

Искусственный интеллект в качественном анализе данных фокус-групп: сравнительное исследование ручного и автоматизированного подходов

Маргарита Рафаильевна Хуснутдинова

Московский психолого-педагогический университет, Москва, Россия, husnutdinovaMR@mgppu.ru

Ирина Игоревна Баранова

независимый исследователь, Россия, Baranovairinaa99@gmail.com

Оксана Геннадьевна Зубова

Московский государственный университет имени М. В. Ломоносова, Москва, Россия, zubovaoksana@bk.ru

Аннотация. В статье рассмотрены возможности применения искусственного интеллекта в качественном анализе данных фокус-групп. Реализован эксперимент с тремя исследователями: первый проводил ручной анализ транскриптов фокус-групп, второй использовал модель DeepSeek, третий сравнивал результаты. Эмпирическая база включает материалы двух фокус-групп с подростками, собранные в Пермском крае в 2022 г. Теоретической основой выступил участвующий подход социологии детства с операционализацией модели участия Р. Олсен. Выявлены преимущества LLM в структурировании данных и автоматизации рутинных задач, однако обнаружены существенные ограничения: искажения цитат, генерация несуществующей информации и односторонность анализа. Сделан вывод, что LLM не могут заменить человека в глубоком анализе качественных данных, но эффективны как инструмент первичной обработки. Разработан алгоритм работы с LLM для анализа качественных данных.

Ключевые слова: искусственный интеллект, большие языковые модели, промпт-инженеринг, качественный анализ, соучаствующее проектирование, фокус-группы, социология детства

Для цитирования: Хуснутдинова М. Р., Баранова И. И., Зубова О. Г. Искусственный интеллект в качественном анализе данных фокус-групп: сравнительное исследование ручного и автоматизированного подходов // Ойкумена. Регионоведческие исследования. 2025. Т. 19. № 3. С. 73–83.
<https://doi.org/10.63973/1998-6785/2025-3/73-83>

Original article
<https://doi.org/10.63973/1998-6785/2025-3/73-83>

Artificial Intelligence in Qualitative Focus Group Data Analysis: A Comparative Study of Manual and Automated Approaches

Margarita R. Khusnutdinova

Moscow State University of Psychology and Education, Moscow, Russia, husnutdinovaMR@mgppu.ru

Irina I. Baranova

Independent Researcher, Russia, Baranovairinaa99@gmail.com

Oksana G. Zubova

Lomonosov Moscow State University, Moscow, Russia, zubovaoksana@bk.ru

Abstract. The article examines artificial intelligence applications in qualitative focus group data analysis. An experiment involved three researchers: the first conducted manual analysis of focus group transcripts, the second used the DeepSeek model, the third compared results. The empirical base includes materials from two focus groups with teenagers collected in Perm region in 2022. The theoretical basis was the participatory approach of childhood sociology with R. Olsen's participation model operationalization. LLM advantages in data structuring and routine task automation were identified, however, significant limitations were found quote distortions, generation of non-existent information and one-sided analysis. It was concluded that LLM cannot replace humans in deep qualitative data analysis but are effective as primary processing tools. An algorithm for working with LLM for qualitative data analysis was developed.

Key words: artificial intelligence, large language models, prompt engineering, qualitative analysis, participatory design, focus groups, sociology of childhood

For citation: Khusnutdinova M. R., Baranova I. I., Zubova O. G. Artificial Intelligence in Qualitative Focus Group Data Analysis: A Comparative Study of Manual and Automated Approaches // Ojkumena. Regional Researches. 2025. Vol. 19. No. 3. P. 73–83. <https://doi.org/10.63973/1998-6785/2025-3/73-83>

Введение

Искусственный интеллект (ИИ) становится неотъемлемой частью научных исследований. Учёные применяют его практически на всех этапах – от генерации идеи до создания итоговых презентаций. Это порождает актуальные научные дискуссии о роли исследователя и его возможном замещении при быстром развитии инновационных технологий, связанных с ИИ. Некоторые эксперты видят в нейросетях огромный потенциал ускорения исследований, когда ИИ может заменить рутинный труд, снизить трудоёмкость часто повторяющихся операций, но основную интеллектуальную работу исследователь делает самостоятельно. Другие видят возможность сотрудничества, самостоятельного решения ИИ более сложных задач, связанных с анализом и интерпретацией данных [1; 2].

В арсенале современных исследователей качественных данных присутствует широкий спектр специализированных программ, таких как NVivo, MAXQDA, Atlas.ti, LIWC, IBM Watson Natural Language Understanding, MonkeyLearn и др. Эти инструменты предлагают автоматизацию транскрипции и кодирования, обнаружение тематических паттернов с помощью методов обработки естественного языка (NLP), анализ эмоционального фона дискуссии и визуализацию результатов.

Появление больших языковых моделей (Large Language Models, LLMs), таких как GPT (Generative Pre-trained Transformer), DeepSeek, Grok4, Gemini (семейство моделей от DeepMind (часть Google)), Claude (семейство моделей от Anthropic), и других знаменует собой новый подход к обработке и интерпретации качественных данных. Традиционные методы анализа, основанные на ручном кодировании и экспертной интерпретации, постепенно дополняются и трансформируются. Это позволяет не только ускорить обработку больших массивов текстовой информации, но и повысить качество и глубину аналитики за счёт выявления скрытых паттернов, эмоциональных оттенков и семантических связей.

LLMs представляют собой нейронные сети, обученные на больших объемах текстовых данных, что позволяет им генерировать связный и контекстуально релевантный текст, понимать естественный язык, отвечать на вопросы, резюмировать информацию и выполнять другие сложные лингвистические задачи [8]. Нейросети – это разновидность искусственного интеллекта, представляющие собой компьютерные модели, вдохновленные работой человеческого мозга. Хотя их функционирование схоже с мозгом, оно не является его точной копией. В мозге сигналы передаются через нейроны, тогда как в нейросетях используются вычислительные элементы, которые взаимодействуют между собой, обмениваются данными и производят результат.

Уникальность нейросетей в контексте анализа качественных данных заключается в способности работать с неструктурированным текстом на естественном языке – именно той форме, в которой представлены данные фокус-групп. В отличие от традиционных специализированных программ для качественного анализа, которые имеют платный доступ и требуют от пользователя освоения интерфейса программного обеспечения, LLMs взаимодействуют через интуитивно понятный чат. Это позволяет сосредоточиться на постановке исследовательских вопросов и интерпретации результатов, а не на финансовых и технических аспектах работы с инструментом. Кроме того, в отличие от специализированных программ, LLMs предоставляет исследователю больше контроля в работе с данными [8].

Традиционная модель предполагает исследователя, использующего инструменты (будь то ручной анализ или использование специализированных программ). Модель, формирующаяся с приходом LLMs, все чаще описывается как взаимное дополнение, симбиоз человеческого интеллекта исследователя и алгоритмических возможностей модели. LLMs берут на себя автоматизацию рутинных и трудоемких задач (транскрипция, первичное кодирование, извлечение цитат, составление кратких резюме), освобождая время и когнитивные ресурсы исследователя для углубленной интерпретации, теоретизирования и критической рефлексии. Более того, интерактивная природа LLMs позволяет исследователю "обсуждать" данные, проверять гипотезы, получать альтернативные интерпретации или формулировки, преодолевая тем самым "академическое одиночество" [8]. LLMs становятся "коллегой-ассистентом", который сохраняет контекстную целостность данных и усиливает, но не заменяет человеческую интерпретацию [12; 7]. Происходит не просто автоматизация работы и быстрое получение выводов, а расширение и усиление аналитических и творческих возможностей исследователя.

Мы придерживаемся позиции, что исследователь остается ключевым агентом в создании данных: тщательно формулирует промпты, критически оценивает валидность всех выходных данных LLM, контекстуализирует результаты, сохраняет авторство итоговой интерпретации и несёт ответственность. LLM – эффективный помощник, но не автономный аналитик.

Динамичное использование LLMs порождает ряд вопросов по методологии и этике. Сокращение времени работы с текстовым материалом может приводить к риску снижения глубины понимания высказываний участников

фокус-групп [12], в том числе из-за того, что не формируется эмпатия [8]. Автоматическое кодирование может быть слишком общим [8]. Часто нейросетевые модели производят галлюцинации, то есть генерируют несуществующую информацию, фантазируют и создают убедительные, но ложные выводы [12]. Кроме того, встают вопросы о том, насколько алгоритмы влияют на самостоятельность мыслительных процессов исследователя, формируют восприятие и задают некий шаблон анализа.

Ключевым инструментом взаимодействия с LLMs является промпт-инженерия – проектирование входных запросов для решения специфических аналитических задач. Промпт представляет собой структурированную инструкцию. Промпт-инженерия включает: формулировку инструкций; выбор формата (one-shot, few-shot, chain-of-thought и др.); подбор примеров и контекста; оптимизацию для конкретной задачи (например, тематический анализ, категоризация, выявление паттернов) [9].

Нейросеть анализирует огромные массивы данных, чтобы находить в них закономерности и на основе этого принимать решения. Следует учитывать, что нейросеть не обладает человеческим мышлением и не способна к осмысленному восприятию текста. Её функционирование основано на моделировании поведения, при котором генерируемые ответы определяются вероятностными связями, а не логическим размышлением. Поскольку результат зависит от характера запроса, предъявляемого системе, ключевым условием для достижения качественного результата является формулировка точного и корректно структурированного запроса.

Современные ученые детально систематизируют процедуру создания промптов, предлагая базовые техники [9; 4] и каталоги паттернов для реализации исследовательских задач [10]. Однако, на данный момент отсутствуют стандартизированные оценки надежности и валидности промптов, что ограничивает воспроизводимость результатов [9]. Необходимым этапом является пилотное тестирование: проверка промптов на подвыборке, анализ ошибок и искажений.

Эффективным решением достижения методической строгости и валидности результатов является использование системного промпта, под которым понимается начальное сообщение, содержащее инструкцию с правилами: роль, стиль, формат вывода ответов и ограничения, которые модель должна соблюдать на протяжении всей сессии [9]. Это существенно снижает риск галлюцинаций, а повторные прогоны дают сопоставимые результаты. В анализе качественных данных системный промпт позволяет исключить смещение, присущее свободной интерпретации LLM, обеспечив привязку выводов к теоретической рамке и дать эмпирическое подтверждение на основе релевантных цитат.

Общими рекомендациями для создания эффективного промпта являются: использование кратких и точных формулировок, уточнение роли и контекста, указание объема (количество слов, абзацев), включение примера для лучшего понимания. После получения ответа от модели необходимо проверить его соответствие целям исследования. В случае необходимости доработать формулировку промпта.

Методология исследования

Цель данного исследования заключается в анализе потенциала применения технологий искусственного интеллекта, в частности больших языковых моделей (LLM), для обработки и интерпретации качественных социологических данных на материале транскриптов фокус-групп. Методологический дизайн исследования основан на сравнительном анализе эффективности традиционного ручного подхода и автоматизированного анализа качественных социологических данных с использованием больших языковых моделей. Исследовательская процедура предполагала независимое выполнение аналитических задач тремя исследователями в рамках экспериментального дизайна. Первый исследователь осуществлял традиционный качественный анализ транскриптов фокус-групп с применением методов интерпретативной социологии. Второй проводил автоматизированный анализ идентичного эмпирического материала посредством большой языковой модели DeepSeek на основе структурированного исследовательского запроса. Третий выполнял функции

независимого эксперта, осуществляя сравнительный анализ результатов двух подходов с целью выявления различий в интерпретации данных и оценки особенностей каждого из применяемых аналитических инструментов.

Эмпирической базой исследования стали материалы фокус-групп, собранные в ходе научно-исследовательской экспедиции в Пермский край в 2022 г. Экспедиция была направлена на изучение результатов совместного проектирования городской среды, выполненного взрослыми (архитекторами, кураторами и родителями) и подростками – участниками проекта "Лаборатория архитектурной мысли", реализуемого НКО "Улица детства". Первый этап включал лекции и конкурс рисунков среди школьников. По его итогам отбирались участники, которых приглашали на следующий этап – командную работу по созданию макетов будущих проектов. По завершении воркшопов проходил конкурс, в ходе которого выбирался проект для дальнейшей реализации, в которой подростки могли принять участие. Например, в г. Чусовой были сделаны гамаки, а в г. Горнозаводск – бетонная лавочка.

Для настоящего исследования были выбраны материалы двух фокус-групп с подростками, ранее участвовавшими в разработке и создании проектов общественных пространств г. Чусовой и г. Горнозаводск Пермского края.

Выбор лишь двух фокус-групп обусловлен исследовательской задачей: сопоставлением особенностей анализа качественных социологических данных, выполненного человеком и LLM. Такой подход требует более глубокого разбора текстов с последующим сравнением интерпретаций, что делает невозможным охват всех доступных материалов в рамках одной статьи. При этом выбранные фокус-группы обеспечивают достаточное тематическое и содержательное разнообразие для постановки и решения задач, связанных с изучением различий в аналитических стратегиях человека и ИИ.

Для обеспечения аналитического единства всем исследователям был предложен единый исследовательский вопрос – *каковы основные трудности в коммуникации взрослых и детей в процессе реализации соучаствующего проектирования городской среды?* В качестве теоретической основы для анализа данных фокус-групп использован участвующий подход социологии детства. Подход базируется на признании детей активными социальными акторами, способными к самостоятельному действию и влиянию на социальные структуры. Соучастие детей реализуется через создание пространства "этической симметрии" между взрослыми и детьми, где дети выступают равноправными участниками исследовательских и социальных процессов [3]. Важным теоретическим вкладом в понимание детского участия является концепция "Лестницы участия" Роджера Харта [5], которая предлагает семиступенчатую модель участия детей от "Я ничего не решаю, ни в чем не участвую" до "Я сам принимаю решения и информирую о них взрослых". Эта модель позволяет различать подлинное участие детей от простого токенизма.

Показатели активного участия, выделенные Р. Олсен стали теоретической рамкой для анализа транскриптов фокус-групп [6]. В системный промпт была вставлена следующая исследовательская модель: информированность (доступ к понятным данным о цели, процессе и своей роли в проекте), автономия (свобода формировать/выражать мнение без давления), признание (ощущение, что взгляды ребёнка услышаны и учтены) и альянс (чувство партнерства и взаимной значимости в совместной работе).

Сравнительный анализ ручного и автоматизированного подходов

В рамках исследования было проведено сопоставление результатов ручного и автоматизированного анализа данных фокус-групп с подростками по следующим 10 пунктам, которые позволяют наглядно выявить особенности каждого подхода.

Стиль изложения текста. Без использования специального промпта LLM выдает информацию в виде конспекта, используя схемы, пункты, таблицы. В рамках эксперимента DeepSeek предложил следующую классификацию мотивации участников: практическая (бесплатный кофе, возможность пропустить школу), творческая (интерес к работе с бетоном, реализация идей) и социальная (общение, совместный труд), что раскрывает мотивацию более широко. Исследователь придерживается научного стиля изложения (использование структуры и терминологии, опора на научные концепции со

ссылками на них), в то время как LLM склонен к воспроизведению, смешению множества разных источников, в т. ч. и нерелевантных для научных текстов, и генерации на их основе собственных суждений, которые видоизменяются при прописывании нового запроса. Структурирование текста языковой модели похоже на первичную кодировку собранных данных при ручной обработке результатов исследования. В обоих случаях в тексте выделяются ключевые слова и темы.

Цитирование. ИИ подвержен семантическим искажениям при цитировании, в том числе фабрикации элементов, отсутствующих в исходном тексте. К примеру, добавлена не существовавшая в оригинале оценка: "это было не актуально". Произведено новое перефразирование, где исходное высказывание "Единственное, что изменилось, – нас начали замечать. Нас поздравили с днём строителя" редуцировано до "Самое важное – нас поздравили с Днём строителя!", что исказило смысл, исключив ключевой контекст "начали замечать".

Использование теоретической рамки. И ручной, и автоматизированный анализ строились на основе теории участия Р. Олсен [6], с использованием четырех категорий подросткового участия: информация, автономия, признание и альянс.

Исследователь даёт развернутые определения основных понятий, а DeepSeek использует более сжатую форму на основе промпта. В категории "информация" исследователь акцентирует внимание на проблемах информирования и возможном принуждении к участию, а LLM подчёркивает неясность критериев отбора, реализации и последующей судьбы проектов.

В блоке "автономия" ИИ трактует вмешательство кураторов как ограничение творческой свободы и указывает на принудительный характер групповой работы. Исследователь признаёт ограниченность автономии на этапе реализации (вследствие материальных и бюрократических факторов), но отмечает учёт мнения подростков и смешанный стиль общения взрослых (директивный и демократический).

В категории "признание" исследователь выделяет ограниченное влияние подростков на важные решения, несмотря на активное вовлечение на ранних этапах. DeepSeek фиксирует игнорирование мнения при выборе локаций и отсутствие связи с результатом.

В категории "альянс" ИИ указывает на низкую степень соучастия, отсутствие совместного целеполагания и разрушение доверия, однако последний аргумент основан на искажении, что свидетельствует о рисках генерации неверных выводов. Исследователь же рассматривает альянс как следствие взаимодействия со взрослыми, указывая на рост интереса подростков к благоустройству городской среды, несмотря на ощущение недостаточного учёта их мнения.

Сравнение демонстрирует, что ручной и машинный анализ по-разному акцентируют внимание на аспектах участия, и их сочетание может углубить интерпретацию полученных данных.

Ошибки и искажения данных. Человек склонен допускать ошибки ввода – например, пропуски или перестановку данных, тогда как LLM может искажать информацию. Так, модель DeepSeek указала на слабую координацию со школами на основе одного высказывания, не учитывая, что другие сообщали об осведомленности учителей и наличии подтверждающих документов.

Кроме того, модель иногда самостоятельно "достраивает" информацию, которая может быть важна для анализа: так, она сгенерировала возраст участников, хотя такие сведения отсутствовали в транскриптах.

При этом ручной анализ также подвержен интерпретационным ошибкам. В одном из выводов исследователь указал на негативную оценку подростками связи проекта с историей Пермского края. В качестве аргумента была приведена цитата: *"Мы должны были это все связать с историей Пермского края. Вот с этим всем. Поэтому все проекты, изделия должны были быть из бетона и связаны с бетоном, потому что в городе бетонный завод"*, в которой не прослеживается однозначное негативное отношение к данной концепции.

Пропуск значимых деталей. Упущение ключевых аспектов может нарушать логическую целостность и повлиять на достоверность выводов. Например, при сравнении двух фокус-групп модель DeepSeek заключает: "В Чусовом выше вовлеченность учителей и разнообразие проектов, но слабее связь с реализованным объектом". Однако в этом выводе не учитывается разное количество участников (Чусовой – 11, Горнозаводск – 4), а также различие в фокусе обсуждений.

Также при машинном анализе не применялась концепция "Лестница участия" [5], из-за отсутствия её упоминания в промпте. Исследователь, напротив, использует её как диагностический инструмент для оценки субъективного восприятия подростками своей роли во взаимодействии со взрослыми.

При этом в анализе исследователя отсутствует аспект, связанный с готовностью рекомендовать проект друзьям, тогда как ИИ раскрывает эту тему с нескольких точек зрения: 1. Ироничное подтверждение ценности полученного опыта. 2. Идея: "Процесс важнее результата: коллективный опыт как самоцель".

Создание несуществующих взаимосвязей. Модель DeepSeek интерпретирует высказывания подростков как выражение фрустрации из-за неполной реализации проекта, утверждая, что участники отказались доделывать гамаки. Однако в действительности подростки отмечали ограниченность по времени: "*У нас было мало времени на реализацию гамаков. Мы сделали всего лишь 3, а надо было 6, поэтому строители сами уже все доделывали*". Также модель указала на логистические трудности: "участники ходили пешком из-за пропущенных автобусов", тогда как подростки, напротив, сообщают, что пеший маршрут был осознанным и удобным выбором.

Восприятие и интерпретация контекста. Подростки отмечали трудности в коммуникации с кураторами, что исследователь трактует как ограничение возможностей выражения мнения и проявление директивного стиля общения, а DeepSeek – как противоречивость указаний от кураторов. ИИ интерполирует эмоциональный контекст, приписывая участникам эмоции обиды и бессилия, чего не подтверждают данные, ручной анализ, напротив, фиксирует в целом положительный эмоциональный фон: "респонденты дают положительные эмоциональные оценки, связанные с новым опытом, возможностью коммуникации с разными людьми, приятной атмосферой", что подтверждается цитатами из фокус-групп.

Критичность выводов и полнота анализа. В эксперименте исследовательский вопрос касался трудностей коммуникации взрослых и подростков. DeepSeek сфокусировался лишь на проблемах, тогда как ручной анализ учёл и позитивные аспекты (уважительное отношение к детям и обучение новому). Кроме того, LLM прибегает к необоснованной стигматизации, утверждая, что "маленькие группы в депрессивных городах (Горнозаводск) склонны к цинизму; большие группы с опытом грантов (Чусовой) – к рефлексии системных проблем", что не подтверждается данными транскриптов и противоречит принципам научной интерпретации.

Рекомендации и прогнозы. Положительной стороной использования LLM является то, что помимо анализа имеющихся данных она генерирует и прогнозные данные в виде рекомендаций и возможных последствий. Так, при раскрытии основных трудностей в коммуникации модель приводит возможные последствия: снижение веры в эффективность соучастия, восприятие процесса как "взрослой инициативы" без реального влияния, эмоциональное выгорание.

Разница во временных затратах. Одним из основных различий является количество затраченного времени на анализ. Использование языковой модели значительно сокращает время работы, но все равно требует тщательной проверки со стороны исследователя.

Этапы реализации анализа данных фокус-групп с помощью LLM

В нашем исследовании мы решили использовать бесплатную и доступную в России модель. Из возможных вариантов была выбрана DeepSeek. Это одна из мощных современных моделей с функцией reasoning (с англ. – рассуждение), что оптимально для реализации исследовательских задач. Экспериментальные исследования подтверждают, что DeepSeek демонстрирует вы-

сокую точность при тематическом кодировании, анализе длинных контекстов (до 128K токенов), распознавании имплицитных социальных паттернов [3]. По данным платформы OpenLM.ai (<https://openlm.ai/chatbot-arena/>) DeepSeek занимает лидирующие позиции. Кроме того, предварительное тестирование нескольких моделей (Qwen, Perplexity, DeepSeek, Gigachat, YandexGPT) с помощью специального инструмента Chatbot Arena показало, DeepSeek хорошо справляется с качественным анализом транскриптов фокус-групп.

В рамках настоящего исследования применены две методики промпт-инженерии:

1. Chain-of-Thought (CoT): Пошаговая декомпозиция рассуждений, усиливающая интерпретируемость выводов.

2. Role Prompting: Назначение LLM специализированной роли (например, "эксперт по социологии") для активации доменно-специфических знаний.

Важно корректно подготовить транскрипты: обеспечить конфиденциальность данных и присвоить участникам соответствующий код, обозначить возраст и пол, сохранять реплики с указанием кода участника. Желательно структурировать по тематическим разделам.

Использовалась простая и короткая цепочка действий без дополнительных промптов на выявление специфики по отдельным темам (в реальном социологическом исследовании необходимо более глубоко и детально изучать). Ниже представлен план.

1. В диалоговом окне прикреплен полный текстовый файл с транскриптом фокус-группы в г. Горнозаводск и написан промпт:

/промпт
Изучи данный файл

2. В диалоговом окне прикреплен полный текстовый файл с транскриптом фокус-группы в г. Чусовой и написан промпт:

/промпт
Изучи данный файл

3. Сформулирован системный промпт, интегрирующий исследовательский вопрос и теоретическую модель, включающий следующие блоки:

- присвоение роли профессионального социолога;
- определение задачи и указание следовать исследовательской модели;
- механизм валидации через цитирование первичных данных и обозначение формата написания цитат;
- формулировка вывода на основе исследовательского вопроса;
- формат вывода итогового ответа.

/промпт
Ты – профессиональный социолог и пишешь научную статью. На основе глубокого анализа транскриптов обеих фокус-групп [ФГ Чусовой, ФГ Горнозаводск] выдели и перечисли ключевые темы, которые часто упоминаются или кажутся значимыми в обсуждении в соответствии с исследовательской моделью:
["Информация". Доступ к понятным данным о цели, процессе и своей роли. Объяснение контекста (напр., "Зачем мы встречаемся?"), визуализация этапов процесса.
"Автономия". Свобода формировать/выражать мнение без давления. Открытые вопросы, время на размышление, отказ от корректировки высказываний ребенка.
"Признание". Ощущение, что взгляды ребенка услышаны и учтены. Активное слушание, вербализация понимания ("Я слышу, ты хочешь...").
"Альянс". Чувство партнерства и взаимной значимости в совместной работе. Совместная постановка целей (напр., "Давай решим эту проблему вместе")]
1. **Структурируй выводы по каждому принципу модели: **
Начни с четкого обозначения принципа (например, **Информация **).
Под принципом **перечисли ключевые выявленные темы (проблемы)**, связанные с его нарушением или неполной реализацией в коммуникации взрослых и детей. **Сформулируй каждую тему полно и содержательно, отражая суть проблемы в контексте соучаствующего проектирования.** Избегай общих фраз, делай акцент на конкретных аспектах взаимодействия.

****После каждой темы**** приведи ****2-3 полные, релевантные цитаты из транскриптов**** (с указанием кода участника, например `Д9`), которые ****наглядно иллюстрируют**** именно эту конкретную тему. Убедись, что цитата напрямую подтверждает сформулированную тему.

Цитату оформляй так – вначале текст с высказыванием участника в кавычках, затем в круглых скобках код участника и город, по следующему примеру "Когда нас объединяли в группы по рисункам, мы не понимали, что там общего" (Д8, Чусовой)

Укажи ***контекст*** (когда, на каком этапе это происходило, если это ясно из данных).

2. ****Ответ на исследовательский вопрос:****

На основе структурированного анализа по пункту 1, ****сформулируй итоговый ответ**** на вопрос [Основные трудности в коммуникации взрослых и детей в процессе реализации соучаствующего проектирования городских проектов].

Ответ должен ****обобщать ключевые проблемы****, выявленные по всем четырем принципам модели, подчёркивая их влияние на процесс соучастия. Свяжи проблемы с нарушением конкретных принципов модели. Кратко укажи последствия этих трудностей (например, снижение мотивации, ощущение декоративности участия, разочарование).

Примечание: с целью лучшего понимания моделью заданий промпта, используются специальные обозначения.

Материал для анализа включается в квадратные скобки.

* – одна звёздочка обозначает акцент, модель обращает повышенное внимание, однако такие фразы не считаются жёсткими инструкциями. Например: уточнения, нюансы, примечания; обозначение переменных-заполнителей внутри шаблона; выделение примеров, чтобы отделить их от основной инструкции.

** – две звёздочки – директива, то есть модель интерпретирует выделенный фрагмент как ключевую директиву / термин, к которому нужно строго "прислушаться".

4. В диалоговое окно отправлен промпт с заданием сформулировать неочевидный инсайт:

/промпт

На основе анализа обсуждения темы [Основные трудности в коммуникации взрослых и детей в процессе реализации соучаствующего проектирования городских проектов] в фокус-группах [ФГ Горнозаводск, ФГ Чусовой], сформулируй один неочевидный, но потенциально важный инсайт о глубинных мотивах или барьерах участников, связанных с этой темой. Подкрепи инсайт конкретными примерами цитат из разных групп.

5. Промпт на сравнительный анализ между двумя фокус-группами.

/промпт

Сравни две фокус-группы [ФГ Горнозаводск и ФГ Чусовой] по обсуждению ключевой темы [Основные трудности в коммуникации взрослых и детей в процессе реализации соучаствующего проектирования городских проектов]. Выдели основные сходства и различия в восприятии, аргументации, эмоциональной окраске. Укажи, какие социодемографические или контекстуальные факторы (если они тебе известны: [указать, например, возрастной состав, регион]) могли повлиять на эти различия?

Системный промпт для формулирования итогового заключения по всему анализу. Здесь мы не будем его приводить, так как он большой и занимает большое количество знаков. Но логика соответствует системному промпту на шаге 3.

В качестве практических комментариев следует добавить, что первые промпты на шагах 1 и 2 написаны в максимальной простой форме с целью получения более качественной информации на шаге 3. На начальном этапе были протестированы различные варианты промптов, включая запросы на выявление ключевых идей, консенсуса и разногласий между участниками, а также определение эмоциональной окраски дискуссии. Однако такие инструкции оказались ограничивающими, что привело к уменьшению глубины и качества выводов при использовании системного промпта на шаге 3.

После серии итеративных экспериментов был выбран более гибкий подход: предоставление модели возможности автономного изучения текста без детализированных указаний исследователя. Этот метод позволил расширить спектр интерпретаций материала, сделав их более содержательными и релевантными для целей качественного анализа. Полученные результаты демон-

стрируют значительное улучшение как в объёме выявленных тем, так и в их глубине.

Особое внимание уделялось разработке эффективной стратегии создания промптов. Основной вывод заключается в том, что наиболее продуктивным решением является делегирование задачи по формированию промпта самой модели. Исследователь может задать свои цели в общем виде, после чего модель способна предложить структурированный запрос или улучшить его на основе обратной связи. Такой подход позволяет минимизировать требования к исследователю в части владения навыками создания промптов и обеспечивает возможность получения содержательных результатов даже на начальных этапах работы с большими языковыми моделями (LLM).

В части повышения методологической точности перспективно применение технологии Retrieval-Augmented Generation (RAG) в анализе качественных данных. Но так как в нашу задачу входил поиск наиболее простого и доступного способа работы с современными нейросетевыми моделями для начинающего пользователя, данный вариант не использован.

Концептуальная архитектура RAG-систем включает два основных компонента: модуль поиска релевантной информации и модуль генерации ответов на основе найденного контекста [9]. То есть мы создаём базу с документами (наша теория и транскрипты) и работаем непосредственно с ней. В контексте анализа транскриптов фокус-групп данная архитектура обеспечивает систематический подход к извлечению и интерпретации качественных данных, минимизируя риск галлюцинаций и повышая обоснованность исследовательских выводов.

Таким образом, комбинирование гибкого подхода к формулировке запросов с использованием RAG представляет собой перспективную методологию для качественного анализа данных фокус-групп, позволяя достигать более точных и информативных выводов.

Заключение

Возможности применения ИИ в анализе качественных данных связаны, прежде всего, со значительной экономией времени и ресурсов из-за автоматизации рутинных задач. Как и профессиональные программы, в короткие сроки мультифункциональный ИИ способен проводить все необходимые процедуры, связанные с кодированием и аналитикой качественных данных, но этот процесс направляет исследователь.

Одной из актуальных проблем при использовании ИИ в качественных исследованиях является его неспособность полноценно обрабатывать невербальные данные, что наиболее важно при анализе фокус-групп, так как метод связан с групповой динамикой. Эти аспекты часто имеют критическое значение для последующих интерпретаций. Важна и методологическая позиция исследователя, который определяет подход к данным и их анализу, пока технология не может заменить глубину понимания и контекстуальные особенности.

Сложности связаны и с особенностями интерпретации. Системы ИИ делают ошибочные выводы из-за неполной информации и изначально заложенных представлений о социально-демографических характеристиках разных социальных групп. Присутствуют и культурные различия, так как все нейросети отражают специфические стереотипы и особенности, в том числе связанные с представлениями о гендерном порядке.

К общим проблемам использования современных ИКТ можно отнести как возрастающую зависимость от технологий, так и безопасность, и приватность данных, ведь системы могут стать объектом хакерских атак, а несанкционированный доступ к большим объемам личной информации иметь серьезные последствия для пользователей.

Проведённое исследование показало, что результаты анализа с использованием больших языковых моделей в значительной степени зависят от компетенций исследователя, работающего с ними. Ключевыми условиями эффективного применения ИИ являются точность формулировки задач, выбор аналитического ракурса и построение чётких методологических установок.

Ручной и машинный анализ акцентируют внимание на разных аспектах изучаемых данных, совместное использование которых может дать более

глубокие интерпретации. LLM может выступать в роли помощника исследователя, способного снизить когнитивную и временную нагрузку при выполнении рутинных задач – таких как первичная кодировка, составление кратких резюме, тематическое структурирование данных и генерация альтернативных точек зрения.

Тем не менее, на данном этапе развития технологии языковые модели не способны заменить человека в задачах глубокого анализа качественных данных. Они не обеспечивают достаточную связность, логичность и стилистическую точность при формировании научных текстов. Важно учитывать, что модели подвержены генерации недостоверной информации и "галлюцинациям", а также могут выдавать непроверенные или фиктивные ссылки, что требует тщательной редакционной и концептуальной доработки со стороны исследователя.

Также следует учитывать этические риски использования LLM в гуманитарных исследованиях. Автоматическая интерпретация данных, особенно собранных у уязвимых групп (в данном случае – подростков), может привести к искажению смыслов, неоправданной стигматизации или ошибочным выводам. Это требует ответственного подхода к внедрению ИИ и чёткой фиксации границ его аналитических возможностей.

Использование ИИ предоставляет множество преимуществ, но несёт в себе и риски, которые требуют внимания со стороны профессионального научного сообщества. Разработка и соблюдение этических стандартов, обучение новым алгоритмам взаимодействия с ИИ, позволяют исследователю более качественно и эффективно решать поставленные задачи. Опыт расширяется, как и меняется статус исследователя как архитектора, выстраивающего методологию своей работы, но с использованием возможностей ИИ, расширяющих его опции, способного даже к научной дискурсивности, но не решающего творческие задачи и не принимающего итоговые решения.

Литература / References

1. Горбачева Т. А. Искусственный интеллект: риски и проблемы внедрения в Российской Федерации // *Инновационная экономика: информация, аналитика, прогнозы*. 2025. № 1. С. 96–105. Gorbacheva T. A. Artificial intelligence: risks and problems of implementation in the Russian Federation // *Innovative Economy: Information, Analytics, Forecasts*. 2025. No. 1. P. 96–105. (In Russ.).
2. Мокшанов М. В. Применение искусственного интеллекта в анализе данных: обзор текущего состояния и будущих направлений // *Universum: технические науки*. 2024. № 5–1 (122). С. 40–48. Mokshanov M. V. Application of artificial intelligence in data analysis: a review of the current state and future directions. // *Universum: Technical Sciences*. 2024. No. 5-1 (122). P. 40–48. (In Russ.).
3. Хуснутдинова М. Р., Филипова А. Г. "Распаковка" агентности в школьном проекте: перспективы участвующего подхода // *Образование и саморазвитие*. 2024. Т. 19, № 2. С. 139–154. Khusnutdinova M. R., Filipova A. G. "Unpacking" agency in a school project: perspectives of a participatory approach // *Education and Self-Development*. 2024. Vol. 19, No. 2. P. 139–154. (In Russ.).
4. Chen B., Zhang Z., Langrené N., Zhu S. Unleashing the potential of prompt engineering for large language models // *Patterns*. 2025. T. 6. No. 6. P. 101–260.
5. Hart R. *Children's Participation: From Tokenism to Citizenship*. Florence: UNICEF / International Child Development Centre, 1992. 39 p.
6. Olsen R. K., Stenseng F., Kvello O. Key Factors in Facilitating Collaborative Research with Children: A Self-Determination Approach // *Academic Quarter. Akademisk Kvarter*. 2022. No. 24. P. 135–148. <https://doi.org/10.54337/academicquarter.vi24.7256>
7. Schroeder H., Quéré M.A.L., Randazzo C., Mimno D., Schoenebeck S. Large Language Models in Qualitative Research: Uses, Tensions, and Intentions. Large Language Models in Qualitative Research // *arXiv*, 2025. URL: <https://arxiv.org/abs/2410.07362> (accessed 05.07.25).
8. Schroeder H., Randazzo C., Mimno D., Schoenebeck S., Aubin Le Quere M. Large Language Models in Qualitative Research: Can We Do the Data Justice? // *ResearchGate In Extended Abstracts of the CHI Conference on Human Factors in Computing Systems (CHI EA'25)*, April 26–May 1, 2025. Yokohama, Japan. ACM, New York, NY, USA, 2025. <https://doi.org/10.48550/arXiv.2410.07362>.
9. Schulhoff S., Ilie M., Balepur N., Kahadze K., Liu A., Si C., Li Y., Gupta A., Han H., Schulhoff S., Dulepet P. S., Vidyadhara S., Ki D., Agrawal S., Pham C., Kroiz G., Li F., Tao H., Srivastava A., Costa H. D., Gupta S., Rogers M. L., Goncearenco I., Sarli G., Galyunker I., Peskoff D., Carpuat M.,

- White J., Anadkat S., Hoyle A., Resnik P. The Prompt Report: A Systematic Survey of Prompt Engineering Techniques. The Prompt Report / arXiv:2406.06608 [cs]. – arXiv, 2025.
10. White J., Fu Q., Hays S., Sandborn M., Olea C., Gilbert H., Elnashar A., Spencer-Smith J., Schmidt D. C. A Prompt Pattern Catalog to Enhance Prompt Engineering with ChatGPT / arXiv:2302.11382 [cs]. – arXiv, 2023. // arXiv, 2023. URL: <https://arxiv.org/abs/2302.11382> (accessed 05.07.25).
 11. Wong L. Z., Bhattacharya P., Loh S. B., Oh H. S., Juraimi S. A., Anant N., Pillay A., Chong M. F.-F., Fogel A., Sheen F., Pink A. E. Utilizing Large Language Models to Conduct Thematic Analysis: A Case Study on Focus Group Transcripts. Utilizing Large Language Models to Conduct Thematic Analysis. Rochester, NY: Social Science Research Network, 2025. 56 p.
 12. Zhang H., Wu C., Xie J., Lyu Y., Cai J., Carroll J. M. Redefining Qualitative Analysis in the AI Era: Utilizing ChatGPT for Efficient Thematic Analysis // arXiv, 2025. <https://arxiv.org/html/2309.10771v3> (accessed 05.07.25).



Информация об авторах

Маргарита Рафаильевна Хуснутдинова, канд. соц. наук, старший научный сотрудник Московского психолого-педагогического университета, Москва, Россия, e-mail: hushnutdinovaMR@mgppu.ru

Ирина Игоревна Баранова, независимый исследователь, Санкт-Петербург, Россия, e-mail: Baranovairinaa99@gmail.com

Оксана Геннадьевна Зубова, канд. полит. наук, доцент кафедры социальных технологий социологического факультета Московского государственного университета имени М.В. Ломоносова, Москва, Россия, e-mail: zubovaoksana@bk.ru

Information about the authors

Margarita R. Khusnutdinova, Candidate of Sociology, Senior Researcher, Center for Interdisciplinary Research on Contemporary Childhood, Moscow State University of Psychology & Education, Moscow, Russia, e-mail: hushnutdinovaMR@mgppu.ru

Irina I. Baranova, Independent Researcher, Saint Petersburg, Russia, e-mail: Baranovairinaa99@gmail.com

Oksana G. Zubova, Candidate of Political Sciences, Associate Professor, Department of Social Technologies, Faculty of Sociology, Lomonosov Moscow State University, Moscow, Russia, e-mail: zubovaoksana@bk.ru

Поступила в редакцию 16.07.2025

Одобрена после рецензирования 12.08.2025

Принята к публикации 25.08.2025

Received 16.07.2025

Approved 12.08.2025

Accepted 25.08.2025