

# ИСКУССТВЕННЫЙ ИНТЕЛЛЕКТ В РАБОТЕ С ТЕКСТОВЫМИ ДАННЫМИ: ПЕРСПЕКТИВЫ И ПОДХОДЫ

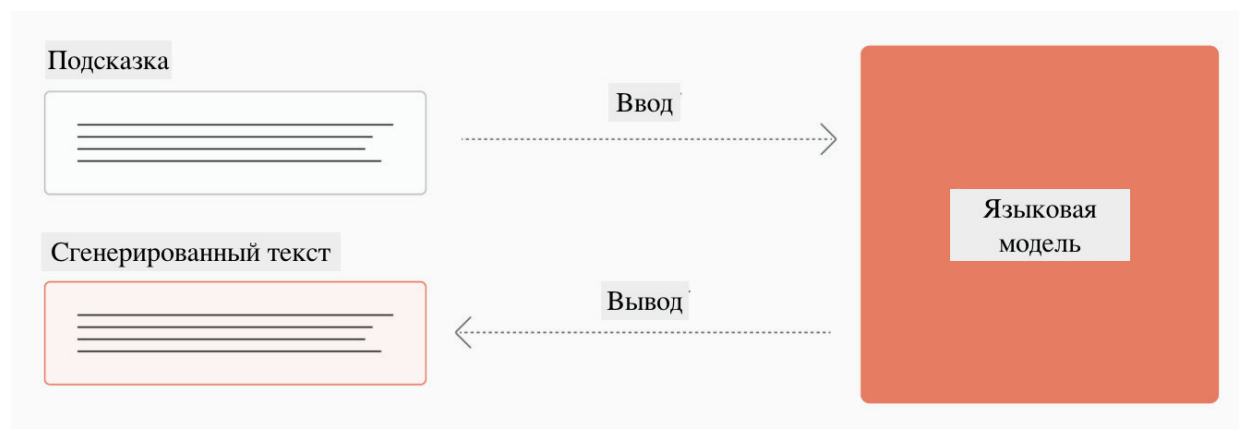
*Вуйкович А.Д., аспирантка кафедры № 65 «Анализ конкурентных систем»  
Национальный исследовательский ядерный университет «МИФИ»*

*Научные руководители: Артамонов А.А., к.т.н.,  
доцент кафедры «Анализ конкурентных систем», AAArtamonov@mephi.ru;  
Проничева Л.В., к.т.н., доцент кафедры «Анализ конкурентных систем»  
LVPrnicheva@mephi.ru*

**Аннотация:** В современном мире роль информационных технологий становится всё более значимой. Одним из наиболее перспективных подходов в работе с текстовыми данными является использование нейросетей, способных генерировать и анализировать контент на различных языках.

**Ключевые слова:** нейросеть, промпт-инжиниринг, техники промпт-инжиниринга, анализ данных.

**Промпт-инжиниринг** – это процесс разработки подробных инструкций и подсказок, то есть промптов, для языковой модели с целью получения пользователем желаемого результата запроса. Этот процесс схематически показан на рисунке [7]. В инженерии подсказок пользователь выбирает наиболее подходящие форматы, фразы, слова и символы, которые помогают искусственному интеллекту или языковой модели создать максимально приближенный к желаемому результат [3].



*Рис. 1. Принцип промпт-инжиниринга*

Для получения приближенного к желаемому ответа от нейросети используют элементы промпта [2] такие, как:

1. **задача или инструкция** – пользователь отправляет языковой модели конкретный запрос, который хочет получить на выходе;

2. **контекст** – информация, которая может направить нейросеть на нужный ответ;
3. **входные данные** – вопрос, на который пользователь хочет получить ответ;
4. **индикатор вывода** – формат, в котором пользователь желает увидеть ответ.

Не обязательно в запросах использовать все четыре элемента, так как это зависит от конкретной задачи.

Помимо элементов промпта используются различные **техники промпт-инжиниринга**, которые помогают пользователю настроить языковую модель или нейросеть:

1. **Zero-shot prompting (обучение без примеров)**: нейросети, такие как ChatGPT версии 3 и выше, обучаются на больших объемах данных, поэтому способны выполнять некоторые задачи без примеров, инструкций и классификаций.

2. **Few-shot prompting (обучение с примерами)** [3] используется для контекстного обучения, когда пользователь приводит пример в запросе, чтобы направить модель на желаемый формат ответа.

3. **Chain-of-Thought prompting (обучение, основанное на рассуждении)** [8] позволяет выполнять сложные рассуждения шаг за шагом.

4. **Generated Knowledge prompting (генерация знаний, как часть промпта)** используется для получения развернутого, обоснованного и доказательного ответа от нейросети.

5. **Prompt chaining (цепочка промптов)** [6] используется для решения задач, которые сложно выполнить с помощью одного большого промпта и они выполняются с помощью деления задачи на мелкие подзадачи. При первом запросе нейросеть выполняет одну подзадачу и далее на основе полученного результата пользователь может задать языковой модели следующий промпт, который будет основан на уже полученном ответе, и так до тех пор, пока задача не будет выполнена.

6. **Directional Stimulus prompting (направляющий промптинг)** [5] используется для суммаризации текста. Помимо самого текста и запроса пользователь отправляет нейросети подсказки – важные слова, фразы, цифры, даты, которые должны составить основу саммари.

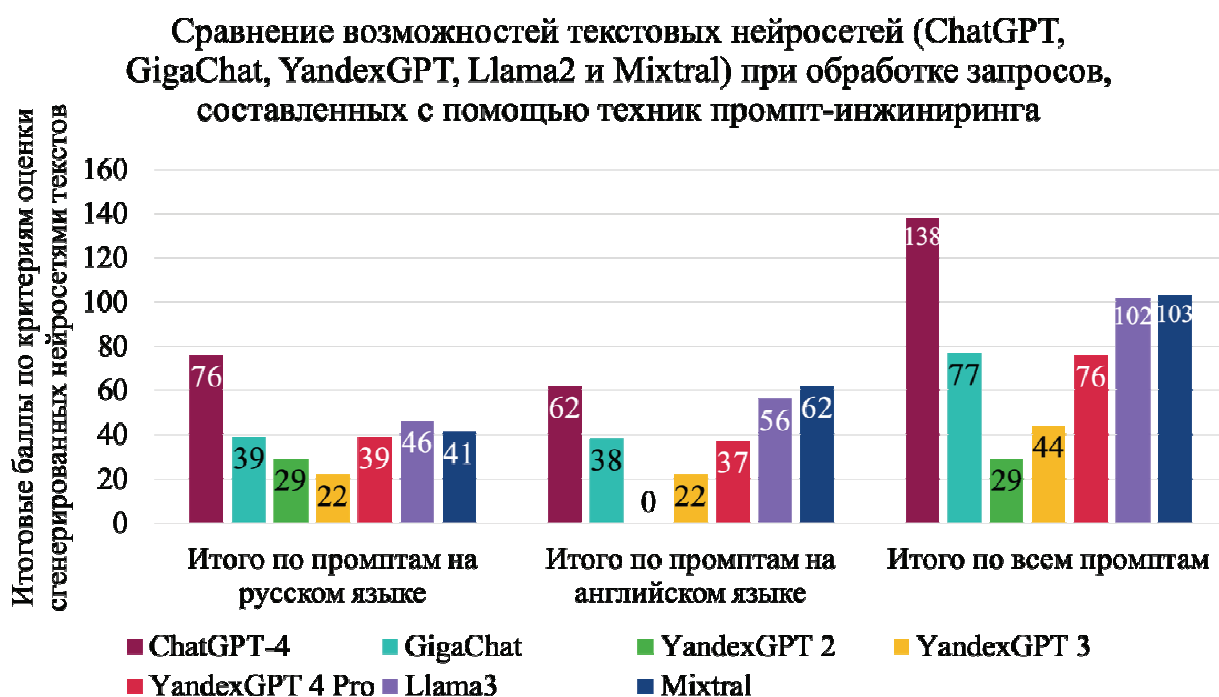
7. **СЕТО** позволяет получить ответ на экспертном уровне. Пользователь задает нейросети роль, в соответствии с которой получает ответ.

Для сравнения сгенерированных нейросетями текстов были составлены и использованы следующие критерии:

- достоверность данных (от 0 до 2 баллов);
- логичность и последовательность структуры текста (от 0 до 1 балла);

- использование специализированной терминологии (от 0 до 1 балла);
- грамматическая правильность текста (от 0 до 1 балла);
- стилистическая правильность текста (от 0 до 1 балла);
- соответствие полученного текста промпту (от 0 до 2 баллов).

По итогам оценки текстов, сгенерированных нейросетями, с помощью техник промпт-инжиниринга была получена диаграмма на рисунке 2 для большей наглядности результатов. Эксперименты с промптами проводились на русском и английском языках. На русском языке было сделано 10 промптов с различными примерами, а на английском – 8 промптов. Максимальный балл за результат, соответствующий всем критериям оценки сгенерированных нейросетями текстов, – 8 баллов. Соответственно, максимальный балл за выполнение всех промптов на русском языке – 80 баллов; максимальный балл за выполнение всех промптов на английском языке – 64 баллов.



*Рис. 2. Сравнение возможностей текстовых нейросетей (ChatGPT, GigaChat, YandexGPT, Llama3и Mixtral) при обработке запросов, составленных с помощью техник промпт-инжиниринга*

Таким образом, после проведения ряда экспериментов, можно заключить, что нейросеть ChatGPT-4o отлично справляется с промптами и на русском, и на английском языках. Нейросеть Giga Chat обучена на данных на русском и английском языках, как заявляют разработчики. По сравнению с предыдущей версией она работает с запросами на английском языке и отвечает на них. Yandex GPT 2 поддерживает только

русский язык и не реагирует на некоторые техники промпт-инжиниринга, такие как `promptchaining` (цепочка промптов) и `directional stimulus prompting` (направляющий промптинг). Yandex GPT 3 – более усовершенствованная версия Yandex GPT 2, которая поддерживает английский язык и лучше реагирует на запросы пользователя, но техника `prompt chaining` и все запросы, состоящие из более 1000 символов, остаются для нейросети нерешаемой задачей. YandexGPT 4 Pro лучше реагирует на запросы пользователя и различные техники промпт-инжиниринга. Как и в предыдущих версиях, запросы от 1000 символов не поддерживает. Локальные нейросети Llama3и Mixtral реагируют на русскоязычные и англоязычные промпты. Стоит отметить, что нейросеть ChatGPT-4o и локальные нейросети Llama 3и Mixtral отвечают на запросы, включающие в себя политический или околополитический контекст, в отличие от нейросетей Giga Chat, Yandex GPT 2, 3 и 4.

По результатам сравнения текстовых нейросетей были выявлены сильные и слабые стороны. На их основе можно составить рекомендации по использованию той или иной нейросети для получения определенного результата:

1. Выбор нейросети в соответствии с языком запроса, текста или желаемого результата:

- английский язык – Chat GPT, Giga Chat, Yandex GPT 3 и 4, Llama 3, Mixtral;

- русский язык – Chat GPT, Giga Chat, Yandex GPT 2–4, Llama 3 и Mixtral.

2. Выбор техники промпт-инжиниринга и составление промпта.

3. Использование подкрепляющего запроса или переформулирование промпта (в случае, если ответ нейросети не устроил пользователя).

4. Проверка результата на ошибки и неточности.

Перспективы использования искусственного интеллекта в работе с текстовыми данными обширны и разнообразны. Технологии позволяют оптимизировать составление аналитических отчетов, улучшая точность и снижая временные затраты. Обработка и структуризация текстовой информации обеспечивает упорядочивание больших массивов данных, что важно для аналитики и принятия решений. Визуализация данных с использованием искусственного интеллекта упрощает восприятие сложных аналитических отчетов. Применение трансформеров, таких как Chat GPT, Giga Chat, Yandex GPT, Llama, Mixtral и других, открывает новые возможности в генерации высококачественного контента. Кроме того, искусственный интеллект активно используется для повышения точности машинного перевода и автоматического реферирования.

Проведённое исследование подтверждает, что современные нейросетевые модели способны создавать понятный и качественный контент для различных аудиторий, предоставляя новые возможности для автоматизации работы с текстом.

#### **Литература и источники:**

1. Информационная аналитика в современном социально-политическом процессе: теория и практика / Ремарчук В.Н., Бочарников И.В., Артемьев А.А., Галаганова С.Г., Гришнова Е.Е., Егоров В.Г., Карась Р.А., Катков О.Н., Ламинина О.Г., Смутьский С.В., Шевчун В.Н., Урсул В.И. Москва, 2024.

2. Руководство по промпт-инжинирингу [Электронный ресурс] // Prompt Engineering Guide: [сайт]. [2024]. URL: <https://www.promptingguide.ai/ru>.

3. Что такое инженерия подсказок? [Электронный ресурс] // Amazon Web Services: [сайт]. URL: <https://aws.amazon.com/ru/prompts-engineering/>.

4. Brown T.B., Mann B., Ryder N., Subbiah M., Kaplan J., Dhariwal P., Neelakantan A., Shyam P., Sastry G., Askell A. и др. Language Models are Few-Shot Learners // arXiv. 2020. URL: <https://arxiv.org/pdf/2005.14165.pdf>.

5. Li Z., Peng B., He P., Galley M., Gao J., Yan X. Guiding Large Language Models via Directional Stimulus Prompting // arXiv. 2023. URL: <https://arxiv.org/pdf/2302.11520.pdf>.

6. Prompt chaining [Электронный ресурс] // Anthropic: [сайт]. [2024]. URL: <https://docs.anthropic.com/claude/docs/prompt-chaining>.

7. Prompt Engineering [Электронный ресурс] // cohere docs: [сайт]. [2023]. URL: <https://docs.cohere.com/docs/prompt-engineering>.

8. Wei J., Wang X., Schuurmans D., Bosma M., Ichter B., Xia F., Chi E.H., Le Q.V., Zhou D. Chain-of-Thought Prompting Elicits Reasoning in Large Language Models // arXiv. 2023. URL: <https://arxiv.org/abs/2303.11903>.