

УДК 81, 004.8

ЦИФРОВЫЕ ИНСТРУМЕНТЫ В ЛИНГВИСТИКЕ: ВОЗМОЖНОСТИ ПРИМЕНЕНИЯ ИСКУССТВЕННОГО ИНТЕЛЛЕКТА

Г.Р. Камалова

*Академия наук Республики Башкортостан
Уфа, Российская Федерация, guldar@mail.ru*

В статье рассмотрен потенциал искусственного интеллекта (ИИ), который открывает новые горизонты для исследований и практических приложений в области языкознания. Рассмотрены вспомогательные инструменты для анализа и работы с текстами, их возможности и применение в междисциплинарных исследованиях.

Ключевые слова. Искусственный интеллект (ИИ), лингвистика, цифровые инструменты, ChatGPT, NER, TTR, FLEx, ГигаЧат.

Стремительное развитие цифровых технологий оказывает значительное влияние на различные научные дисциплины, включая лингвистику. В данном контексте особое внимание привлекает потенциал искусственного интеллекта (ИИ), который открывает новые возможности для исследований и практических приложений в области языкознания. Искусственный интеллект активно трансформирует лингвистические исследования и образование, предлагая новые инструменты для анализа, обработки и изучения языка. Его применение охватывает широкий спектр задач – от лингвистического анализа текстов до персонализированного обучения языкам.

В лингвистике ИИ находит применение в различных аспектах, начиная от автоматической обработки естественного языка (Natural Language Processing – NLP) и заканчивая анализом больших текстовых массивов.

Одним из ключевых направлений является разработка алгоритмов машинного обучения, способных анализировать и интерпретировать языковые данные. Эти алгоритмы позволяют не только классифицировать тексты по различным категориям, но и выявлять скрытые закономерности и тенденции в использовании языка. Таким образом, ИИ становится мощным инструментом для лингвистических исследований, способствуя более глубокому пониманию структуры и функционирования естественных языков.

Применение ИИ в лингвистике также открывает новые возможности для разработки интеллектуальных систем, таких как диалоговые агенты и виртуальные ассистенты. Эти системы способны взаимодействовать с пользователями на естественном языке, предоставляя им информацию и выполняя различные задачи. Благодаря развитию технологий обработки естественного языка, такие системы становятся все более совершенными и могут использоваться в различных сферах.

Несмотря на все преимущества, применение ИИ в лингвистике сопряжено с рядом вызовов и ограничений. Одним из них является проблема интерпретации результатов, полученных с помощью ИИ-алгоритмов. В некоторых случаях результаты могут быть неоднозначными или требовать дополнительной проверки и анализа со стороны специалистов. Кроме того, существует риск предвзятости в алгоритмах, что может привести к искажению данных и некорректным выводам.

На сегодняшний день имеется широкий выбор цифровых инструментов, которые применяются в языкознании, речь идет, прежде всего, о вспомогательных инструментах для анализа и работы с текстами, которые используют лингвисты [1]. Во-первых, Named Entity Recognition (NER), который может применяться в качестве вспомогательного инструмента для анализа медиадискурса и литературных текстов. NER позволяет выявлять закономерности в употреблении имен собственных, анализировать связи между персонажами, а также определять частотность упоминания различных объектов. Этот метод особенно ценен для исследования стратегий убеждения в медиа, так как он позволяет глубже понять структуру и содержание коммуникативных актов.

Второй инструмент – Type Token Ratio (TTR) – широко используется для оценки лексического разнообразия текстов. TTR работает по формуле, где выявляется отношение числа типов (V – количество уникальных слов в тексте) к общему количеству слов (N – токенов) в тексте, иногда необходимо получившийся количественный показатель умножить на 100 для получения результата в процентах (например, в тексте объемом 100 слов (N) использовано 50 разных уникальных слов (V), то в этом случае TTR составит 50% ($TTR = V / N \times 100\%$). При этом TTR зависит от длины текста [2]. TTR применяется в изучении языковой личности, анализе переводов и локализации текстов под целевую аудиторию. Данный показатель позволяет исследовать сложность и разнообразие речи различных авторов, что является значимым аспектом дискурсивного анализа текстов.

Следующий цифровой инструмент – **FieldWorks Language Explorer™ (FLEx)** представляет собой специализированное программное обеспечение, предназначенное для комплексного анализа, документирования и моделирования языковых структур. Разработан SIL International. Как отмечают эксперты, данный инструмент является передовым решением в области лингвистики, позволяющим осуществлять детальный анализ грамматических систем, создание и управление многоязычными лексическими базами данных, поддержку широкого спектра языков и систем письма, позволяет работать с исчезающими языками [3]. FLEx обеспечивает возможность компиляции текстовых материалов, их структурированного хранения и анализа, что делает его незаменимым для исследователей, занимающихся документированием исчезающих языков, а также для специалистов в области теоретической и прикладной лингвистики.

Также можно отметить инструмент ResearchRabbit с Zotero, который значительно ускоряет процесс подбора и обработки источников для написания «Обзора литературы». Эта платформа предоставляет доступ к обширной базе оригинальных статей, что облегчает поиск релевантных данных и способствует повышению качества научных исследований.

Наконец, возможности ChatGPT позволяют проводить анализ мультимодальных текстов, исследовать стилистические приемы и тропы, а также осуществлять визуальный анализ, учитывая не только вербальный, но и визуальный контент. Этот инструмент открывает новые горизонты для междисциплинарных исследований и комплексного анализа коммуникативных практик.

Российский аналог ChatGPT – ГигаЧат=GigaChat (русскоязычная нейросеть от Сбера, <https://giga.chat/>) является одним из лидеров, который предложил широкие возможности – «от генерации текста и изображений до виртуальных ассистентов и аналитических платформ» [4]. Нейросетью можно пользоваться через Сбер ID, редактировать массивы текстов, использовать ИИ-агентов, делать подкасты на заданную тему и т.д. (фото 1, 2).

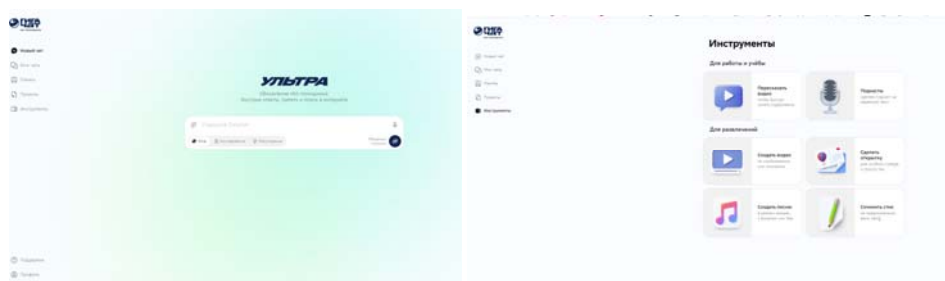


Фото 1, 2. – Русскоязычная нейросеть ГигаЧат от Сбера

Также выделяется своими возможностями сервис Kandinsky от «Сбера», который «разработан при поддержке Института искусственного интеллекта AIRI на объединенных датасетах SberAI и SberDevices» [5]. Нейросеть позволяет генерировать изображения с конкретными задачами, выбирает стиль изображений по текстовому запросу и т.д. [4].

Таким образом, потенциал ИИ в лингвистике остается значительным и продолжает привлекать внимание исследователей и разработчиков. Цифровые инструменты расширяются и видоизменяются, наполняются новыми функциональными возможностями. В будущем можно ожидать дальнейшего совершенствования технологий и расширения областей их применения, что будет способствовать развитию науки и практики в данной области.

БИБЛИОГРАФИЧЕСКИЕ ССЫЛКИ

1. Цифровые инструменты в лингвистике: как превратить ИИ в полезный инструмент, а не в конкурента. НИУ ВШЭ [сайт]. – URL: [https:// lang.hse.ru/news/1015919565.html](https://lang.hse.ru/news/1015919565.html) (дата обращения: 09.05.2026).
2. Type-Token Ratio. Поле цифровой дидактики [сайт]. – URL: https://digida.mgpu.ru/index.php/Type-Token_Ratio (дата обращения: 09.05.2026).
3. FieldWorks Language Explorer™: программное обеспечение для создания словарей. SIL Global. Технологии языка SIL [сайт]. – URL: <https://software.sil.org/ru/fieldworks/> (дата обращения: 10.05.2026).

4. Сбер научил своих виртуальных ассистентов общаться как люди. Благодаря продвинутой языковой модели. Проект Ferra.ru. Новости технологий и обзоры цифровой техники [сайт]. – URL: <https://www.ferra.ru/news/techlife/sber-nauchil-svoikh-virtualnykh-assistentov-obshatsya-kak-lyudi-12-11-2021.htm> (дата обращения: 10.05.2026).

5. Как пользоваться нейросетью Kandinsky в 2026 году. Hi-Tech [сайт]. – URL: <https://hi-tech.mail.ru/review/109792-kak-polzovatsya-nejrosetyu-kandinsky/> (дата обращения: 10.05.2026).

DIGITAL TOOLS IN LINGUISTICS: POSSIBILITIES OF APPLYING ARTIFICIAL INTELLIGENCE

G.R. Kamalova

*Academy of Sciences of the Republic of Bashkortostan, Ufa, Russian Federation
guldar@mail.ru*

The article discusses the potential of artificial intelligence (AI), which opens up new horizons for research and practical applications in the field of linguistics. It examines auxiliary tools for analyzing and working with texts, their capabilities, and their application in interdisciplinary research.

Keywords: Artificial intelligence (AI), linguistics, digital tools, ChatGPT, NER, TTR, FLEx, Gigachat.