

Виктория Сергеевна Машошина, к. филол. н.
Московский городской педагогический уни-
верситет, Москва, Россия
э-почта: MashoshinaVS@mgpu.ru

Victoria Sergeevna Mashoshina, Cand. of Sc.
(Philology)
Moscow City University, Moscow, Russia
e-mail: MashoshinaVS@mgpu.ru

ОТ ГУМБОЛЬДА ДО ГУГЛА: ПОТЕНЦИАЛ ГЕНЕРАТИВНОГО ИСКУССТВЕННОГО ИНТЕЛЛЕКТА В ДИСКУРСИВНЫХ ИССЛЕДОВАНИЯХ

В работе обобщается опыт использования доступных инструментов искусственного интеллекта, которые могут быть задействованы при проведении научных дискурсивных исследований в области лингвистики в целях развития цифровой грамотности преподавателей и студентов, эффективного выполнения рутинных исследовательских задач, а также оптимизации способов обработки естественного языка.

Ключевые слова: ИИ ассистенты исследователя; лингвистика.

FROM HUMBOLDT TO GOOGLE: THE POTENTIAL OF GENERATIVE AI IN DISCURSIVE RESEARCH

The paper summarizes the experience of using available artificial intelligence tools that can be used in discourse-analytical studies in the field of linguistics in order to develop educators' and students' digital literacy, effectively perform routine research tasks, and leverage natural language processing methods.

Key words: AI research tools; linguistics.

«Язык есть не продукт деятельности (*ergon*),
а деятельность (*energeia*)».
В. Гумбольдт

С момента представления на Берлинской конференции 1820 г. одной из ранних работ В. фон Гумбольдта «О сравнительном изучении языков применительно к различным эпохам их развития» прошло немногим более двухсот лет. Однако остаётся надеяться, что «языкотворческая сила в человечестве», о которой много говорил этот выдающийся учёный, по-прежнему будет стремиться к совершенству, невзирая на технический прогресс и стремительное развитие технологий генеративного искусственного интеллекта (ГИИ) в XXI в.

Современные темпы и тренды развития ГИИ поистине впечатляют. В 2024 г. Нобелевскую премию по физике присудили за исследования в области искусственных нейронных сетей. Лауреатами стали американский ученый, профессор Принстонского университета Дж. Хопфилд (*John J. Hopfield*), который в 1982 г. изобрел ассоциативную нейронную сеть, и британец Дж. Хинтон (*Geoffrey E. Hinton*), профессор Университета Торонто (Канада), известный как крёстный отец глубокого обучения (*Deep Learning*).

По данным немецкой компании *Statista*, специализирующейся на маркетинговых и статистических исследованиях, в период с 2024 по 2030 гг. совокупный среднегодовой темп роста (CAGR 2024-2030) рынка ИИ составит 28,46% и в результате к 2030 г. прогнозируемый объем мирового рынка технологий ИИ может достигнуть 826 70 млрд долларов США (URL: <https://www.statista.com/outlook/tmo/artificial-intelligence/worldwide>).

Сегодня созданы целые библиотеки инструментов ГИИ для различных целей, которые пополняются на буквально ежедневной основе (Free AI Tools Library, TopAI.tools, AI Tools Directory и другие).

Как известно, «идеи о расширении границ лингвистики разрабатываются исследователями в течение достаточно длительного периода времени» [1, с. 6]. При этом именно сегодня мы наблюдаем устойчивое партнерство лингвистики и искусственного интеллекта в таких областях, как анализ обширных массивов лингвистических данных, обработка естественного языка, когнитивное моделирование и др.

Репертуар активно используемых по всему миру ИИ ассистентов исследователя, направленных на фасилитацию автоматического анализа и обработки научных текстов, также достаточно широк, и далеко не исчерпывается представленными ниже ресурсами: Elicit (URL: <https://elicit.com/>); QillBot (URL: <https://quillbot.com/>); Scholarcy (URL: <https://www.scholarcy.com/>); Scispace (URL: <https://typeset.io/>); SciSummary (URL: <https://scisummary.com/>); Scopus AI (URL: <https://www.elsevier.com>) и др.

Ведущие образовательные и научно-исследовательские организации активно разрабатывают воросы внедрения инструментов ГИИ в систему высшего лингвистического образования, что, безусловно, соотносится с необходимостью:

- проведения исследований по оценке эффективности применения ГИИ в области лингвистики;
- разработки методов обучения студентов и преподавателей эффективной работе с ГИИ-системами;
- создания инструментов для контроля и верификации результатов, полученных с помощью ГИИ.

Сегодня вопросы интеграции систем ГИИ в научно-исследовательскую деятельность студентов активно обсуждаются в ведущих университетах мира. Внедряемые в образовательный процесс средства ГИИ нацелены на решение самых разных задач от индивидуализации обучения до решения рутинных исследовательских задач, в том числе при написании выпускных квалификационных работ студентов [3, с. 202].

При этом главный вызов как для учёных, так и для начинающих исследователей представляет решение прикладных задач с применением инструментов искусственного интеллекта. Важно в данной связи способствовать синергии лингвистики и ГИИ, центрировать инструментарий ГИИ вокруг решения практических задач, в том числе направленных на верификацию гипотез; полимодальную обработку языка, изучение неочевидных языковых фактов и взаимосвязей и пр.

К числу эффективных стратегий применения ГИИ можно отнести различные компоненты исследовательской деятельности, в том числе в системе переводческого дискурса: разработка информационных систем энтропийного анализа машинных переводов произведений современной художественной литературы с использованием предобученных больших языковых моделей (GPT (Generative Pre-trained Transformer), BERT (Bidirectional Encoder Representations from Transformers), а также систем DeepL, Microsoft Translator; создание собственных корпусов текстов (AntConc, TextSTAT, NLTK); визуализация данных и работа с графами (Voyant Tools, Gephi); аннотирование звуковых и видеофайлов (ELAN); автоматическое аннотирование и реферирование научных текстов, представляющее собой «ключевую составляющую дигитализации» [2, с. 7] (SciSummary, Scopus AI, Scholarcy и др.).

Несмотря на множество преимуществ, которые дает использование генеративного ИИ при проведении дискурсивных исследований различного рода, следует учитывать всевозможные риски, с которыми ассоциируются продукты дигитализации. Широко распространены опасения по поводу возможной дезинформации, вопросов академической этики и достоверности контента, создаваемого ГИИ, потенциальных угроз для процесса обучения и развития критического мышления его участников. При этом мотивация к речемыслительной деятельности, язык как «обязательная предпосылка мышления» (В. Гумбольдт) и способность порождать новые смыслы по-прежнему остаются прерогативами человека, а не машины.

ЛИТЕРАТУРА

1. Основы концептного анализа художественного произведения: Учебное пособие для студентов магистратуры, обучающихся по направлению подготовки «Лингвистика» / О. В. Афанасьева, К. М. Баранова, В. С. Машошина, О. Г. Чупрына. Москва: Общество с ограниченной ответственностью «Диона». 2019. 130 с.
2. Петрова И. М., Иванов А. М., Никитина В. В. Современные цифровые технологии в лингвистических исследованиях : учеб. пособие для обучающихся по направлению «Лингвистика»; Московский городской педагогический университет, Институт иностранных языков. М. : Общество с ограниченной ответственностью «Языки Народов Мира». 2022. 259 с.
3. Тивьяева И. В. Регламентирование использования средств генеративного искусственного интеллекта в выпускной квалификационной работе / И. В. Тивьяева, С. В. Михайлова, А. А. Казанцева // Вестник МГПУ. Серия: Филология. Теория языка. Языковое образование. 2024. № 2(54). С. 202–218.