

Изучение влияния больших языковых моделей на пользовательские истории, созданные студентами, и тестирование приемки в разработке программного обеспечения

Дата: 2025-02-04 00:00:00

Ссылка на исследование: <https://arxiv.org/pdf/2502.02675>

Рейтинг: 75

Адаптивность: 85

Ключевые выводы:

Исследование направлено на изучение влияния LLM (больших языковых моделей) на способность студентов-программистов преобразовывать отзывы пользователей в пользовательские истории (user stories) в рамках Agile-методологии. Главные результаты показали, что LLM значительно улучшают способность студентов создавать критерии приемки (acceptance criteria) и повышают ценность пользовательских историй, но при этом студенты без помощи LLM лучше справляются с созданием историй подходящего объема.

Объяснение метода:

Исследование дает конкретные данные о том, где LLM помогают (критерии приемки +88%, ценность формулировок +23%) и где мешают (определение объема -24%). Методика работы с LLM универсально применима. Пользователи получают важное концептуальное понимание: LLM эффективны для детализации, но требуют контроля объема задач.

Ключевые аспекты исследования: 1. Исследование изучает влияние LLM на способность студентов трансформировать отзывы пользователей в пользовательские истории (user stories) в контексте разработки программного обеспечения.

Студенты создавали пользовательские истории дважды: без помощи LLM и с помощью LLM (ChatGPT 3.5), следуя принципам INVEST (Independent, Negotiable, Valuable, Estimable, Small/Scope, Testable).

Использование LLM значительно улучшило способность студентов создавать критерии приемки (acceptance criteria) для пользовательских историй (+88%) и

повысило ценность (value) историй (+23%).

Однако студенты, не использующие LLM, лучше справлялись с созданием историй подходящего объема (small/score) — использование LLM снизило показатели по этому параметру на 24%.

Не было обнаружено статистически значимых различий по другим атрибутам INVEST (независимость, возможность обсуждения и возможность оценки).

Дополнение:

Применимость методов в стандартном чате

Исследование не требует дообучения моделей или специального API для применения описанных методов. Все подходы можно использовать в стандартном чате с любой современной LLM. Исследователи использовали ChatGPT 3.5, который доступен широкой аудитории.

Ключевые концепции для применения в стандартном чате:

Структурированные промпты для пользовательских историй: Создание отдельного промпта для каждой истории Включение контекста приложения Указание конкретных принципов (например, INVEST) Запрос на генерацию критериев приемки

Критический анализ результатов:

Особенно внимательная проверка объема предложенных задач Акцент на использование LLM для детализации критериев успеха Ручное редактирование результатов при необходимости

Избирательное использование LLM:

Использование для структурирования критериев приемки Использование для повышения ценности формулировок Самостоятельное определение объема задач
Ожидаемые результаты: - Более детальные и тестируемые критерии успеха для любых проектов - Более ценные и понятные формулировки задач - Избежание проблем с завышенным объемом работ при правильном применении

Анализ практической применимости: 1. **Улучшение критериев приемки с помощью LLM** - Прямая применимость: Высокая. Пользователи могут применять LLM для создания детальных критериев приемки в любых проектных задачах, не только в программной разработке. - Концептуальная ценность: Средняя. Понимание того, что LLM особенно эффективны в задачах, требующих структурирования и детализации критериев. - Потенциал для адаптации: Высокий. Подход может быть адаптирован для формулирования критериев успеха в любых проектах, планирования, создания чек-листов.

Повышение ценности историй пользователей Прямая применимость: Средняя.

Пользователи могут использовать LLM для улучшения формулировок и повышения ясности требований. Концептуальная ценность: Высокая. Исследование показывает, что LLM помогают сфокусироваться на ценности для конечного пользователя. Потенциал для адаптации: Высокий. Метод применим к формулировке любых задач, где важно подчеркнуть ценность результата.

Ограничения в определении объема

Прямая применимость: Высокая. Пользователи должны критически оценивать предложения LLM по объему задач. Концептуальная ценность: Высокая. Важное понимание того, что LLM склонны к расширению объема задач. Потенциал для адаптации: Средний. Знание о необходимости корректировки объема задач применимо во всех сферах планирования.

Методика направленного использования LLM

Прямая применимость: Высокая. Описанный подход работы с LLM (создание отдельного промпта, предоставление контекста, анализ и редактирование) может быть непосредственно использован. Концептуальная ценность: Высокая. Понимание необходимости структурированного взаимодействия с LLM. Потенциал для адаптации: Высокий. Методика применима к любым задачам взаимодействия с LLM.

Целенаправленное использование LLM для конкретных задач

Прямая применимость: Высокая. Вывод о том, что LLM следует использовать для конкретных задач, а не универсально. Концептуальная ценность: Очень высокая. Понимание сильных и слабых сторон LLM в различных аспектах задач. Потенциал для адаптации: Высокий. Принцип селективного использования LLM применим во всех областях.

Prompt:

Применение исследования о влиянии LLM на пользовательские истории ##
Ключевые знания из исследования

Исследование показало, что использование LLM: - Улучшает качество **критериев приемки** (+88%) - Повышает **ценность** пользовательских историй (+23%) - Но ухудшает **размер** историй (делает их слишком большими, -24%)

Пример промпта на основе исследования

[=====] Помоги мне создать пользовательскую историю и критерии приемки на основе следующего отзыва пользователя: [ОТЗЫВ ПОЛЬЗОВАТЕЛЯ]

При создании следуй этим правилам: 1. Сформулируй историю в формате: "Как [роль пользователя], я хочу [функциональность], чтобы [ценность/польза]" 2. Сделай историю ЦЕННОЙ - четко объясни, какую пользу получит пользователь 3. Разработай ПОДРОБНЫЕ критерии приемки, которые можно легко преобразовать в

тест-кейсы 4. ВАЖНО: Убедись, что история имеет небольшой объем и сфокусирована на одной конкретной функции 5. Следуй принципам INVEST (Independent, Negotiable, Valuable, Estimable, Small, Testable)

После создания истории, пожалуйста, проверь, не является ли она слишком большой, и при необходимости раздели на несколько меньших историй. [=====]

Почему этот промпт работает

Данный промпт учитывает основные выводы исследования:

Усиливает сильные стороны LLM: Запрашивает подробные критерии приемки (область, где LLM показали +88% улучшение) Подчеркивает важность ценности для пользователя (область с +23% улучшением)

Компенсирует слабые стороны LLM:

Специально обращает внимание на необходимость небольшого размера истории
Просит проверить и разделить историю, если она получилась слишком большой

Использует структурированный подход, опираясь на принципы INVEST, которые были частью методологии исследования

Такой подход позволяет максимизировать преимущества LLM, минимизируя их недостатки при создании пользовательских историй.