perivolaris A., Teferra B.G., Samavi R., Rambhatla S., Wu Y., Zhang Y., Cao B., Sharma D., Krishnan S., Bhatet V. Understanding LLM Scientific Reasoning through Prompts and Model's Explanation on the Answers. arXiv preprint arXiv:2505.01482. 25 Jul. 2025 [In English]. <https://doi.org/10.48550/arXiv.2505.01482>.
32. Sebastian R., Kottekkadan N.N., Thomas T.K., Niyas M. Generative AI Tools (ChatGPT*) in Social Science Research. Journal of Information, Communication and Ethics in Society. 2025; 23(2):284–290 [In English]. <https://doi.org/10.1108/JICES-10-2024-0145>.
33. Sun Y., Kok S. Investigating the Effects of Cognitive Biases in Prompts on Large Language Model Outputs. arXiv preprint arXiv:2506.12338. 14 Jun. 2025 [In English]. <https://doi.org/10.48550/arXiv.2506.12338>.
34. Tao Y., Shen Q. Academic Discourse on ChatGPT in Social Sciences: A Topic Modeling and Sentiment Analysis of Research Article Abstracts. PLoS one. 2025; 20(10):e0334331 [In English]. <https://doi.org/10.1371/journal.pone.0334331>.
35. Törnberg P. Large Language Models Outperform Expert Coders and Supervised Classifiers at Annotating Political Social Media Messages. Social Science Computer Review. 2024; 43(6) [In English]. <https://doi.org/10.1177/08944393241286471>.
36. Wang X., Salmani M., Omid P., Ren X., Rezagholizadeh M., Eshaghiet A. Beyond the Limits: A Survey of Techniques to Extend the Context Length in Large Language Models. arXiv preprint arXiv:2402.02244. 29 May 2024 [In English]. <https://doi.org/10.48550/arXiv.2402.02244>.
37. Zabaleta M., Lehman J. Simulating Tabular Datasets through LLMs to Rapidly Explore Hypotheses about Real-World Entities. arXiv preprint arXiv:2411.18071. 27 Nov. 2024 [In English]. <https://doi.org/10.48550/arXiv.2411.18071>.

About the authors

Yury Yu. KOLOTAEV. CandSc. (Polit.). Assistant Professor, European Studies Department, School of International Relations, Saint-Petersburg State University. <https://orcid.org/0000-0001-8372-1193>. Address: 7-9, Universitetskaya nab., St. Petersburg, 199034, Russian Federation. yury.kolotaev@mail.ru.

Mariya M. BAZLUTCKAYA. CandSc. (Polit.). Executive Director, Autonomous Non-Commercial Research Organisation "Coordination Lab" (ANO Colaboratoria). <https://orcid.org/0000-0002-7118-8632>. Address: Room 5-N., 16A, Rayevskiy ave., St. Petersburg, 194064, Russian Federation. m.bazlutckaya@gmail.com

Contribution of the authors

The authors declare no conflicts of interests.

Article info

Received: November 16, 2025. Approved after review: February 10, 2026. Accepted for publication: February 20, 2026. Published: March 1, 2026.

The authors have read and approved the final manuscript.

Peer review info

«Russia & World: Scientific Dialogue» thanks the anonymous reviewers for their contribution to the peer review of this work.