ДИСКУРС. 2025. Т. 11, № 1. С. 52–70
DISCOURSE. 2025, vol. 11, no. 1, pp. 52–70

Оригинальная статья
УДК 316; 004.8
http://doi.org/10.32603/2412-8562-2025-11-1-52-70

Использование генеративного искусственного интеллекта для социологических исследований

***Владимир Евгеньевич Драч*[1]✉*, Юлия Владимировна Торкунова*[2]**

[1, 2]*Сочинский государственный университет, Сочи, Россия*
[2]*Казанский государственный энергетический университет, Казань, Россия*
[1]✉*vladimir@drach.pro, https://orcid.org/0009-0009-6280-3160*
[2]*torkynova@mail.ru, https://orcid.org/0000-0001-7642-6663*

Введение. В статье рассматривается использование генеративного искусственного интеллекта (ГИИ) в социологических исследованиях. Актуальность темы определяется возрастанием интереса к применению новых технологий для повышения эффективности и точности исследований в области социальных наук. ГИИ предоставляет новые возможности для сбора, обработки и анализа данных, что может существенно изменить традиционные подходы в социологии.

Методология и источники. Исследование базируется на анализе доступных публикаций и экспериментальных данных, полученных в ходе общения с социологами, использующими ГИИ в своих проектах. В работе рассматриваются методики генерации описаний, анализа изображений и генерации синтетических изображений с использованием алгоритмов машинного обучения. Акцент сделан на конкретных случаях применения ГИИ в социологических исследованиях, а также на примерах успешных проектов.

Результаты и обсуждение. Результаты исследования показывают, что использование ГИИ позволяет значительно ускорить процесс обработки данных и повысить их качество. Выявлены новые паттерны и тенденции в социологических исследованиях, благодаря чему ученые могут получать более точные и обоснованные выводы. Обсуждаются также этические аспекты, связанные с использованием ГИИ, такие как вопросы конфиденциальности и алгоритмической предвзятости.

Заключение. Генеративный искусственный интеллект представляет собой мощный инструмент, способный трансформировать социологические исследования. Несмотря на существующие вызовы, он открывает новые горизонты для сбора и анализа данных, способствуя более глубокому пониманию социальных процессов и явлений. Важно продолжать исследовать возможности и ограничения ГИИ для развития социологической науки.

Ключевые слова: генеративный искусственный интеллект, социологические исследования, анкеты, машинное обучение, этические аспекты

Для цитирования: Драч В. Е., Торкунова Ю. В. Использование генеративного искусственного интеллекта для социологических исследований // ДИСКУРС. 2025. Т. 11, № 1. С. 52–70. DOI: 10.32603/2412-8562-2025-11-1-52-70.

ДИСКУРС. 2025. Т. 11, № 1. С. 52–70
DISCOURSE. 2025, vol. 11, no. 1, pp. 52–70

Original paper

Implementation of Generative Artificial Intelligence in Sociological Research

***Vladimir E. Drach*[1]✉*, Yulia V. Torkunova*[2]**

[1, 2]*Sochi State University, Sochi, Russia*
[2]*Kazan State Energy University, Kazan, Russia*
[1]✉*vladimir@drach.pro, https://orcid.org/0009-0009-6280-3160*
[2]*torkynova@mail.ru, https://orcid.org/0000-0001-7642-6663*

Introduction. This article discusses the use of generative artificial intelligence (GAI) in sociological research. The relevance of the topic is determined by the increasing interest in applying new technologies to enhance the efficiency and accuracy of research in social sciences. GAI provides new opportunities for data collection, processing, and analysis, which can significantly change traditional approaches in sociology.
Methodology and sources. The research is based on an analysis of available publications and experimental data obtained during discussions with sociologists using GAI in their projects. The paper examines methodologies for generating surveys, processing respondents' answers, and analyzing big data using machine learning algorithms. The focus is on specific cases of GAI applications in sociological research, as well as examples of successful projects.
Results and discussion. The results of the study demonstrate that the use of GAI allows for significantly accelerating the data processing process and enhancing the quality of the data. New patterns and trends in sociological research have been identified, enabling researchers to draw more accurate and justified conclusions. Ethical aspects related to the use of GAI are also discussed, such as issues of confidentiality and algorithmic bias.
Conclusion. Generative artificial intelligence represents a powerful tool capable of transforming sociological research. Despite existing challenges, it opens new horizons for data collection and analysis, fostering a deeper understanding of social processes and phenomena. It is important to continue exploring the possibilities and limitations of GAI for the advancement of sociological science.

Keywords: generative artificial intelligence, sociological research, surveys, machine learning, ethical aspects

For citation: Drach, V.E. and Torkunova, Yu.V. (2025), “Implementation of Generative Artificial Intelligence in Sociological Research”, *DISCOURSE*, vol. 11, no. 1, pp. 52–70. DOI: 10.32603/2412-8562-2025-11-1-52-70 (Russia).

Введение. Взрывной рост возможностей и популярности нейросетей после 2020 г. вызвал огромный интерес к генеративному искусственному интеллекту (Generative artificial intelligence, generative AI, GenAI или GAI), который широко обсуждается в СМИ и вызывает целый спектр эмоций – от восторга до серьезных опасений [1]. Некоторые исследователи предупреждают о возможных катастрофических последствиях, включая угрозу для человечества, в то время как другие указывают на риск создания вредоносного контента и экологические издержки [2–4]. В академической среде преподаватели сталкиваются с проблемой GAI, когда студенты выдают результаты работы нейросетей за свои. Важно рассмотреть влияние этих технологий на социологические исследования, поскольку GAI может способствовать прогрессу в анализе текста, экспериментах и моделировании.

В настоящей работе представлены исследования на стыке социальных наук и информационных технологий для демонстрации применения GAI в социологии. Рассматриваются две основные методологические области: анализ текста и изображений. Во-первых, большие языковые модели (LLM) улучшают социологический анализ текстовых данных, предлагая более точную классификацию документов и возможности для экспериментов. Рассматривается, как эти модели могут быть адаптированы для качественного анализа текста, позволяя выполнять более интерактивное и интерпретативное исследование. Во-вторых, рассматривается, как современные модели GAI, такие как Fooocus, DALL·E, Perplexity и ChatGPT, могут анализировать и генерировать изображения. Показано, как эти модели могут кодировать визуальные данные и создавать синтетические изображения для экспериментов. Сделано предположение, что GAI дополняет, но не заменяет существующие социологические методы.

Несмотря на потенциал GAI, важно использовать эти технологии осторожно. Сложность интерпретации выходных данных, непрозрачность данных для обучения и возможная ненадежность результатов вызывают вопросы о воспроизводимости и приватности. Эти модели поднимают также этические проблемы, особенно в исследованиях, связанных с персональными данными. Рассматривается вопрос «предвзятости» GAI и его влияние на социальные исследования, а также необходимость разработки моделей, более подходящих для социальных наук.

Методология и источники. Методология исследования основана на интеграции количественных и качественных подходов, что позволяет более глубоко анализировать применение генеративного искусственного интеллекта в социологических исследованиях. В рамках работы используются большие языковые модели для анализа текстовых данных, что обеспечивает высокую точность классификации и аннотирования. Эти модели, обученные на обширных текстовых корпусах, позволяют исследователям взаимодействовать с данными в диалоговом формате, задавая вопросы и получая интерпретации, что значительно расширяет возможности качественного анализа. Кроме того, применяется мультимодальный подход, который включает анализ изображений с использованием современных моделей компьютерного зрения, способных извлекать информацию из визуальных данных и генерировать синтетические изображения.

Источниками данных для исследования служат как научные публикации, так и открытые наборы данных, содержащие текстовые и визуальные материалы. Это позволяет обеспечить разнообразие и достоверность информации. Важным аспектом является также учет этических вопросов, связанных с использованием ГИИ, включая предвзятость моделей и защиту персональных данных. Таким образом, методология статьи сочетает в себе современные технологии анализа с критическим осмыслением их применения в социологии, что способствует выявлению как потенциала, так и рисков использования ГИИ в данной области.

Результаты и обсуждение

Анализ текстов с применением нейросетей

Классификация, аннотации и методологические рекомендации. Большие языковые модели (англ. Large language model, LLM) или сокращенно БЯМ – это тип искусственного интеллекта, который специализируется на обработке и генерации натурального (человеческого)

языка. Они представляют собой машинные модели, обученные на огромных объемах текстовых данных, что позволяет им «понимать» вопросы и генерировать ответ на языке человека. Ранние версии моделей не могли обрабатывать объемные входные данные и генерировали короткие последовательности текста. Серьезный прорыв в преодолении этих ограничений случился около 2023 г. с появлением ChatGPT версии 3 и подобных. Скачкообразный рост качества ответов современных GAI, как правило, обуславливается тремя факторами: развитие вычислительных мощностей, разработка «трансформеров» – нейронных сетей, способных обрабатывать длинные тексты, и доступность обширных текстов для обучения. Современные модели называют «большими» из-за огромного объема текста, на которых они обучались, и миллиардов параметров, используемых для их обучения. Например, модель Llama 3 обучена на 15 трл токенов и содержит около 70 млрд параметров.

Наиболее актуальные применения больших языковых моделей для социологических исследований связаны с вычислительной социологией. Социологи используют эти модели для классификации текстов – задачи, при которой модель обучается предсказывать категории для новых текстов на основе обучающих данных. Этот метод часто применяется для обработки больших объемов текстов, таких как статьи или посты в соцсетях. БЯМ превосходят традиционные методы машинного обучения в задачах классификации текстов благодаря предварительной подготовке на огромных объемах данных, что позволяет им адаптироваться к новым задачам без дополнительного обучения.

БЯМ также способны ускорить процесс аннотирования данных. Недавние исследования показывают, что БЯМ могут справляться с задачей аннотирования, достигая качества на уровне с людьми, значительно снижая затраты. Однако, как и люди, модели могут ошибаться или неверно интерпретировать инструкции. Комбинирование работы людей и БЯМ помогает улучшить процесс аннотирования и снизить его стоимость.

Применение БЯМ выходит за рамки классификации текстов. Универсальность позволяет использовать их в различных задачах анализа, например в тематическом моделировании или в качестве вероятностных моделей языка. Способность работать с запросами дает возможность ученым создавать индивидуальные решения без глубоких знаний в области компьютерных наук, что ускоряет прототипирование и исследовательский анализ. БЯМ могут использоваться для выявления тем в текстах или проверки различных схем кодирования, а также для подтверждающего анализа. Хотя БЯМ не всегда являются идеальным инструментом для всех задач анализа текста, их универсальность создает возможности для методологических инноваций.

Вычислительный анализ содержания в диалоговом формате. БЯМ обещают усилить использование вычислительных методов в качественных исследованиях. Качественные исследователи требуют более строгих, прозрачных и гибких процедур контент-анализа и предлагают использовать вычислительные методы для достижения этих целей. БЯМ выгодны для качественного анализа по сравнению с традиционными методами. Стандартные подходы, такие как тематическое моделирование, лучше работают с большими объемами данных, а качественные данные, такие как интервью или полевые заметки, содержат несколько элементов, которые сложно анализировать без потери информации. БЯМ, обладая обширными языковыми представлениями, могут напрямую анализировать такие тексты без пред-

варительной обработки. Кроме того, их гибкость позволяет адаптировать запросы под конкретные данные.

Самым значимым новшеством является возможность диалогового вычислительного анализа контента. БЯМ позволяют исследователям «задавать вопросы» своим данным. Например, модель может резюмировать интервью, выделять темы и анализировать паттерны. Это отличается от традиционного качественного анализа, который начинается с тщательного чтения документов и создания тематических категорий. БЯМ изменяют этот процесс, выполняя итерации между вычислительным и интерпретативным подходами, позволяя распределять интерпретацию между людьми и машинами. Исследователь интерпретирует ответы модели и может оспаривать или уточнять их с помощью дальнейших запросов.

Эта диалоговая методика вызывает новые вопросы. БЯМ могут использовать внешние знания, полученные на этапе обучения, что усложняет понимание их интерпретации данных. Важно изучить «предустановленные предположения», которые влияют на ответы моделей. Например, модели могут отражать общественные когнитивные схемы и даже воспроизводить стереотипы. Для решения этих вопросов модели адаптируют с помощью обратной связи от людей, направляя их к определенным ценностям.

Исследователи должны экспериментировать с формулировкой запросов, чтобы понять, как модели реагируют, и сравнивать результаты с ручными примерами для валидации. Важно учитывать не только свою позицию в исследовании, но и роль модели, анализируя в ней заложенные взгляды. Модели не следует воспринимать как объективные или нейтральные; их выводы зависят от данных, на которых они обучены.

Современные качественные проекты часто включают большое количество интервью или наблюдений и комбинируют их с другими источниками данных для смешанного анализа. БЯМ помогают качественным исследователям более эффективно и масштабно анализировать данные, сохраняя при этом интерпретативные нюансы, которые могут быть утрачены при использовании традиционных методов.

Применение мультимодальных моделей для анализа и генерации изображений

Представление визуальных данных. Помимо работы с текстами, ИИ открывает новые возможности в социологических исследованиях изображений и визуальных данных. Параллельно с развитием языкового моделирования компьютерное зрение значительно продвинулось благодаря увеличению вычислительных мощностей и разработке сверточных нейронных сетей (СНС), которые могут распознавать сложные паттерны в изображениях, такие как контуры, текстуры и оттенки. Как и БЯМ, модели компьютерного зрения проходят предварительное обучение на общих наборах изображений, а затем адаптируются для конкретных задач. Эти модели успешно применялись в социологических исследованиях, например, для выявления изображений, связанных с протестами в социальных сетях Китая, или для анализа черт городских районов, таких как мусор на улицах по данным Google Street View.

Новые ИИ являются мультимодальными, это означает, что они обучены на данных разных типов, например текстах и изображениях, и могут использовать их как входные и выходные данные. Это открывает уникальные возможности для работы с визуальными данными, так как текстовые запросы можно применять для анализа изображений без дополнительной настройки.

ДИСКУРС. 2025. Т. 11, № 1. С. 52–70
DISCOURSE. 2025, vol. 11, no. 1, pp. 52–70

Пример, связанный с социальными движениями, показывает, как ИИ умеет извлекать информацию из изображений. Так, можно передать в GAI фотографию антивоенного протеста в США (College students from various nearby schools march down Commonwealth Avenue in Boston on Oct. 16, 1965 to attend rally on Boston Common protesting U.S. involvement in Vietnam [AP Photo/Frank C. Curtin, File]) [5] и провести ее анализ (рис. 1).

Рис. 1. Исходная фотография для анализа
Fig. 1. Original photograph for analysis

Для анализа фотографии был использован инструментарий ChatGPT-4, модель o1. Фрагмент диалога с ChatGPT-4 приведен ниже.

Авторы: опиши фотографию.

ChatGPT: на изображении показана группа людей, участвующих в антивоенном марше или демонстрации. Они держат плакат с надписью «End the War in Vietnam» («Прекратите войну во Вьетнаме»), что указывает на протест против участия США во вьетнамской войне. Эта фотография, вероятно, была сделана в 1960-е или начале 1970-х годов, когда движение против войны во Вьетнаме достигло своего пика в США. Люди на фото одеты в одежду того времени, что соответствует контексту массовых протестов той эпохи.

Авторы: перечисли объекты на фотографии, выбери категории описания и конвертируй ответ в формат JSON.

ChatGPT:

{
 "scene": "Демонстрация против войны во Вьетнаме",
 "banner_text": "End the War in Vietnam",
 "participants": [
  {
   "gender": "мужчина",
   "approximate_age": "молодой",
   "clothing": "пальто, галстук",
   "position": "в центре, держит плакат"
  },
  {
   "gender": "мужчина",

```
      "approximate_age": "молодой",
      "clothing": "костюм, очки",
      "position": "слева от центра, держит плакат"
    },
    {
      "gender": "женщина",
      "approximate_age": "молодая",
      "clothing": "пальто, головной убор",
      "position": "справа, держит бумагу"
    },
    {
      "gender": "мужчина",
      "approximate_age": "молодой",
      "clothing": "куртка",
      "position": "слева, идет в стороне от основной группы"
    },
    {
      "gender": "мужчина",
      "approximate_age": "молодой",
      "clothing": "очередной костюм",
      "position": "в центре, рядом с другими участниками"
    }
  ],
  "objects": [
    {
      "type": "плакаты",
      "text": "Прекратите войну во Вьетнаме"
    },
    {
      "type": "машины",
      "approximate_year": "60-е",
      "location": "на заднем плане"
    }
  ]
}.
```


Начальный запрос был простым и не раскрывал никакой информации о фотографии. Ответ ChatGPT показал, что модель распознала изображение протеста и описала обстановку, протестующих и тексты на их плакатах. Чтобы получить более точные данные, был дан запрос на извлечение детальной информации из фотографии и представление ее в формате JSON для облегчения последующей обработки, например, на языке Python. Как видно, нейросети удалось распознать текст с плакатов и даже получить перевод, а также получить структурированную информацию, часто необходимую для дальнейших статистических и

вычислительных исследований. Результаты были не идеальными: распознаны не все тексты, ввиду того, что некоторые надписи частично перекрываются. По-видимому, уточнение запросов или доработка на основе примеров могли бы улучшить точность.

Этот пример подчеркивает, как мультимодальный ИИ может кодировать и описывать изображения без необходимости специализированного обучения. Такой подход можно масштабировать для анализа больших наборов изображений, расширяя возможности изучения коллективных действий с использованием визуальных данных. В отличие от существующих применений компьютерного зрения в социологии, которые часто ограничиваются задачами классификации изображений или обнаружения объектов, мультимодальный ИИ может извлекать более полную информацию из визуальных данных. Так же, как и в случае с анализом текста, взаимодействие с моделями через запросы упрощает эксперименты и создает новые возможности для использования визуальных данных в качественных исследованиях. Например, этнографы или архивные исследователи могли бы применять модели компьютерного зрения для создания описаний фотографий или видео, что упростило бы работу с большими коллекциями изображений и видео.

Генерация синтетических изображений. LLM могут генерировать тексты, а самые передовые мультимодальные модели способны генерировать изображения на основе текстовых запросов. Эти синтетические данные представляют собой нечто новое, что выходит за рамки привычных категорий. Как и специально собранные данные, запросы могут адаптировать сгенерированные материалы к нашим требованиям, но результирующие синтетические изображения тоже могут представлять интерес, поскольку они синтезируются на базе огромных объемов данных, использованных для обучения моделей.

До недавнего времени тексты и изображения, сгенерированные ИИ, рассматривались как курьез и могли вызывать улыбку или недоумение, однако технологический прогресс привел к значительному улучшению качества результатов. Тексты, созданные LLM, сложно отличить от написанных людьми [6], а изображения лиц, сгенерированные ИИ, могут быть неотличимы от настоящих [7].

Способность ИИ создавать реалистичные тексты и изображения открывает новые возможности для экспериментальных исследований. В частности, можно использовать инструменты наподобие Fooocus или DALL·E для создания реалистичных синтетических изображений, например для иллюстрации статей или статистических диаграмм. Такие изображения могут улучшить внешнюю валидность экспериментов, передавая социальную информацию более естественным образом, чем традиционные таблицы атрибутов, применяемые в конъюнктурных исследованиях.

В процессе выполнение данной работы, с целью иллюстрации возможностей GAI, были синтезированы фотореалистичные изображения с использованием инструментария Fooocus, серверная часть которого была запущена в Google Colab. Использованы только модели, подключенные по умолчанию. Для получения изображения (рис. 2) был сформулирован запрос: «Спортсмен на зимних Олимпийских играх в Сочи, 2014 г.». Дополнительно передавались параметры о поле, возрасте и виде спорта спортсмена для создания изображений четырех вымышленных персонажей.

ДИСКУРС. 2025. Т. 11, № 1. С. 52–70
DISCOURSE. 2025, vol. 11, no. 1, pp. 52–70

Рис. 2. Результат генерации изображений
Fig. 2. Result of image generation

Полученные изображения весьма правдоподобны (хотя можно заметить некоторые несостыковки в геометрии и физике) и демонстрируют различия по половому признаку и возрасту. Даже относительно простые запросы позволяют генерировать изображения, передающие массу информации, которая способна оказать эмоциональное влияние на зрителя. Интересно отметить, что изображения включают множество деталей, таких как одежда, выражения лиц и окружение, которые не были заданы явно, но были добавлены нейросетью; при этом более детализированные запросы могут явно управлять подобными элементами.

При использовании моделей для анализа или генерации изображений важно учитывать предвзятость в данных обучения, которая может приводить к тому, что модель лучше отображает одни характеристики по сравнению с другими. Как и LLM, мультимодальные модели после предварительного обучения приобретают определенное «видение» и описание мира. Модели компьютерного зрения часто допускают ошибки при определении демографических данных на основе изображений, что приводит к систематическим предвзятостям в отношении маргинализированных групп [8]. Эти предвзятости также встречаются в GAI: в частности, аудит модели Stable Diffusion выявил склонность модели к воспроизведению стереотипов, включая представление женских лиц при запросе изображения «эмоционального человека», белых людей при запросе «привлекательного человека» и чернокожих мужчин при запросе «бандит» [9]. Учитывая такие нюансы, можно рекомендовать проводить дополнительные исследования для оценки качества этих представлений и скрытых предвзятостей, чтобы подтвердить возможность использования синтетического контента в экспериментальных условиях. Кроме того, такие атрибуты, как знание языка или продолжительность проживания [10], не всегда легко визуально представить, что ограничивает применимость сгенерированных изображений в зависимости от исследовательского вопроса.

Помимо создания изображений людей, мультимодальные модели могут генерировать сценарии, представляющие контекстуальные и окружающие факторы, а также большие коллективы. Пример, опирающийся на недавние эксперименты по оценке влияния характеристик протестов на их поддержку [11], иллюстрирует возможность использования таких моделей для создания высококачественных изображений различных демонстраций, протестов и тактик. В данном случае были сгенерированы четыре изображения протестов с вариациями по численности, сохраняя неизменным место проведения и тематику. На рис. 3 можно

видеть синтезированные изображения по запросу «Небольшой митинг Движения за права женщин в Иране», а на рис. 4 – синтезированные изображения по запросу «Массовая демонстрация Движения за права женщин в Иране».

Действительно, все упомянутые нами изображения демонстрируют группы протестующих на улицах Ирана в дневное время. Видны различия в численности протестов и расположении «виртуальной камеры». Протестующие на небольшом митинге держат плакаты с надписями, которые весьма напоминают письменность фарси (иранский язык). Многие участницы носят хиджаб (у них покрыта голова). Весьма правдоподобно показаны флаги Ирана. Сложно оценить национальную принадлежность синтезированных персонажей, но они вполне напоминают среднестатистических жителей Ирана. Есть и очевидные недостатки: модели генерации изображений пока не справляются с отображением текста, и некоторые плакаты, очевидно, содержат бессмысленные надписи. Поэтому исследователям важно внимательно проверять сгенерированный контент, хотя эту проблему, вероятно, решат по мере развития технологий.

Рис. 3. Результат генерации изображений по запросу «Небольшой митинг Движения за права женщин в Иране»
Fig. 3. Result of image generation on the query «Small meeting of the Women's Rights Movement in Iran»

Рис. 4. Результат генерации изображений по запросу «Массовая демонстрация Движения за права женщин в Иране»
Fig. 4. Results of image generation based on the query «Mass demonstration of the Women's Rights Movement in Iran»

ДИСКУРС. 2025. Т. 11, № 1. С. 52–70
DISCOURSE. 2025, vol. 11, no. 1, pp. 52–70

Хотя можно решать проблемы синтетических медиа, подбирая фотографии реальных протестов, представленный здесь подход позволяет создавать контрфактические сценарии, которые невозможно найти среди существующих изображений. Чтобы продемонстрировать этот потенциал, было сгенерировано изображение демонстрации Движения за права женщин в Иране с горящей машиной, символизирующей разрушение собственности и насилие, которые, как известно, снижают поддержку протестов [12], хотя такие действия не характерны для женских протестов. В данном примере используется модель Fooocus; окончательная версия запроса указана в подписи к изображению (рис. 5). Очевидно, данный пример наглядно демонстрирует, насколько легко в настоящее время ввести читателя или зрителя в заблуждение, предлагая реалистичную фотографию с искаженным смыслом.

Рис. 5. Два варианта синтезированных изображений по запросу
«Демонстрация Движения за права женщин в Иране. Горящий автомобиль на переднем плане»
Fig. 5. Two variants of synthesized images based on the query
«Demonstration of the Women's Rights Movement in Iran. Burning car in the foreground»

Мультимодальные модели открывают новые возможности для использования визуальных данных в социологии. Эти модели не только способны выполнять классификационные задачи, но и могут создавать подробные описания изображений и генерировать реалистичные изображения, отражающие социологически значимую информацию. Текстовые, визуальные и мультимедийные возможности GAI предоставляют социологам новые способы взаимодействия с текстами, изображениями и другими медиа. Это не означает, что все социологические исследования должны использовать эти инструменты, или что GAI заменит другие вычислительные методы и роль аналитика, но технологические новшества имеют уникальные возможности, дополняющие существующие методологии. GAI обещает сделать существующие методы более гибкими и эффективными, а также открыть новые формы анализа, компьютерной интерпретации и экспериментов с синтетическим контентом.

Возможные проблемы использования GAI. Относительно нейросетей и искусственного интеллекта достаточно часто высказываются опасения общего характера, в основном по поводу потенциального злоупотребления генеративным ИИ, такого как киберпреступность, использование заведомо ложных новостей или генерация видеоклипов с подменой лиц персонажей (deepfake) для обмана, мошенничества или манипулирования людьми, а также массовая замена рабочих мест людьми [1, 3]. Но если сфокусировать внимание на применении

ДИСКУРС. 2025. Т. 11, № 1. С. 52–70
DISCOURSE. 2025, vol. 11, no. 1, pp. 52–70

GAI для решения прикладных задач в сферах социологии, то набор проблем может быть локализован, что рассматривается далее.

Хотя GAI открывает методологические возможности, эти технологии сталкиваются с проблемами интерпретируемости, прозрачности, воспроизводимости, надежности и этики. Масштаб моделей, обеспечивая высокую производительность, также является источником слабых мест. Интерпретировать сложные нейронные сети трудно, особенно модели с миллиардами параметров. Область «механической интерпретируемости» пытается объяснить работу таких моделей, но она находится на ранней стадии развития.

Одной из самых серьезных проблем является то, что современные модели искусственного интеллекта (AI) для генерации изображений лучше справляются с тематикой, на которой они больше всего обучались. Рассмотрим, чем это обусловлено. Во-первых, эти модели обучаются на огромных наборах данных, которые включают миллионы изображений и соответствующих текстовых описаний. Этот процесс позволяет им выявлять сложные взаимосвязи между словами и визуальными элементами. Например, когда модель обучается на изображениях и текстовых описаниях, она начинает понимать, что такое «голубое небо» или «зеленая трава» и как эти элементы обычно появляются вместе в реальных сценах. Во-вторых, современные подходы к обучению AI, такие как контрастивное обучение и использование синтетических изображений, позволяют моделям глубже понимать высокоуровневые концепции, а не просто запоминать отдельные пиксели. Например, система StableRep, разработанная исследователями из MIT, использует синтетические изображения, сгенерированные текстово-изображающими моделями, для обучения моделей, что позволяет им лучше понимать концепции и вариации, а не просто данные. В-третьих, модели, обученные на конкретной тематике, могут лучше заполнять пробелы в информации, используя общие знания о мире. Например, если модель обучена на изображениях животных, она может генерировать изображение «голубя на улице» с учетом того, что голуби обычно бывают серыми, а улицы часто имеют определенные архитектурные особенности.

Также возникает проблема с прозрачностью данных обучения, так как многие модели управляются корпорациями, которые не раскрывают информацию о своих данных. Это затрудняет оценку влияния обучающих данных на результаты. Например, синтетические изображения протестов во время 2021–2022 гг. могут отображать участников в масках из-за данных, собранных во время пандемии, что может внести нежелательные искажения при использовании этих изображений в экспериментах. Поэтому важно учитывать, как данные обучения могут повлиять на результаты, и тщательно их оценивать. В долгосрочной перспективе необходимо исследовать данные, использованные для обучения, чтобы лучше понимать взаимосвязь между входами и выходами моделей.

Связанные с GAI проблемы включают воспроизводимость. Хотя воспроизводимость вычислительных подходов часто считается сильной стороной по сравнению с гуманитарными [13], она не гарантирована из-за стохастической структуры моделей. Одинаковые запросы могут давать разные результаты, и подобные проблемы воспроизводимости возникают даже в более простых моделях машинного обучения. Кроме того, GAI-технологии, контролируемые корпорациями, могут изменяться без уведомления, что делает невозможным воспроизведение ранее полученных результатов.

ДИСКУРС. 2025. Т. 11, № 1. С. 52–70
DISCOURSE. 2025, vol. 11, no. 1, pp. 52–70

Способность GAI генерировать правдоподобный вывод может также отрицательно сказаться на некоторых исследованиях. Модели, как правило, не основаны на формальной логике и могут выдавать недостоверную информацию, что вызывает опасения по поводу надежности и стабильности. Разработчики GAI работают над созданием более надежных моделей, но исследователи должны внимательно проверять выводы и избегать использования этих моделей для задач, требующих фактической точности.

Применение GAI в социальных науках поднимает новые этические и конфиденциальные проблемы. Например, работа с чувствительными данными, такими как интервью, может нарушать правила IRB [14], так как ввод данных в коммерческие модели означает, что данные передаются третьей стороне. Кроме того, GAI может выдавать неожиданные ответы в экспериментальных условиях, что может привести к расистским или сексистским стереотипам, негативно влияющим на участников. Университеты и IRB должны разработать протоколы для безопасного использования GAI в исследованиях с участием людей. Исследователи должны осознавать риски, связанные с этими технологиями, и тщательно оценивать, как использование GAI может повлиять на субъектов исследования.

Последствия предвзятости и усилия по смягчению ее последствий. Проблема предвзятости в моделях GAI требует более тщательного рассмотрения. Исследования показывают, что традиционные методы машинного обучения могут воспроизводить предвзятости и стереотипы. Например, модели, обученные распознавать пол по изображениям, менее точны для темнокожих женщин, так как обучающие примеры преимущественно содержали белых мужчин [8]. GAI также страдает от предвзятостей, что связано с использованием большого объема неподтвержденных данных [15]. Это может негативно сказаться на социальных науках; например, при запросах, связанных с мусульманами, нейросеть ChatGPT версии 3 часто генерировала ответы, связанные с терроризмом [16].

Компании по разработке ИИ пытаются минимизировать риск предвзятости, идентифицируя проблемный контент и используя дополнительные инструкции для избежания таких выводов. Например, компания OpenAI изменила модель DALL·E, добавляя демографическую информацию к запросам, чтобы увеличить разнообразие выходных данных. Однако такие меры могут привести к абсурдным и исторически неточным изображениям.

Несмотря на важность усилий по уменьшению предвзятости, они могут подорвать использование GAI в исследованиях. С социальной научной точки зрения возможность генерировать «предвзятые» тексты или изображения часто делает их аналитически полезными [17]. Например, ограничения по поводу контента «для взрослых» или незаконной деятельности могут помешать анализу интервью, связанных с сексуальным здоровьем и наркотиками. В частности, нейросеть DALL·E 3 отказалась показывать иммигрантов с уголовными записями или протесты с насилием. Однако у нейросети Fooocus не оказалось ограничений на данную тематику. Эта ситуация подчеркивает очевидное напряжение между предвзятостью и ее смягчением. Если предвзятость является проблемой, исследователи могут предпочесть более контролируемые модели, такие как Claude от Anthropic, которые стремятся быть «полезными, честными и безвредными».

Независимо от того, как используются GAI в социологических исследованиях, важно тщательно проверять, как предвзятости, встроенные в эти технологии, могут повлиять на

ДИСКУРС. 2025. Т. 11, № 1. С. 52–70
DISCOURSE. 2025, vol. 11, no. 1, pp. 52–70

выполнение конкретных задач. Ведущие ученые в сфере информационных технологий предложили включать документацию к моделям и обучающим наборам данных [15], указывая на известные или скрытые предвзятости. Однако такая документация должна быть лишь начальной точкой в изучении взаимодействия между моделями, обученными на огромных наборах данных, и социологическими данными. Исследователи должны сообщать о любых подобных проблемах в опубликованных работах, детализируя, где они были замечены, аналогично тому, как качественные исследователи рефлексируют по поводу своей позиционности и предвзятостей. Если GAI используются для кодирования текстов или визуальных данных, мы должны оценивать их предвзятости так же, как оцениваем работу исследовательского ассистента, тщательно проверяя и сравнивая их выводы с другими примерами. Эта проверка критически важна для обеспечения того, чтобы исследования с использованием этих технологий были не только методически инновационными, но и этически обоснованными и научно достоверными.

Сравнение моделей с открытым исходным кодом и коммерческих. Кроме того, интересно рассмотреть сравнение двух типов моделей GAI: с открытым исходным кодом и коммерческими, а также их последствия для социологических исследований. В настоящее время самые мощные модели разрабатываются коммерческими компаниями, такими как OpenAI и Google. Ключевые особенности этих моделей, включая внутренние параметры, архитектуру и обучающие данные, являются собственностью компаний. В отличие от них, такие модели, как Fooocus или BLOOM, являются открытыми, и любой может ознакомиться с исходным кодом и входными данными.

С технической точки зрения закрытые модели, управляемые несколькими корпорациями, обычно мощнее и проще в использовании, чем открытые альтернативы. Например, взаимодействовать с ChatGPT можно через браузер, а программные интерфейсы (API) позволяют пользователям оплачивать выполнение больших объемов запросов. Однако такие модели могут быть чрезмерно дорогими для приложений, требующих больших объемов выводов, например для классификации текста в миллионах документов. Открытые модели начинают конкурировать с коммерческими решениями, особенно с учетом того, что такие компании, как Meta и Google, открывают некоторые свои модели для общественности. Но эти модели часто требуют доступа к дорогостоящей технической инфраструктуре, что делает их недоступными для большинства социологов.

Закрытые коммерческие модели сейчас являются более жизнеспособным решением для многих социальных ученых, но существуют несколько причин, по которым следует предпочесть открытые модели. Такие модели не обязательно легче интерпретировать, но исходный код и веса модели доступны, что дает возможность исследователям пытаться интерпретировать их. Кроме того, прозрачность обучающих данных позволяет лучше оценивать потенциальные предвзятости и стереотипные выводы. Возможность запускать эти модели на оборудовании, контролируемом исследователем, делает открытые модели более совместимыми с целью воспроизводимости, так как исследователи могут хранить и делиться точными конфигурациями модели. Это устраняет риски конфиденциальности, поскольку данные не передаются третьим лицам, что делает эти модели подходящими для задач с чувствительными данными. В отношении предвзятости открытые модели, которые

можно свободно модифицировать, предпочтительнее, если исследование затруднено усилиями по устранению предвзятости. В целом, мы призываем исследователей тщательно рассмотреть эти вопросы и рекомендуем экспериментировать с несколькими различными моделями, чтобы проверить, как разные реализации влияют на выполнение задач.

В долгосрочной перспективе можно предположить, что академические исследователи будут лучше обеспечены использованием GAI, специально разработанных для научных целей, и не придется полагаться на крупные корпорации. Такие модели могут устранить ограничения существующих технологий, позволяя проверенным исследователям оценивать взаимосвязь между различными обучающими данными и результатами, а также генерировать более чувствительные выводы с меньшими ограничениями.

Создание GAI для социальных наук потребует значительных ресурсов для сбора и курирования огромных объемов обучающих данных, а также для управления и обслуживания дорогого оборудования, необходимого для их обучения и развертывания. Эти инициативы, вероятно, будут выходить за рамки возможностей одной группы исследователей и потребуют сотрудничества между несколькими учреждениями. Критически важно, чтобы социологи принимали участие в этих усилиях как для анализа социальных последствий этих технологий, так и для использования их потенциала на пользу социологии.

Заключение. В публичной сфере GAI вызывают крайне неоднозначные оценки. Данная статья предлагает более взвешенный подход, подчеркивая методологические возможности, которые открывают технологические инновации, при этом признавая их ограничения. LLM могут превосходить существующие методы классификации текста, делая их более эффективными и доступными, а также легко адаптироваться к различным областям и задачам. Эти же модели могут использоваться для более разговорного анализа содержания, позволяя качественным исследователям взаимодействовать с данными новыми способами.

Существуют и другие возможности, например, мультимодальные модели могут генерировать высококачественные изображения людей и протестов, включая контрфактические сценарии, что может быть использовано в экспериментальных исследованиях. Способность применять сложные вычислительные методы с минимальной подготовкой и техническими знаниями, а также адаптировать их под конкретные нужды, безусловно, будет способствовать распространению машинного обучения в социологии.

Приведенные в статье примеры далеко не исчерпывающие. Исследователи разрабатывают креативные применения этих технологий в разных областях социальных наук. Например, LLM могут заполнять пропуски в данных опросов, помогать проводить онлайн полуструктурированные интервью и строить более реалистичные симуляции. Анализ мультимодальных данных делает GAI особенно перспективными для изучения огромных объемов онлайн- и цифровых текстов, изображений, аудио и видео.

Изучение репрезентаций, встроенных в эти модели, также позволит заглянуть изнутри в культуру и когнитивные процессы. Эти трансформационные возможности выходят за рамки методологии, так как GAI могут служить «виртуальными научными помощниками», помогать обучать навыкам кодирования и даже генерировать исследовательские вопросы. Исследование AI станет важной областью социологического анализа по мере распространения этих технологий в различных сферах социальной жизни. Хотя эти технологические

ДИСКУРС. 2025. Т. 11, № 1. С. 52–70
DISCOURSE. 2025, vol. 11, no. 1, pp. 52–70

изменения принесут множество преимуществ, социологи могут возглавить усилия по выявлению того, как AI воспроизводит существующие неравенства и создает новые формы социальной дивизии.

Использование GAI в социологических исследованиях также порождает значительные проблемы, поскольку сильные стороны этих моделей могут восприниматься как слабости. Большие и сложные модели трудны для интерпретации, не всегда ясно, как их выводы зависят от текстов и изображений, использованных в предобучении, и они могут давать ненадежную или вводящую в заблуждение информацию. Предобучение на больших объемах текста приводит к усвоению и воспроизведению предвзятостей, что может затруднить их использование в научной среде. При этом усилия по смягчению предвзятости усложняются, так как предвзятость часто является ключевым объектом интереса для исследователей. Ученые, использующие эти инструменты, должны быть особенно внимательны к таким ограничениям и тщательно оценивать и документировать, как они могут повлиять на применение GAI к конкретным исследовательским вопросам. Особую осторожность следует проявлять при работе с чувствительными данными или взаимодействии с участниками исследований.

Некоторые проблемы существующих реализаций GAI усугубляются тем, что самые мощные модели контролируются корпорациями, предоставляющими мало информации о внутренней структуре своих систем или данных, на которых они обучаются. Открытые альтернативы могут частично решить эти проблемы, но они труднее в использовании, так как требуют более сложной вычислительной подготовки и ресурсов.

В заключение важно подчеркнуть, что GAI не заменит исследователя или существующие методологические инструменты. Как показывают исследования использования ИИ на рабочем месте и в других областях, мы должны меньше беспокоиться о том, что эти технологии заменят нас, и больше сосредоточиться на их потенциале для расширения наших возможностей. Очевидно, что GAI может улучшить существующие методологии, такие как классификация и аннотация данных, а также может позволить новые формы исследования, такие как разговорный анализ содержания и синтез изображений. При осторожном использовании эти технологии будут способствовать методологическим инновациям в социальных науках и формированию новых подполей и специализаций. Можно рекомендовать социологам воспользоваться методологическими возможностями, которые открывает GAI.

СПИСОК ЛИТЕРАТУРЫ

1. Generative artificial intelligence // Wikipedia. URL: https://en.wikipedia.org/wiki/Generative_artificial_intelligence (дата обращения: 20.08.2024).

2. Глухих В. А., Елисеев С. М., Кирсанова Н. П. Искусственный интеллект как проблема современной социологии // Дискурс. 2022. Т. 8, № 1. С. 82–93. DOI: https://doi.org/10.32603/2412-8562-2022-8-1-82-93.

3. Ильичев В. Ю., Драч В. Е., Чукаев К. Е. Морально-нравственные проблемы всеобщего применения нейронных сетей // Рефлексия. 2023. № 5. С. 8–13.

4. Цифровая экономика и системная цифровая трансформация / А. С. Копырин, Е. В. Видищева, В. В. Коваленко и др. / под общ. ред. А. С. Копырина. Сочи: РИЦ ФГБОУ ВО «СГУ», 2023.

5. Deepti Hajela. The American paradox of protest: Celebrated and condemned, welcomed and muzzled // The Associated Press. 05.05.2024. URL: https://apnews.com/article/american-protest-paradox-israel-hamas-war22b1325188e0808db7389c8c3f04c331 (дата обращения: 20.08.2024).

6. All That's 'Human' Is Not Gold: Evaluating Human Evaluation of Generated Text / E. Clark, T. August, S. Serrano et al. // Proc. of the 59th Annual Meeting of the Association for Computational Linguistics and the 11th Int. Joint Conf. on Natural Language Proc. Vol. 1. Stroudsburg, PA: Association for Computational Linguistics, 2021. P. 7282–7296. DOI: 10.48550/arXiv.2107.00061.

7. Wang Y., Wang H. A face template: improving the face generation quality of multi-stage generative adversarial networks using coarse-grained facial priors // Multimedia Tools and Applications. 2024. Vol. 83, no. 7. P. 21677–21693. DOI: https://doi.org/10.1007/s11042-023-16183-2.

8. Buolamwini J. Unmasking AI: My Mission to Protect What Is Human in a World of Machines. NY: Random House, 2023.

9. Easily Accessible Text-to-Image Generation Amplifies Demographic Stereotypes at Large Scale / F. Bianchi, P. Kalluri, E. Durmus et al. // FAccT '23 Proc. of the 2023 ACM Conf. on Fairness, Accountability, and Transparency. NY: Association for Computing Machinery, 2023. P. 1493–1504. DOI: https://doi.org/10.1145/3593013.3594095.

10. Flores R. D., Schachter A. Who Are the "Illegals"? The Social Construction of Illegality in the United States // American Sociological Review. 2018. Vol. 83, iss. 5. P. 839–868. DOI: 10.1177/0003122418794635.

11. How Tilly's WUNC Works: Bystander Evaluations of Social Movement Signals Lead to Mobilization / E. R. Bailey, D. Wang, S. A. Soule, R. Hayagreeva // American J. of Sociology. 2023. Vol. 128, no. 4. P. 1206–1262. DOI: https://doi.org/10.1086/723489.

12. Feinberg M., Willer R., Kovacheff Ch. The Activist's Dilemma: Extreme Protest Actions Reduce Popular Support for Social Movements // J. of Personality and Social Psychology. 2020. Vol. 119, iss. 5. P. 1086–1111. DOI: 10.1037/pspi0000230.

13. Nelson L. K. Computational Grounded Theory: A Methodological Framework // Sociological Methods & Research. 2020. Vol. 49, iss. 1. P. 3–42. DOI: https://doi.org/10.1177/0049124117729703.

14. Institutionalreviewboard // Wikipedia. URL: https://en.wikipedia.org/wiki/Institutionalreviewboard (дата обращения: 20.08.2024).

15. On the Dangers of Stochastic Parrots: Can Language Models Be Too Big? / E. M. Bender, T. Gebru, A. McMillan-Major, Sh. Shmargaret // FAccT '21: Proc. of the 2021 ACM Conf. on Fairness, Accountability, and Transparency. NY: Association for Computing Machinery, 2021. P. 610–623. DOI: https://doi.org/10.1145/3442188.3445922.

16. Abubakar A., Maheen F., Zou J. Persistent Anti-Muslim Bias in Large Language Models // Proc. of the 2021 AAAI/ACM Conf. on AI, Ethics, and Society, Virtual Event. NY: Association for Computing Machinery, 2021. P. 298–306. DOI: 10.48550/arXiv.2101.05783.

17. Out of One, Many: Using Language Models to Simulate Human Samples / L. P. Argyle, E. C. Busby, N. Fulda et al. // Political Analysis. 2023. Vol. 31, iss. 3. P. 337–351. DOI: https://doi.org/10.1017/pan.2023.2.

Информация об авторах.

Драч Владимир Евгеньевич **–** кандидат технических наук (2005), доцент (2006), доцент кафедры информационных технологий и математики Сочинского государственного университета, ул. Пластунская, д. 94, г. Сочи, 354000, Россия. Автор 106 научных публикаций. Сфера научных интересов: радиоэлектроника, информационные технологии.

Торкунова Юлия Владимировна **–** доктор педагогических наук (2015), профессор кафедры информационных технологий и интеллектуальных систем Казанского государственного энергетического университета, ул. Красносельская, д. 51, г. Казань, 420066, Россия; профессор кафедры информационных технологий и математики Сочинского государственного университета, ул. Пластунская, д. 94, г. Сочи, 354000, Россия. Автор 171 научной публикации. Сфера научных интересов: информационные технологии в образовании и экономике.

О конфликте интересов, связанном с данной публикацией, не сообщалось.
Поступила 29.09.2024; принята после рецензирования 17.10.2024; опубликована онлайн 20.02.2025.

ДИСКУРС. 2025. Т. 11, № 1. С. 52–70
DISCOURSE. 2025, vol. 11, no. 1, pp. 52–70

REFERENCES

1. "Generative artificial intelligence", *Wikipedia*, available at: https://en.wikipedia.org/wiki/Generative_artificial_intelligence (accessed 20.08.2024).

2. Glukhikh, V.A., Eliseev, S.M. and Kirsanova, N.P. (2022), "Artificial Intelligence as a Problem of Modern Sociology", *DISCOURSE*, vol. 8, no. 1, pp. 82–93. DOI: 10.32603/2412-8562-2022-8-1-82-93.

3. Ilyichev, V.Yu., Drach, V.E. and Chukaev, K.E. (2023), "Moral and Ethical Problems of Universal Use of Neural Networks", *Refleksiya* [Reflection], no. 5, pp. 8–13.

4. Kopyrin, A.S., Vidishcheva, E.V., Kovalenko, V.V. et al. (2023), *Tsifrovaya ekonomika i sistemnaya tsifrovaya transformatsiya* [Digital Economy and Systemic Digital Transformation], in Kopyrin, A.S. (ed.), RIC FGBOU VO 'SSU", Sochi, RUS.

5. Deepti Hajela (2024), "The American paradox of protest: Celebrated and condemned, welcomed and muzzled", *The Associated Press*, 05.05.2024, available at: https://apnews.com/article/american-protest-paradox-israel-hamas-war22b1325188e0808db7389c8c3f04c331 (accessed 20.08.2024).

6. Clark, E., August, T., Serrano, S. et al. (2021), "All That's 'Human' Is Not Gold: Evaluating Human Evaluation of Generated Text", *Proc. of the 59th Annual Meeting of the Association for Computational Linguistics and the 11th Int. Joint Conf. on Natural Language Proc*., vol. 1, Stroudsburg, PA: Association for Computational Linguistics, pp. 7282–7296. DOI: 10.48550/arXiv.2107.00061.

7. Wang, Y. and Wang, H. (2024), "A face template: improving the face generation quality of multi-stage generative adversarial networks using coarse-grained facial priors", *Multimedia Tools and Applications*, vol. 83, no. 7, pp. 21677–21693. DOI: https://doi.org/10.1007/s11042-023-16183-2.

8. Buolamwini, J. (2023), *Unmasking AI: My Mission to Protect What Is Human in a World of Machines*, Random House, NY, USA.

9. Bianchi, F., Kalluri, P., Durmus, E. et al. (2023), "Easily Accessible Text-to-Image Generation Amplifies Demographic Stereotypes at Large Scale", *FAccT '23 Proc. of the 2023 ACM Conf. on Fairness, Accountability, and Transparency*, Association for Computing Machinery, NY, USA, pp. 1493–1504. DOI: https://doi.org/10.1145/3593013.3594095.

10. Flores, R.D. and Schachter, A. (2018), "Who Are the "Illegals"? The Social Construction of Illegality in the United States", *American Sociological Review*, vol. 83, iss. 5, pp. 839–868. DOI: 10.1177/0003122418794635.

11. Bailey, E.R., Wang, D., Soule, S.A. and Hayagreeva, R. (2023), "How Tilly's WUNC Works: Bystander Evaluations of Social Movement Signals Lead to Mobilization", *American J. of Sociology*, vol. 128, no. 4, pp. 1206–1262. DOI: https://doi.org/10.1086/723489.

12. Feinberg, M., Willer, R. and Kovacheff, Ch. (2020), "The Activist's Dilemma: Extreme Protest Actions Reduce Popular Support for Social Movements", *J. of Personality and Social Psychology*, vol. 119, iss. 5, pp. 1086–1111. DOI: 10.1037/pspi0000230.

13. Nelson, L.K. (2020), "Computational Grounded Theory: A Methodological Framework", *Sociological Methods & Research*, vol. 49, iss. 1, pp. 3–42. DOI: https://doi.org/10.1177/0049124117729703.

14. "Institutionalreviewboard", *Wikipedia*, available at: https://en.wikipedia.org/wiki/Institutional reviewboard (дата обращения: 20.08.2024).

15. Bender, E.M., Gebru, T., McMillan-Major, A. and Shmargaret, Sh. (2021), "On the Dangers of Stochastic Parrots: Can Language Models Be Too Big?", *FAccT '21: Proc. of the 2021 ACM Conf. on Fairness, Accountability, and Transparency*, Association for Computing Machinery, NY, USA, pp. 610–623. DOI: https://doi.org/10.1145/3442188.3445922.

16. Abubakar, A., Maheen, F. and Zou, J. (2021), "Persistent Anti-Muslim Bias in Large Language Models", *Proc. of the 2021 AAAI/ACM Conf. on AI, Ethics, and Society, Virtual Event*, Association for Computing Machinery, NY, USA, pp. 298–306. DOI: 10.48550/arXiv.2101.05783.

17. Argyle, L.P., Busby, E.C., Fulda, N. et al. (2023), "Out of One, Many: Using Language Models to Simulate Human Samples", *Political Analysis*, vol. 31, iss. 3, pp. 337–351. DOI: https://doi.org/10.1017/pan.2023.2.

ДИСКУРС. 2025. Т. 11, № 1. С. 52–70
DISCOURSE. 2025, vol. 11, no. 1, pp. 52–70

Information about the authors.

Vladimir E. Drach **–** Can. Sci. (Engineering, 2005), Docent (2006), Associate Professor at the Department of Information Technologies and Mathematics, Sochi State University, 94 Plastunskaya str., Sochi 354000, Russia. The author of 106 scientific publications. Area of expertise: radio electronics and information technologies.

Yulia V. Torkunova – Dr. Sci. (Pedagogic, 2015), Professor at the Department of Information Technologies and Intelligent Systems, Kazan State Power Engineering University, 51 Krasnoselskaya str., Kazan 420066, Russia; Professor at the Department of Information Technologies and Mathematics, Sochi State University, 94 Plastunskaya str., Sochi 354000, Russia. The author of 171 scientific publications. Area of expertise: information technology in education and economics.

No conflicts of interest related to this publication were reported.
Received 29.09.2024; adopted after review 17.10.2024; published online 20.02.2025.