

# Доверяйте на свой страх и риск: смешанное исследование способности крупных языковых моделей генерировать артефакты системной инженерии, похожие на экспертные, и характеристика режимов их сбоя

Дата: 2025-02-13 00:00:00

Ссылка на исследование: <https://arxiv.org/pdf/2502.09690>

Рейтинг: 75

Адаптивность: 80

## Ключевые выводы:

Исследование направлено на оценку способности многоцелевых больших языковых моделей (LLM) генерировать артефакты системной инженерии (SE), сравнимые с созданными экспертами-людьми. Основные результаты показывают, что хотя LLM могут создавать артефакты, семантически похожие на экспертные, они демонстрируют серьезные недостатки в качестве, включая преждевременное определение требований, необоснованные числовые оценки и склонность к избыточной детализации.

## Объяснение метода:

Исследование высоко полезно для пользователей LLM благодаря выявлению конкретных паттернов ошибок при генерации артефактов системной инженерии. Идентифицированные "режимы отказа" (преждевременное определение требований, необоснованные оценки, чрезмерная детализация) и рекомендации по формулированию эффективных промптов имеют прямую практическую ценность для критической оценки ответов LLM и более эффективного взаимодействия с ними.

**## Ключевые аспекты исследования:** 1. **Исследование способности LLM генерировать артефакты системной инженерии:** Авторы проверили, могут ли многоцелевые LLM (GPT-3.5, GPT-4, Claude) генерировать артефакты системной инженерии, похожие на созданные экспертами-людьми, без какой-либо дополнительной настройки или обучения.

**Методология сравнения:** Использован смешанный подход — количественное сравнение с помощью алгоритма MAUVE для измерения семантического сходства

текстов и качественный анализ для выявления содержательных различий между AI-генерированными и экспертными артефактами.

**Влияние промптов:** Исследование показало, что формулировка запросов критически важна для качества результатов. Более конкретные промпты с указанием желаемой длины и структуры ответа значительно повышают качество генерируемых артефактов.

**Выявленные режимы отказа:** Определены три ключевых паттерна ошибок LLM: преждевременное определение требований (неспособность отличить потребности от требований), необоснованные числовые оценки и склонность к чрезмерной детализации.

**Предостережение о риске доверия к LLM:** Хотя LLM могут создавать тексты, которые выглядят профессионально и похожи на экспертные, они содержат серьезные ошибки, которые могут привести к катастрофическим последствиям при принятии проектных решений.

## Дополнение:

### Применимость методов исследования в стандартном чате

Исследование не требует дообучения или специального API для применения основных выводов. Все выявленные концепции и подходы могут быть реализованы в стандартном диалоге с LLM.

#### Ключевые концепции и подходы для стандартного чата:

**Эффективное составление промптов:** Конкретизация запроса (указание точного контекста) Указание желаемой длины ответа Определение структуры ответа Поэтапное уточнение запросов

**Распознавание режимов отказа:**

Проверка числовых оценок на обоснованность Критическая оценка преждевременных требований Фильтрация чрезмерной детализации

**Стратегии улучшения результатов:**

Разделение сложных задач на более простые Запрос обоснования для любых числовых оценок Явное указание уровня абстракции в запросе Использование LLM для задач, где они показывают лучшие результаты (форматирование, суммаризация)

**Применение "экспертной проверки":**

Запрос у LLM критической оценки своего предыдущего ответа Просьба указать ограничения и потенциальные проблемы в предоставленном решении Применяя эти

концепции, пользователи могут значительно улучшить качество взаимодействия с LLM в стандартном чате, без необходимости в специальных API или дообучении.

## Анализ практической применимости: **1. Преждевременное определение требований - Прямая применимость:** Высокая. Понимание этого режима отказа помогает пользователям критически оценивать ответы LLM при формулировании проблем и требований. Пользователи могут запрашивать у LLM общие границы проблемы, а не детальные требования на ранних этапах. - **Концептуальная ценность:** Очень высокая. Исследование показывает фундаментальное ограничение LLM в понимании процесса системной инженерии, что помогает пользователям осознать необходимость собственной экспертизы при работе с моделями. - **Потенциал для адаптации:** Средний. Пользователи могут адаптировать свои промпты, чтобы явно указывать LLM, что они ищут общие границы проблемы, а не конкретные требования.

**2. Необоснованные числовые оценки - Прямая применимость:** Высокая. Пользователи должны игнорировать конкретные числовые оценки, предоставляемые LLM без обоснования, или использовать их только как отправную точку для собственного анализа. - **Концептуальная ценность:** Высокая. Исследование ясно демонстрирует, что LLM не имеют инженерных аналитических возможностей и не могут выполнять расчеты, что помогает установить правильные ожидания. - **Потенциал для адаптации:** Средний. Пользователи могут адаптировать свои запросы, чтобы получать качественную информацию или методологические рекомендации вместо конкретных числовых оценок.

**3. Склонность к чрезмерной детализации - Прямая применимость:** Средняя. Пользователи могут научиться распознавать избыточную детализацию в ответах LLM и отфильтровывать ненужную информацию. - **Концептуальная ценность:** Высокая. Понимание этого режима отказа помогает пользователям более критически оценивать ответы LLM и уметь отделять полезную информацию от "шума". - **Потенциал для адаптации:** Высокий. Пользователи могут запрашивать более краткие и целенаправленные ответы, а также учиться формулировать промпты, которые минимизируют эту проблему.

**4. Влияние формулировки промптов - Прямая применимость:** Очень высокая. Результаты исследования дают практические рекомендации по формулированию эффективных промптов, включая указание желаемой длины ответа и конкретных элементов, которые должны быть включены. - **Концептуальная ценность:** Высокая. Исследование демонстрирует, что LLM лучше работают с конкретными, хорошо структурированными запросами, что помогает пользователям понять важность формулировки промптов. - **Потенциал для адаптации:** Очень высокий. Пользователи могут применять принципы эффективного составления промптов к различным задачам системной инженерии и другим областям.

**5. Понимание ограничений LLM - Прямая применимость:** Высокая. Исследование дает четкое представление о том, в каких аспектах системной инженерии LLM могут быть полезны (форматирование, суммаризация), а в каких они ненадежны (аналитические расчеты, определение требований). - **Концептуальная ценность:**

Очень высокая. Результаты помогают пользователям установить реалистичные ожидания от LLM и понять необходимость экспертной проверки их выводов. - **Потенциал для адаптации:** Высокий. Понимание ограничений LLM может быть применено к различным задачам и областям, помогая пользователям определить, когда и как использовать эти инструменты.

## Prompt:

Использование исследования LLM в системной инженерии для создания эффективных промптов ## Ключевые уроки из исследования

Исследование показывает, что LLM могут создавать артефакты системной инженерии, похожие на экспертные, но имеют определенные режимы отказа: - Преждевременное определение требований - Необоснованные числовые оценки - Избыточная детализация

При этом качество результатов сильно зависит от структуры и конкретности промптов.

## Пример эффективного промпта

[=====] # Задание для создания артефакта системной инженерии

## Контекст Я работаю над проектом автономного робота для инспекции трубопроводов. Мне нужно разработать структурированный артефакт системной инженерии.

## Требования к формату - Используй формат IEEE для документации требований - Ограничь длину документа до ~500 слов - Используй иерархическую нумерацию разделов - Представь информацию в виде структурированных списков

## Содержание 1. Сначала определи общий контекст проблемы и границы системы 2. Выделяй требования только на основе предоставленной информации, не добавляй преждевременных требований 3. Не указывай конкретные числовые значения, если они не предоставлены 4. Для неопределенных параметров укажи диапазоны или методологию определения 5. Структурируй документ по следующим разделам: - Обзор системы - Функциональные требования - Нефункциональные требования - Ограничения и допущения - Интерфейсы

## Важно - Не вводи необоснованные числовые оценки - Не добавляй избыточных деталей, которые могут ограничить пространство проектирования - Четко отделяй факты от предположений [=====]

## Почему этот промпт работает

**Конкретность и структура:** Исследование показало, что MAUVE-оценки выросли с 0.0000 до 0.9932 при использовании более структурированных промптов.

**Предотвращение режимов отказа:**

Явно запрещает преждевременное определение требований Предупреждает о недопустимости необоснованных числовых оценок Ограничивает избыточную детализацию

**Четкие указания по формату:** Использует преимущество LLM следовать указаниям по форматированию.

**Разделение контекста и требований:** Следует рекомендации использовать LLM для обобщения контекста, а не для определения конкретных требований.

Такой подход к составлению промптов позволяет максимизировать полезность LLM при создании артефактов системной инженерии, минимизируя их известные недостатки.