

Санкт-Петербургский политехнический университет
Петра Великого
Физико-механический институт
Высшая школа прикладной математики и вычислительной физики

Отчет по лабораторной работе №7
по дисциплине
"Автоматизация научных исследований"

Поиск научной литературы

Выполнил студент:

Липс Екатерина Константиновна
группа: 5040102/50201

Преподаватель:

Новиков Федор Александрович

Санкт-Петербург
2025

Содержание

1	Задание	2
2	Промпты	2
2.1	Промпт 1:	2
2.2	Промпт 2:	2
2.3	Промпт 3:	3
2.4	Промпт 4:	3
2.5	Авторский промпт	3
3	Результаты	5
3.1	GPT5	5
3.1.1	Результаты генерации по промπτу 1	5
3.1.2	Результаты генерации по промπτу 2	6
3.1.3	Результаты генерации по промπτу 3	7
3.1.4	Результаты генерации по промπτу 4	8
3.1.5	Результаты генерации по промπτу 5	8
3.2	DeepSeek	9
3.2.1	Результаты генерации по промπτу 1	9
3.2.2	Результаты генерации по промπτу 2	10
3.2.3	Результаты генерации по промπτу 3	10
3.2.4	Результаты генерации по промπτу 4	11
4	Сравнение	12
4.1	Сравнительная таблица оценок	12
4.2	Анализ результатов	12
4.3	Анализ слабых мест моделей	13
5	Выводы	14

1 Задание

Цель работы

Сформулировать запросы к системе поиска научной литературы для нахождения релевантных работ, относительно приведенного научного текста. Для поиска использовать arXiv (<https://arxiv.org/search/>).

Задача:

1. Подготовка исходной статьи (желательно по теме НИР/ВКР).
2. Выбор моделей. Выбрать две различные ИИ-модели для проведения эксперимента.
3. Эксперимент. Отправить текст статьи моделям, используя каждый из четырех предложенных промптов.
4. Анализ результатов. Провести детальный сравнительный анализ 8 полученных ответов, выявляя сильные и слабые стороны различных моделей при работе с различными промптами.
5. Оптимизация. Выбрать одну ИИ-модель, которая, по вашему мнению, показала наилучший потенциал, и на основе выявленных недостатков существующих промптов, составить собственный промпт, который, по вашему мнению, должен дать наилучший результат по всем критериям.
6. Финальная проверка. Прогнать статью через выбранную модель, используя свой авторский промпт.

Критерии оценки:

- Достоверность и проверяемость источников
- Релевантность теме
- Покрытие подтем (полнота)
- Разнообразие без дублей
- Воспроизводимость и прозрачность (можно ли повторить ваш результат и проверить путь получения)

2 Промпты

2.1 Промпт 1:

Сгенерируй 12 простых поисковых запросов для сайта arXiv (<https://arxiv.org/search/>), которые точно дают выдачу. Требования:

- Английский язык, 2–4 общих ключевых слова без кавычек, без двоеточий и логических операторов.
- Не используй редкие аббревиатуры и длинные составные фразы.
- В каждом запросе должен быть хотя бы один «якорный» термин: learning, neural, model, algorithm, method, system, analysis, dataset — выбери подходящие к теме.
- Выведи только готовые строки запросов, по одному на строку.

2.2 Промпт 2:

Подготовь 10 пакетов запросов для сайта arXiv (<https://arxiv.org/search/>). Каждый пакет — два web-safe варианта, которые дают выдачу:

- Basic: 3–6 ключевых слов без кавычек и операторов (только пробелы).
- Soft-phrase: одна короткая фраза в кавычках до 3 слов + 1–2 общих слова (например: “graph neural” models).

Ограничения:

- Только английский; избегай редких узких терминов и длинных фраз.

- Не используй NOT/AND/OR, поля (ti:, abs:, cat:) и двоеточия.
- В каждом варианте добавляй 1 «якорное» слово (learning, neural, model, algorithm, method, system, analysis, dataset).

Формат вывода: для каждого пакета две строки подряд — сначала Basic, затем Soft-phrase.

2.3 Промпт 3:

Сформируй 10 наборов эскалируемых запросов для arXiv по теме. Для каждого набора дай три строки:

- Q0 (web-safe): 3–5 общих ключевых слов без кавычек и операторов — ориентирован на <https://arxiv.org/search> (All fields).
- Q1 (web-safe+): одна короткая фраза в кавычках (до 3 слов) + 1–2 общих слова. Без операторов.
- Q2 (fielded мягкий, для API или Advanced Search): abs:«краткая фраза до 3 слов» OR ti:«краткая фраза до 3 слов» — без NOT; категорию не добавляй.

Правила:

- Только английский; избегай редких аббревиатур (кроме общеизвестных вроде GNN, LLM, NLP).
- Не используй длинные точные фразы.
- Не показывай рассуждения; выведи только тройки строк в порядке Q0, Q1, Q2 для каждого набора.

2.4 Промпт 4:

Выдели 6 фасетов (подтем) по тексту ниже (каждый 2–3 слова), НО не выводи их отдельно. Для каждого фасета сгенерируй по 3 запроса, предназначенные для arXiv и дающие выдачу:

- Minimal (web-safe): 2–4 общих ключевых слова, без кавычек и операторов.
- Broadened (web-safe): добавь 1–2 синонима через пробел (без OR), оставь якорные слова (learning, neural, model, method, system).
- Review-oriented (web-safe): включи слова survey или review вместе с темой, без кавычек и операторов.

Ограничения:

- Только английский. Запрещены NOT/AND/OR, поля (ti:, abs:, cat:), двоеточия и длинные точные фразы.
- Избегай узких редких терминов; предпочитай общепотребимые.

Формат вывода:

- 18 строк (по 3 на каждый из 6 фасетов) в группах по 3 строки подряд: Minimal, затем Broadened, затем Review-oriented.

2.5 Авторский промпт

Задача

Сформировать воспроизводимый, релевантный и полно покрывающий набор поисковых запросов для arXiv на основе научного текста ниже, минимизируя избыточную обобщённость и чрезмерное сужение поиска, при сохранении структурного покрытия темы.

Инструкция модели

На основе приведённого научного текста:

1. **Не выводя их отдельно**, внутренне выдели 5–6 фасетов (подтем), отражающих:
 - объект исследования,
 - используемые методы,

- тип данных,
- пространственные метрики,
- прикладной (биомедицинский) контекст.

Фасеты в ответе не показывай.

2. Для **каждого фасета** сгенерируй **4 взаимосвязанных поисковых запроса**, образующих логическую цепочку от общего к точному:

- **Q0 — Broad (web-safe):**
 - 3–5 общих ключевых слов,
 - без кавычек, операторов и полей,
 - обязательно содержит одно якорное слово:
learning, neural, model, method, system, analysis, dataset.
- **Q1 — Focused phrase (web-safe):**
 - одна короткая фраза в кавычках (2–3 слова),
 - + 1–2 общих слова,
 - без операторов.
- **Q2 — Method-aware (web-safe):**
 - 3–5 слов,
 - допускается одно общеизвестное название метода или метрики, если оно явно присутствует в тексте
(например: DBSCAN, Moran index, nearest neighbor),
 - без кавычек и операторов.
- **Q3 — Review-oriented (web-safe):**
 - 3–5 слов,
 - обязательно включает *review* или *survey*,
 - сохраняет связь с тем же фасетом.

3. **Формат вывода (строгий):**

- для каждого фасета — ровно 4 строки подряд: Q0, Q1, Q2, Q3;
- всего 20–24 строки;
- без заголовков, нумерации, пояснений и комментариев.

Ограничения

- Только английский язык.
- Запрещены: NOT / AND / OR, поля (ti:, abs:, cat:), двоеточия.
- Избегай редких узких терминов, если они не являются центральными в тексте.
- Каждый запрос должен давать выдачу на arXiv.
- Не повторяй запросы и не используй семантические дубликаты.

Цель

Получить структурированный, масштабируемый и проверяемый набор поисковых запросов, пригодный как для быстрого обзора литературы, так и для целенаправленного углублённого поиска по методам и пространственным метрикам.

Обоснование

Разработанный авторский промпт был сформирован на основе анализа сильных и слабых сторон ранее использованных промптов P1–P4 и направлен на системное устранение выявленных ограничений при генерации поисковых запросов для arXiv.

Устранение избыточной обобщённости (недостаток P1). Промпт P1 ориентирован на минимальные “web-safe” запросы, что гарантирует выдачу, но приводит к чрезмерно общим формулировкам и снижению тематической релевантности. В авторском промпте эта проблема решается за счёт фасетного разбиения темы и обязательного формирования нескольких запросов для каждого фасета. Это позволяет сохранить общую доступность поиска, не теряя связи с предметной областью исследования.

Баланс между точностью и полнотой (ограничения P2). Промпт P2 повышает точность за счёт коротких фраз в кавычках, однако не обеспечивает системного покрытия темы и может приводить к фрагментарности результатов. В P5 фазированная структура (Q0–Q3) обеспечивает последовательный переход от широких запросов к более уточнённым, что позволяет одновременно сохранять полноту охвата и повышать адресность поиска.

Снижение риска чрезмерного сужения поиска (недостаток P3). Хотя промпт P3 обеспечивает высокую воспроизводимость и управляемость за счёт эскалации запросов, использование точных фраз и методов может приводить к излишнему сужению выдачи. В авторском промпте это компенсируется введением отдельного уровня Q2 (method-aware), где допускается не более одного названия метода или метрики, и сохранением более общего уровня Q0, что предотвращает потерю релевантных работ на ранних этапах поиска.

Сохранение и развитие фасетного подхода (преимущество P4). Промпт P4 продемонстрировал наилучшее покрытие подтем за счёт фасетной структуры и включения обзорных запросов. В P5 данный подход сохранён и расширен: фасеты используются как внутренняя структура генерации, а для каждого фасета формируется полный цикл запросов (от обзорных до методо-ориентированных). Это обеспечивает как горизонтальное (между подтемами), так и вертикальное (по глубине) покрытие литературы.

Повышение воспроизводимости и прозрачности. В отличие от P1 и P2, где результат представляет собой неструктурированный список запросов, авторский промпт задаёт жёсткий формат вывода и фиксированную логику генерации. Это позволяет повторять процедуру поиска, сравнивать результаты между моделями и документировать путь получения библиографии, что является важным критерием научной воспроизводимости.

3 Результаты

3.1 GPT5

3.1.1 Результаты генерации по промпту 1

- neural morphology analysis
- dendritic spine model
- spatial neural analysis
- neural system structure
- learning neural structure
- neural dataset morphology
- spine distribution analysis
- neural pattern method
- neural clustering algorithm
- spatial neural model

- neural network structure
- neural data analysis

3.1.2 Результаты генерации по промπτу 2

- Пакет 1
 - dendritic spine neural analysis
 - "dendritic spine"model
- Пакет 2
 - neural morphology spatial model
 - "neural morphology"analysis
- Пакет 3
 - spatial neural distribution analysis
 - "spatial neural"model
- Пакет 4
 - neural clustering method analysis
 - "neural clustering"method
- Пакет 5
 - neural pattern detection algorithm
 - "pattern detection"neural
- Пакет 6
 - neural data spatial analysis
 - "neural data"analysis
- Пакет 7
 - neural structure learning model
 - "neural structure"learning
- Пакет 8
 - neural system spatial organization
 - "spatial organization"neural
- Пакет 9
 - neural dataset spatial model
 - "neural dataset"analysis
- Пакет 10
 - neural morphology learning analysis
 - "neural morphology"learning

3.1.3 Результаты генерации по промπτу 3

- Набор 1
 - dendritic spine neural analysis
 - "dendritic spine"model
 - abs:"dendritic spine"OR ti:"dendritic spine"
- Набор 2
 - neural morphology spatial model
 - "neural morphology"analysis
 - abs:"neural morphology"OR ti:"neural morphology"
- Набор 3
 - spatial neural distribution analysis
 - "spatial neural"model
 - abs:"spatial neural"OR ti:"spatial neural"
- Набор 4
 - neural clustering spatial method
 - "neural clustering"method
 - abs:"neural clustering"OR ti:"neural clustering"
- Набор 5
 - neural pattern spatial analysis
 - "neural pattern"analysis
 - abs:"neural pattern"OR ti:"neural pattern"
- Набор 6
 - neural data spatial system
 - "neural data"system
 - abs:"neural data"OR ti:"neural data"
- Набор 7
 - neural structure learning model
 - "neural structure"learning
 - abs:"neural structure"OR ti:"neural structure"
- Набор 8
 - neural spatial organization analysis
 - "spatial organization"neural
 - abs:"spatial organization"OR ti:"spatial organization"
- Набор 9
 - neural dataset spatial analysis
 - "neural dataset"analysis
 - abs:"neural dataset"OR ti:"neural dataset"
- Набор 10
 - neural morphology spatial learning
 - "neural morphology"learning
 - abs:"neural morphology"OR ti:"neural morphology"

3.1.4 Результаты генерации по промπτу 4

- Фасет 1
 - dendritic spine neural
 - dendritic spine neural structure model
 - dendritic spine neural review
- Фасет 2
 - neural morphology analysis
 - neural morphology structure system
 - neural morphology survey
- Фасет 3
 - spatial neural organization
 - spatial neural organization structure
 - spatial neural organization review
- Фасет 4
 - neural clustering method
 - neural clustering grouping method
 - neural clustering survey
- Фасет 5
 - neural pattern analysis
 - neural pattern structure system
 - neural pattern review
- Фасет 6
 - neural data distribution
 - neural data spatial distribution model
 - neural data distribution survey

3.1.5 Результаты генерации по промπτу 5

- dendritic spine neural analysis
- "dendritic spine"neural
- dendritic spine nearest neighbor analysis
- dendritic spine neural review
- neural morphology spatial model
- "neural morphology"analysis
- neural morphology PCA analysis
- neural morphology survey
- spatial neural point analysis
- "spatial point"neural
- spatial point Moran index analysis

- spatial point pattern review
- neural clustering spatial method
- "neural clustering" method
- neural clustering DBSCAN method
- neural clustering survey
- neural data spatial dataset
- "neural data" analysis
- neural data Shannon entropy analysis
- neural data spatial review
- neurodegeneration neural system analysis
- "Alzheimer disease" neural
- neurodegeneration dendritic spine analysis
- neurodegeneration neural review

3.2 DeepSeek

3.2.1 Результаты генерации по промпту 1

- neural networks learning analysis
- dendritic spine spatial analysis
- spatial point pattern method
- spatial clustering neural data
- dendritic spine morphology dataset
- point pattern analysis algorithm
- neuron dendritic spines model
- spatial autocorrelation analysis method
- neural morphology clustering system
- spine distribution analysis neural
- DBSCAN clustering algorithm method
- autocorrelation Moran analysis neural

3.2.2 Результаты генерации по промπτу 2

- Пакет 1
 - spatial analysis dendritic spines clustering
 - "spatial point pattern"analysis neural
- Пакет 2
 - dendritic spine morphology clustering method
 - "dendritic spines"clustering algorithm
- Пакет 3
 - DBSCAN clustering neural data
 - "DBSCAN clustering"neural method
- Пакет 4
 - neural networks spatial autocorrelation analysis
 - "Moran I"spatial analysis
- Пакет 5
 - point pattern analysis neural model
 - "nearest neighbor distance"analysis
- Пакет 6
 - dendritic spine distribution dataset
 - "spine distribution"neural system
- Пакет 7
 - neuron morphology spatial algorithm
 - "spatial clustering"neural method
- Пакет 8
 - spatial statistics neural networks
 - "spatial statistics"neural analysis
- Пакет 9
 - image analysis dendritic spines model
 - "biomedical image analysis"neural

3.2.3 Результаты генерации по промπτу 3

- Набор 1
 - spatial point pattern analysis
 - "spatial point pattern"clustering
 - abs:"spatial point pattern"OR ti:"spatial point pattern"
- Набор 2
 - dendritic spine clustering
 - "dendritic spine"clustering algorithm
 - abs:"dendritic spine"OR ti:"dendritic spine"
- Набор 3
 - neural morphology analysis

- "neural morphology"spatial analysis
- abs:"neural morphology"OR ti:"neural morphology"
- Набор 4
 - spatial autocorrelation neural
 - "spatial autocorrelation"Moran analysis
 - abs:"spatial autocorrelation"OR ti:"spatial autocorrelation"
- Набор 5
 - DBSCAN clustering neural
 - "DBSCAN clustering"neural networks
 - abs:"DBSCAN clustering"OR ti:"DBSCAN clustering"
- Набор 6
 - point pattern statistics
 - "point pattern"neural analysis
 - abs:"point pattern"OR ti:"point pattern"
- Набор 7
 - spine distribution neural
 - "dendritic spines"distribution
 - abs:"dendritic spines"OR ti:"dendritic spines"
- Набор 8
 - neural networks spatial analysis
 - "neural networks"spatial statistics
 - abs:"neural networks"OR ti:"neural networks"
- Набор 9
 - spatial clustering method
 - "spatial clustering"neural data
 - abs:"spatial clustering"OR ti:"spatial clustering"
- Набор 10
 - Moran index analysis
 - "Moran index"spatial autocorrelation
 - abs:"Moran index"OR ti:"Moran index"

3.2.4 Результаты генерации по промпту 4

- Фасет 1
 - spatial point pattern
 - spatial point pattern analysis neural
 - survey spatial point pattern analysis
- Фасет 2
 - dendritic spine clustering
 - dendritic spine morphology clustering method
 - review dendritic spine clustering
- Фасет 3

- spatial autocorrelation neural
- spatial autocorrelation analysis Moran method
- survey spatial autocorrelation neural
- Фасет 4
 - DBSCAN clustering neural
 - DBSCAN clustering algorithm method
 - review DBSCAN clustering neural
- Фасет 5
 - neural morphology analysis
 - neuron dendritic morphology analysis system
 - survey neural morphology analysis
- Фасет 6
 - spine distribution analysis
 - dendritic spine distribution analysis neural
 - review spine distribution analysis

4 Сравнение

Для количественного сравнения качества генерации научных источников моделями GPT и DeepSeek была проведена оценка по нескольким критериям. Оценка проводилась отдельно для каждого промпта, после чего вычислялось среднее значение по источникам.

4.1 Сравнительная таблица оценок

Для корректного сравнения результатов генерации поисковых запросов использовались критерии из задания:

1. достоверность/проверяемость (в контексте arXiv: даёт выдачу и не ломает синтаксис),
2. релевантность теме,
3. покрытие подтем (полнота),
4. разнообразие без дублей,
5. воспроизводимость и прозрачность (насколько легко повторить и получить сопоставимую выдачу).

Оценка выполнена экспертно по 10-балльной шкале на основе полученных списков запросов.

4.2 Анализ результатов

Сравнение показало, что качество итоговых поисковых запросов существенно зависит от структуры промпта: чем сильнее формализованы требования, тем выше воспроизводимость и тем лучше обеспечивается покрытие подтем исследования.

Промпт 1 (короткие “web-safe” запросы) дал наиболее простые и универсальные строки. Это гарантирует выдачу на arXiv, но приводит к недостаточной специфичности: запросы часто смещаются в сторону общих слов (*neural*, *model*, *analysis*), из-за чего релевантность к предметной области статьи (пространственные закономерности на дендрите, метрики NNdist/PCF, Moran’s I, DBSCAN) оказывается умеренной. Для первичной разведки темы промпт подходит хорошо, но для систематического подбора литературы его недостаточно.

Промпт 2 (пакеты Basic + Soft-phrase) улучшил точность за счёт коротких фраз в кавычках, сохранив “web-safe” формат. У обеих моделей выросла адресность запросов: появляются устойчивые связки (*“dendritic spine”*, *“neural morphology”*, *“spatial point pattern”*). При этом существует риск, что

Таблица 1: Сравнительная таблица результатов генерации запросов (модели × промпты)

		Критерии					
Модель	Промпт	1	2	3	4	5	Среднее
GPT5	1	10	6	5	6	10	7.4
	2	10	7	6	7	10	8.0
	3	9	8	7	7	9	8.0
	4	10	8	9	8	10	9.0
	5	10	9	10	9	10	9.6
DeepSeek	1	9	8	6	7	9	7.8
	2	8	8	7	7	8	7.6
	3	9	9	8	7	9	8.4
	4	9	9	9	8	9	9.0

слишком общая фраза в кавычках сузит выдачу не по теме (например, “pattern detection”) или, наоборот, будет избыточно широкой. В целом P2 оказался хорошим компромиссом между полнотой и точностью.

Промпт 3 (эскалация Q0/Q1/Q2) дал наиболее управляемый результат: Q0 обеспечивает широкий охват, Q1 уточняет запрос фразой, а Q2 позволяет перейти к более точному поиску по полям *abs/ti*. На практике это повышает воспроизводимость и прозрачность: можно фиксировать цепочку расширения/уточнения запроса и повторять её. Для DeepSeek P3 оказался особенно удачным из-за более предметных фраз (например, *Moran index*, *nearest neighbor distance*, *spatial autocorrelation*, *DBSCAN clustering*).

Промпт 4 (фасеты + review/survey) дал наилучшее покрытие подтем, поскольку принуждает модель мыслить структурно (несколько независимых направлений поиска) и дополнительно добавляет обзорные запросы (*survey/review*). Это критически важно для быстрого построения базы литературы: обзорные статьи позволяют определить основные подходы и затем углубляться в первичные работы. По совокупности критериев P4 показал наилучший потенциал и для GPT5, и для DeepSeek, особенно по полноте и разнообразию запросов.

Итог: для первичного входа в тему достаточно P1, для систематической работы лучше P3–P4; максимальную практическую пользу для построения библиографии дают фасетные запросы (P4) и эскалируемые схемы (P3).

На основании данного анализа промптов 1-4 был разработан авторский промпт, текст которого приведен в разделе 2.5. Данный промпт был протестирован в комбинации с моделью GPT5 и показал более высокие результаты, нежели промпты 1-4.

4.3 Анализ слабых мест моделей

GPT5. Основное слабое место проявилось в склонности к чрезмерной обобщённости при строгих ограничениях “web-safe” (особенно в P1–P2). В результате запросы становятся корректными по формату и гарантированно дают выдачу, но теряется связь с ключевыми сущностями статьи: *PCF* (*pair correlation function*), *NNdist*, *Moran’s I*, *DBSCAN*, а также контекст биомедицинских данных (дендритные шипики, нейродегенерация, сравнение групп Wt/Ab). Это снижает релевантность и делает последующий ручной отбор статей более трудоёмким. Также наблюдается повторяемость высокоуровневых формулировок (*neural morphology / spatial analysis*), что уменьшает разнообразие.

DeepSeek. Сильной стороной стала более высокая предметность формулировок (чаще появляются названия конкретных методов и метрик), однако слабое место — повышенная вероятность выхода за ограничения промпта или появления терминов, которые потенциально сужают поиск (например, использование *Moran I* как фразы может ограничить выдачу на arXiv, если это реже встречающийся термин в конкретной области). Кроме того, в некоторых случаях модель склонна формировать запро-

сы на грани “редких” (метод/метрика становится центром, а биологический контекст отходит на второй план), что может приводить к выдаче по общей пространственной статистике без фокуса на дендритных шипиках.

Общее ограничение для обеих моделей. Обе модели не могут гарантировать, что каждый запрос приведёт к полностью релевантным статьям без ручной проверки: arXiv-поиск чувствителен к выбору общих слов, и слишком короткие запросы дают шум, а слишком точные — теряют полноту. Поэтому оптимальной стратегией является комбинирование: сначала фасеты/обзоры (P4) для карты темы, затем эскалация (P3) для точного отбора, и только после этого ручная фильтрация результатов.

5 Выводы

В работе была выполнена генерация поисковых запросов для arXiv по научному тексту, посвящённому анализу пространственной организации дендритных шипиков и сравнительному исследованию нормальных (Wt) и патологических (Ab) дендритов с использованием метрик пространственной статистики и кластеризации (NNdist, PCF, энтропия Шеннона, Moran's I, DBSCAN).

В эксперименте были использованы две ИИ-модели (GPT5 и DeepSeek) и четыре базовых промпта, различающиеся степенью формализации требований к запросам. Установлено, что увеличение детализации промпта повышает качество результатов по критериям воспроизводимости, полноты и релевантности. Наиболее слабым оказался P1: он обеспечивает корректный формат и выдачу, но даёт слишком общие запросы и требует существенной ручной доработки. Промпт P2 улучшает точность за счёт коротких фраз в кавычках, сохраняя web-safe ограничения. Промпт P3 обеспечивает наилучшую управляемость поиска благодаря эскалации (Q0/Q1/Q2) и удобен для прозрачного повторения процедуры подбора литературы. Наиболее эффективным из базовых промптов для построения карты темы оказался P4, поскольку фасетный подход и добавление review/survey позволяют быстро получить обзорные работы и одновременно покрыть несколько подтем без существенной потери выдачи.

На основе анализа сильных и слабых сторон промптов P1–P4 был разработан авторский промпт (P5), сочетающий фасетный подход с эскалируемой структурой запросов и контролируемым использованием методологических терминов. Применение данного промпта позволило достичь наилучших результатов по всем критериям оценки.

Сравнение моделей показало, что GPT5 стабильнее соблюдает форматные ограничения и обеспечивает высокую воспроизводимость результатов, тогда как DeepSeek чаще включает методологические сущности статьи (пространственная автокорреляция, nearest neighbor distance, DBSCAN), что повышает предметную релевантность, но может усиливать риск излишнего сужения запроса и требует дополнительного контроля соответствия ограничениям промпта.

Таким образом, было показано, что большие языковые модели являются эффективным инструментом ускорения поиска научной литературы, однако максимальная практическая эффективность достигается при использовании формализованных и структурированных промптов. Разработанный авторский промпт (P5) может рассматриваться как универсальный и воспроизводимый подход к генерации поисковых запросов для arXiv, пригодный для систематического анализа литературы и адаптируемый к различным предметным областям при сохранении обязательной экспертной оценки результатов.