УДК 004.82

МЕТОДОЛОГИЯ ПОСТРОЕНИЯ ЭФФЕКТИВНЫХ ПРОМПТОВ ДЛЯ ГЕНЕРАТИВНЫХ ЯЗЫКОВЫХ МОДЕЛЕЙ НА ПРИМЕРЕ CHATGPT

Василевский В.А. – магистрант,
Фисков С.А. – магистрант,
Кузбасский государственный технический университет имени Т.Ф. Горбачева
Россия, г. Кемерово

Стремительное развитие технологий искусственного интеллекта, в частности генеративных языковых моделей (Large Language Models, LLM), привело к их активному внедрению в различные сферы человеческой деятельности. Этот процесс детерминирован резким ростом вычислительной мощности, объемов тренировочных данных и архитектурными инновациями, такими как механизм внимания (attention mechanism), лежащий в основе трансформеров [1]. Модель ChatGPT, как наиболее наглядный пример коммерциализированной LLM, демонстрирует способность генерировать связные и контекстуально релевантные тексты, что обусловило ее популярность [2]. Однако, как показывают многочисленные исследования, потенциал подобных моделей раскрывается в полной мере только при условии грамотного формулирования входных запросов – промптов, что выдвигает на первый план дисциплину промт-инжиниринга.

Составление эффективных промптов представляет собой сложную коммуникативную и когнитивную задачу, находящуюся на стыке компьютерной лингвистики, психологии человеко-машинного взаимодействия и теории управления. С когнитивной точки зрения, данный процесс можно охарактеризовать как "перевод" интенции пользователя на формальный язык, понятный статистической модели [3]. В академической и профессиональной среде этот навык трансформируется из факультативного в критически важный [4]. Целью данной статьи является синтез лучших практик промт-инжиниринга на основе анализа современных исследований и представление структурированной методологии построения эффективных промптов.

Эффективный промпт – это детализированная инструкция, содержащая исчерпывающий набор параметров для достижения желаемого результата. Его качество определяется системой взаимосвязанных принципов, эмпирически выведенных из анализа взаимодействия "человек-ИИ". Конкретность и однозначность минимизируют энтропию выходного распределения модели, сужая пространство возможных ответов [5]. Контекстуализация компенсирует отсутствие у модели устойчивого долговременного состояния (statefulness) между сессиями, требуя от пользователя явного включения релевантных фактов и предыстории в каждый запрос [6]. Принцип итеративности отражает итеративный характер решения сложных задач, при котором промпт эволюционирует

от общего к частному через серию уточнений, что методологически близко к технике "уточнения запроса" (query refinement) в информационном поиске [7].

Для систематизации процесса конструирования запросов предлагается использовать универсальную структурную формулу, декомпозирующую запрос на семантические компоненты: [Роль] + [Задача] + [Детали/Контекст] + [Формат] + [Тон]. Назначение Роли (например, «эксперт-вирусолог») модифицирует внутренние приоритеты модели, активируя соответствующие домену паттерны в ее весах [8]. Задача, выраженная императивным глаголом действия («синтезируй», «сравни», «критикуй»), задает целевую функцию для генерации. Детали/Контекст снабжают модель необходимой информацией для разрешения неоднозначностей. Указание Формата вывода (JSON, Markdown, таблица) структурирует данные, облегчая их последующую машинную обработку. Определение Тона позволяет управлять стилистической регистровой принадлежностью текста. Для задач повышенной сложности в промпт инкорпорируются дополнительные компоненты: ограничения (constraints), примеры (few-shot learning) и запреты, которые действуют как негативные инструкции, подавляя генерацию нежелательного контента [9].

Практическая реализация методологии демонстрирует свою эффективность на примере промпта: «Выступайте в роли магистранта по прикладной информатике. Напишите введение к научной статье на тему "Влияние промт-инжиниринга на качество генерации текста в LLM". Объем – 300 слов, используйте подзаголовки. Стиль – научный, избегайте разговорных выражений». Данный промпт явно задает все критические параметры, минимизируя вероятность нерелевантного ответа.

Таксономия промптов по сценариям использования позволяет выделить несколько ключевых типов. Образовательные промпты используются для объяснения концепций и генерации тестовых вопросов, реализуя функцию адаптивного тьютора [10]. Научно-исследовательские применяются для подготовки планов статей и написания аннотаций, выступая инструментом интеллектуального ассистента. Творческие служат для генерации художественных текстов и сценариев, а технические – для написания, комментирования и отладки кода, демонстрируя высокую эффективность в задачах автодополнения (code completion) [11]. Бизнес-ориентированные промпты помогают в составлении отчетов и при проведении SWOT-анализа. Специализированные техники, такие как шаблонизация (создание шаблонов для повторяющихся задач), управление уникальностью (явное требование к парафразированию) и корректировка через обратную связь (refinement through feedback), представляют собой продвинутые методы управления выводом модели [12].

Несмотря на мощь современных LLM, важно осознавать их фундаментальные гносеологические ограничения. Модели не обладают сознанием или пониманием в человеческом смысле и подвержены генерации правдоподобных, но семантически или фактологически ложных утверждений – «галлюцинаций» (hallucinations) [13]. Это обусловлено их статистической природой и требует от пользователя осуществления строгой верификации всех

полученных данных. С этической точки зрения использование промт-инжиниринга в академической среде требует формулирования четких регуляторных рамок [14]. Данный инструмент не должен служить для обхода требований к оригинальности работ, а должен рассматриваться как средство усиления интеллектуальных возможностей исследователя (intelligence amplification). Ответственность за достоверность, качество и атрибуцию финального текста остается на человеке.

Промт-инжиниринг утвердился как критически важный междисциплинарный навык в эпоху повсеместного распространения генеративного ИИ. Представленная в статье методология, основанная на структурной формуле и ключевых принципах ясности, контекста и итеративности, предоставляет систематизированный инструментарий для эффективного взаимодействия с языковыми моделями. Освоение этих техник позволяет перейти от пассивного потребления информации, генерируемой моделью стохастически, к активному, детерминистическому процессу управления генерацией контента, соответствующего строгим академическим и профессиональным стандартам. Дальнейшие исследования в данной области видятся в направлении разработки формальной семантики промптов и создания специализированных моделей-интерпретаторов, способных к более глубокому пониманию интенции пользователя.

Список литературы:

1. Большая языковая модель // Википедия. – URL: https://ru.wikipedia.org/wiki/ Большая_языковая_модель (дата обращения: 19.11.2025).

2. Что такое большая языковая модель (LLM)? // AWS. – URL: https://aws.amazon.com/ru/what-is/large-language-model/ (дата обращения: 19.11.2025).

3. С развитием LLM рынок труда, образование и коммуникации ждет масштабная трансформация // Татар-информ. – URL: https://www.tatar-inform.ru/news/s-razvitiem-llm-rynok-truda-obrazovanie-i-kommunikacii-zdet-masstabnaya-transformaciya-5980935 (дата обращения: 19.11.2025).

4. Большие языковые модели в бизнесе: как компании используют LLM для роста и оптимизации // TechCult. – URL: https://www.techcult.ru/promo/15365-bolshie-yazykovye-modeli-v-biznese (дата обращения: 19.11.2025).

5. LLM (Large Language Model) // Энциклопедия BigDataSchool. – URL: https://bigdataschool.ru/wiki/large-language-models/ (дата обращения: 19.11.2025).

6. Практическое руководство по промт-инжинирингу // Хабрахабр. – URL: https://habr.com/ru/articles/789066/ (дата обращения: 19.11.2025).

7. ИИ в образовании: сможет ли большая языковая модель стать учителем // РБК Тренды. –

URL: https://trends.rbc.ru/trends/industry/cmrm/67fe5ba19a79477651fc7b47 (дата обращения: 19.11.2025).

8. Большие языковые модели (LLM) в здравоохранении // Webiomed. – URL: https://webiomed.ai/blog/perspektivy-primeneniia-llm-v-zdravookhranenii/ (дата обращения: 19.11.2025).

9. LLM требуют внимания к образованию, мышлению и эмоциям // ANNS.URL: https://anns.ru/articles/news/2025/04/17/llm_trebujut_vnimanija_k_obrazovaniju__mishleniju_i_emotsijam (дата обращения: 19.11.2025).

10. Зачем гуманитариям изучать искусственный интеллект // Фоксфорд. – URL: https://media.foxford.ru/artificial-intelligence (дата обращения: 19.11.2025).

11. Технический промт-инжиниринг для разработчиков // Хабрахабр. – URL: https://habr.com/ru/companies/ruvds/articles/769044/ (дата обращения: 19.11.2025).

12. Глушков, П.Н. Продвинутые методы управления выводом больших языковых моделей. – Санкт-Петербург: Питер, 2024. – 298 с.

13. Проблема "галлюцинаций" в больших языковых моделях и методы их детекции // Открытые системы. СУБД. – URL: https://www.osp.ru/os/2023/02/13070306 (дата обращения: 19.11.2025).

14. Этические принципы использования искусственного интеллекта в научной деятельности / Ассоциация "Национальный этический комитет". – Москва, 2023. – 45 с.