



Министерство науки и высшего образования Российской Федерации
Федеральное государственное автономное образовательное учреждение высшего образования
«Московский государственный технический университет имени Н.Э. Баумана
(национальный исследовательский университет)»
(МГТУ им. Н.Э. Баумана)

Курсовой проект

Программное решение на основе искусственного интеллекта для разбиения сложных вопросов на простые с использованием больших языковых моделей

Студент: Гиричев Марк Сергеевич ИУ7-46Б

Научный руководитель: Строганов Юрий Владимирович

Консультанты: Якуба Дмитрий Васильевич
Москвичёв Николай Владимирович
Левиев Дмитрий Олегович

Цель и задачи работы

Цель работы: Разработка программного решения для разбиения сложных вопросов на простые вопросы с использованием больших языковых моделей.

Задачи:

1. Проанализировать существующие методы декомпозиции вопросов.
2. Разработать алгоритм декомпозиции на основе больших языковых моделей.
3. Реализовать программное решение.
4. Провести исследование эффективности разработанного решения.
5. Проанализировать правовые аспекты программной реализации и выбрать лицензию.

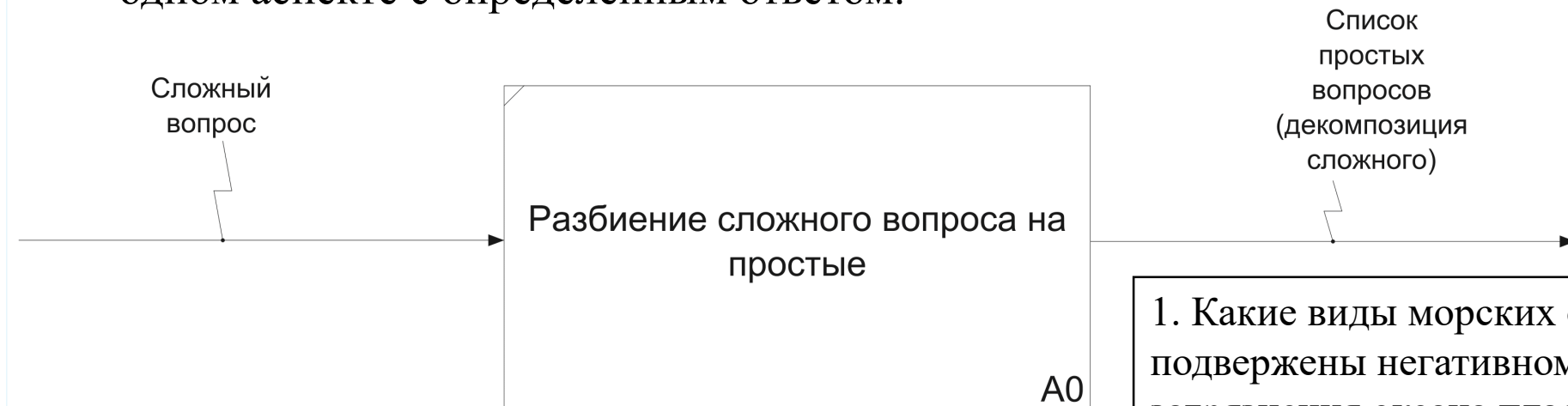
Рейтинг языковых моделей (LLM Arena)

Модель	Общий балл	Доверит. интервал	Ср. кол-во токенов в отв.	Станд. отклон. кол-ва токенов	Оценка языковой компетенц.
DeepSeek-V3-Chat	96.30	+0.8/-0.7	665.97	504.83	56.62
chatgpt-4o-latest	94.74	+0.8/-1.0	693.15	634.20	56.40
o1-mini	93.46	+0.7/-0.8	791.18	647.74	56.22
yi-lightning	93.46	+0.9/-1.0	636.68	469.74	56.22
RuadaptQwen-32B-Pro_v1	92.18	+1.2/-1.2	563.43	387.83	56.04
claude-3-opus-20240229	91.31	+1.0/-1.1	468.69	254.10	55.92
T-Tech-T-pro-it-1.0	90.87	+0.8/-1.1	502.00	380.68	55.85
SberDevices-GigaChatMax	89.96	+1.2/-1.4	523.95	421.87	55.73

Функциональная модель процесса

Сложный вопрос — вопрос, содержащий несколько аспектов и требующий многокомпонентного ответа.

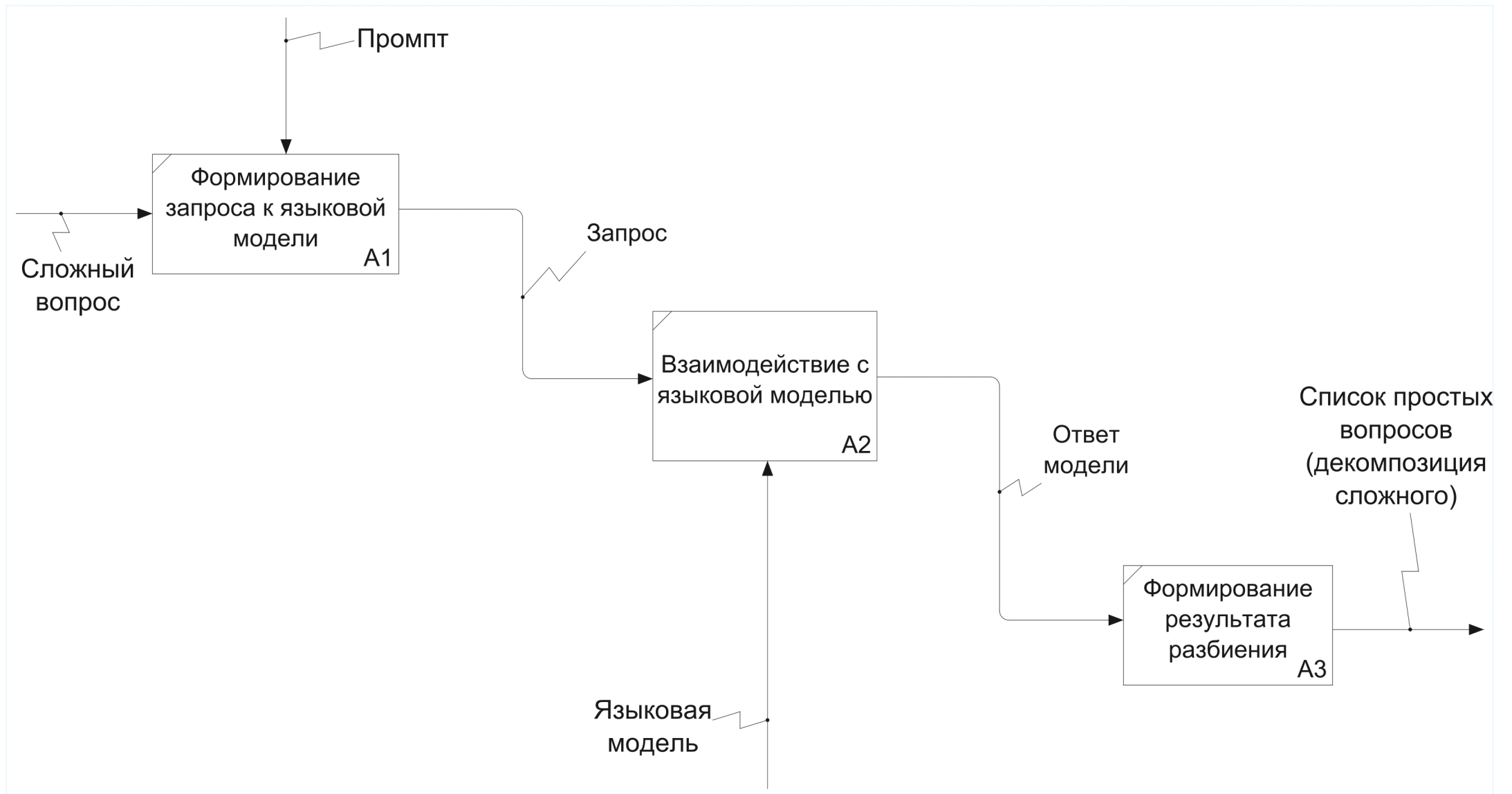
Простой вопрос —самодостаточный атомарный вопрос, фокусирующийся на одном аспекте с определенным ответом.



Пример: Как экологические проблемы, связанные с загрязнением океана пластиком, влияют на морских обитателей, прибрежные экосистемы и здоровье человека, и какие существуют технологические и законодательные решения для минимизации этого воздействия?

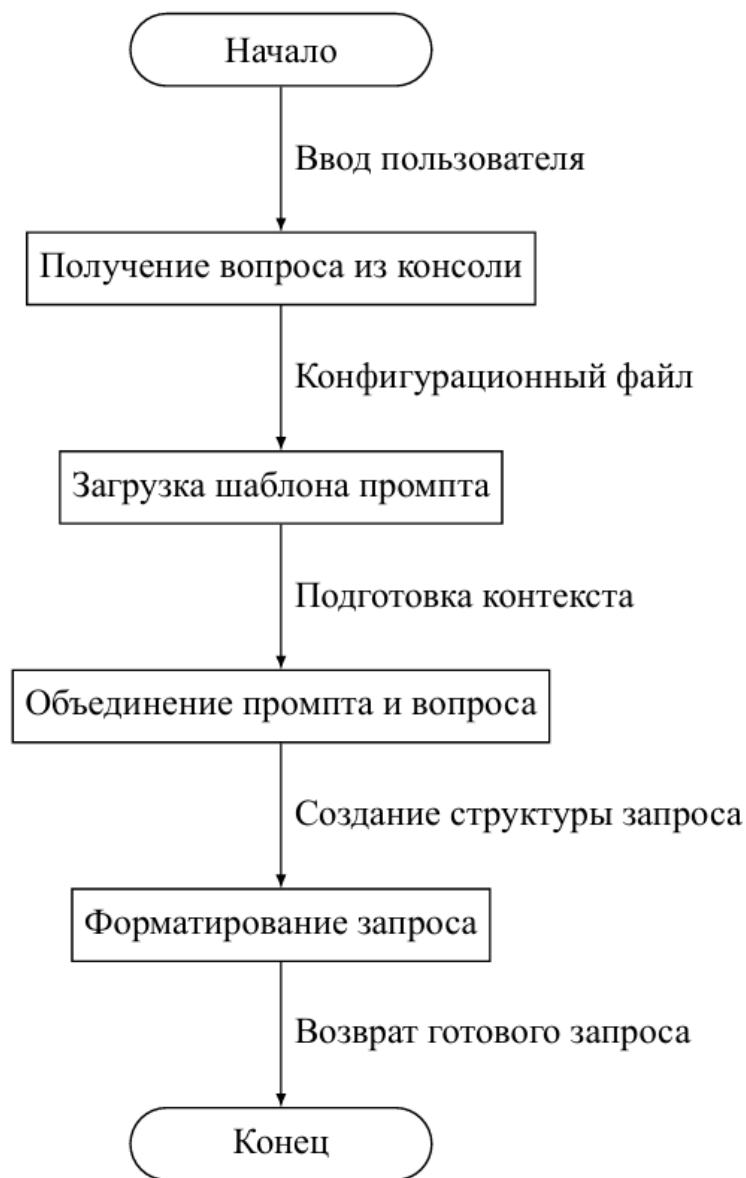
1. Какие виды морских обитателей наиболее подвержены негативному воздействию загрязнения океана пластиком?
 2. Какие конкретные физиологические и поведенческие последствия вызывает загрязнение пластиком у морских обитателей?
 3. Как загрязнение океана пластиком влияет на прибрежные экосистемы?
 4. Какие виды пластикового загрязнения наиболее опасны для прибрежных экосистем?
- и т.д. ещё 14 простых вопросов...

Функциональная модель системы



Алгоритм формирования запроса к БЯМ (А1)

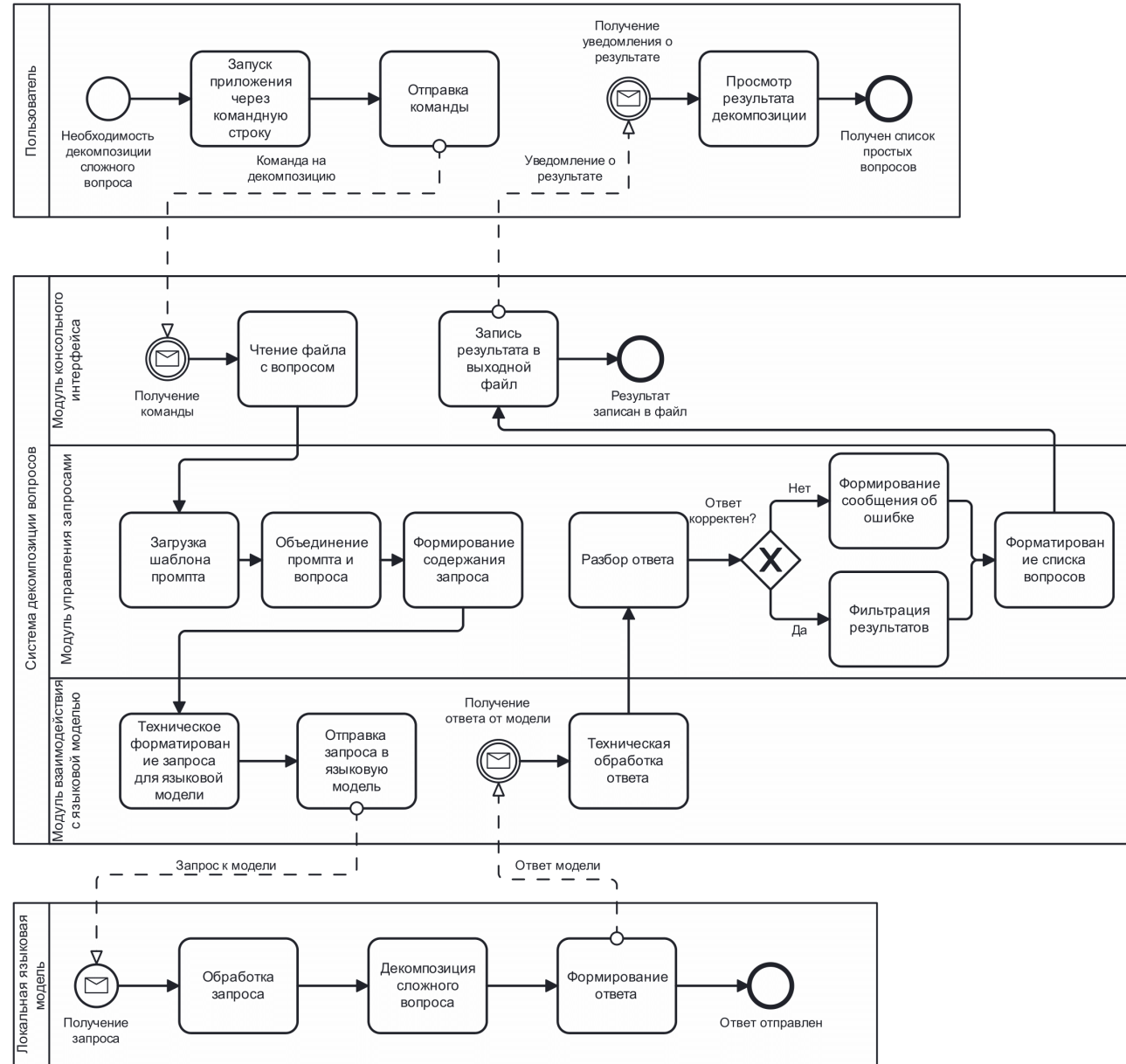
БЯМ – большая
языковая модель



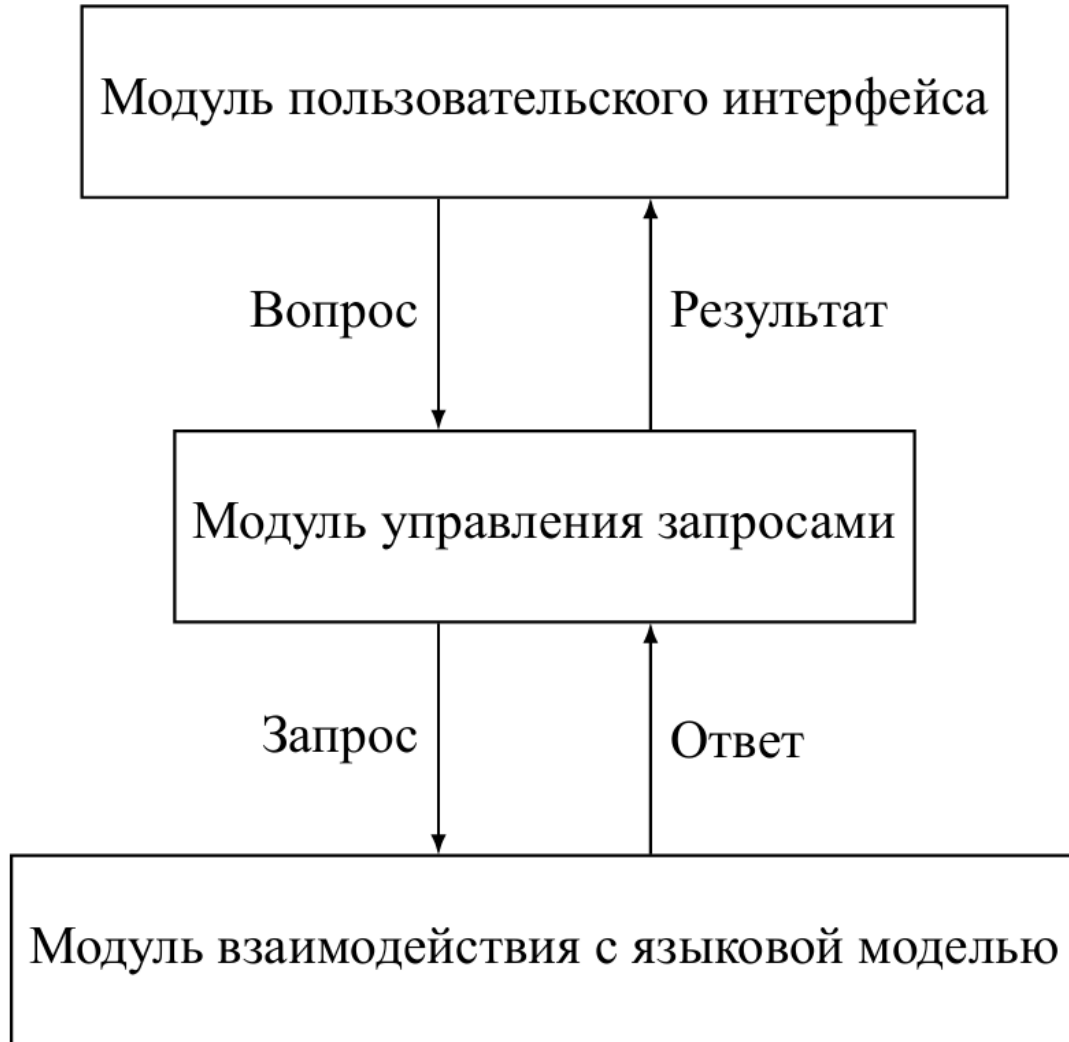
Алгоритм взаимодействия с БЯМ (А2/А3)



VRMN-схема декомпозиции вопроса



Структура программного решения



Используемые библиотеки:

- **llama-cpp-python** – работа с локальными моделями
- **typer** – консольный интерфейс
- **pydantic** – валидация данных
- **rich** - форматированный вывод

Пример работы программы

```
(venv) PS C:\MSTU\kursach\prog\test_code_v1\code> python .\src\cli_module.py
Исходный вопрос: Какие факторы повлияли на развитие письменности в древних цивилизациях и как это изменило политическую, экономическую и социальную структуру общества?
llama_context: n_ctx_per_seq (4096) < n_ctx_train (32768) -- the full capacity of the model will not be utilized
Результат валидации: Прошел
Завершение обработки... 100% 0:00:00
Результат декомпозиции:
1. Какие основные факторы способствовали развитию письменности в древних цивилизациях?
2. Как письменность повлияла на политическую структуру древних обществ?
3. Как письменность изменила экономическую структуру древних цивилизаций?
4. Как письменность повлияла на социальную структуру древних обществ?
5. Какие аспекты политической структуры древних цивилизаций были изменены благодаря письменности?
6. Как письменность способствовала развитию торговли и экономики в древних обществах?
7. Как письменность повлияла на социальную мобильность и статус в древних обществах?
8. Какие конкретные примеры древних цивилизаций иллюстрируют влияние письменности на их политическую, экономическую и социальную структуру?
9. Как письменность способствовала централизации власти в древних государствах?
10. Как письменность повлияла на развитие правовых систем в древних обществах?
11. Как письменность изменила методы ведения учета и бухгалтерии в экономике древних цивилизаций?
12. Какие социальные группы в древних обществах первыми начали использовать письменность и почему?
13. Как письменность повлияла на сохранение культурной и исторической информации в древних цивилизациях?
14. Как письменность способствовала распространению знаний и идей в древних обществах?
15. Какие ограничения или недостатки письменности могли повлиять на её развитие в древних цивилизациях?
Результаты сохранены в файл: results\answers1.txt
```

Результаты экспертной оценки (1/2)

Все критерии в диапазоне [0, 5], где 0 - полное несоответствие критерию, а 5 - полное соответствие.

Полнота: 0 - отсутствие охвата аспектов исходного вопроса. 5 - набор подвопросов полностью покрывает информационную потребность исходного вопроса.

Атомарность: 0 - подвопросы содержат множество разнородных аспектов, требуя дальнейшей декомпозиции. 5 - каждый подвопрос идеально фокусируется на одном смысловом компоненте.

Корректность: 0 - грамматические ошибки, нелогичные формулировки, стилистическая несогласованность. 5 - безупречная грамматическая и логическая целостность подвопросов.

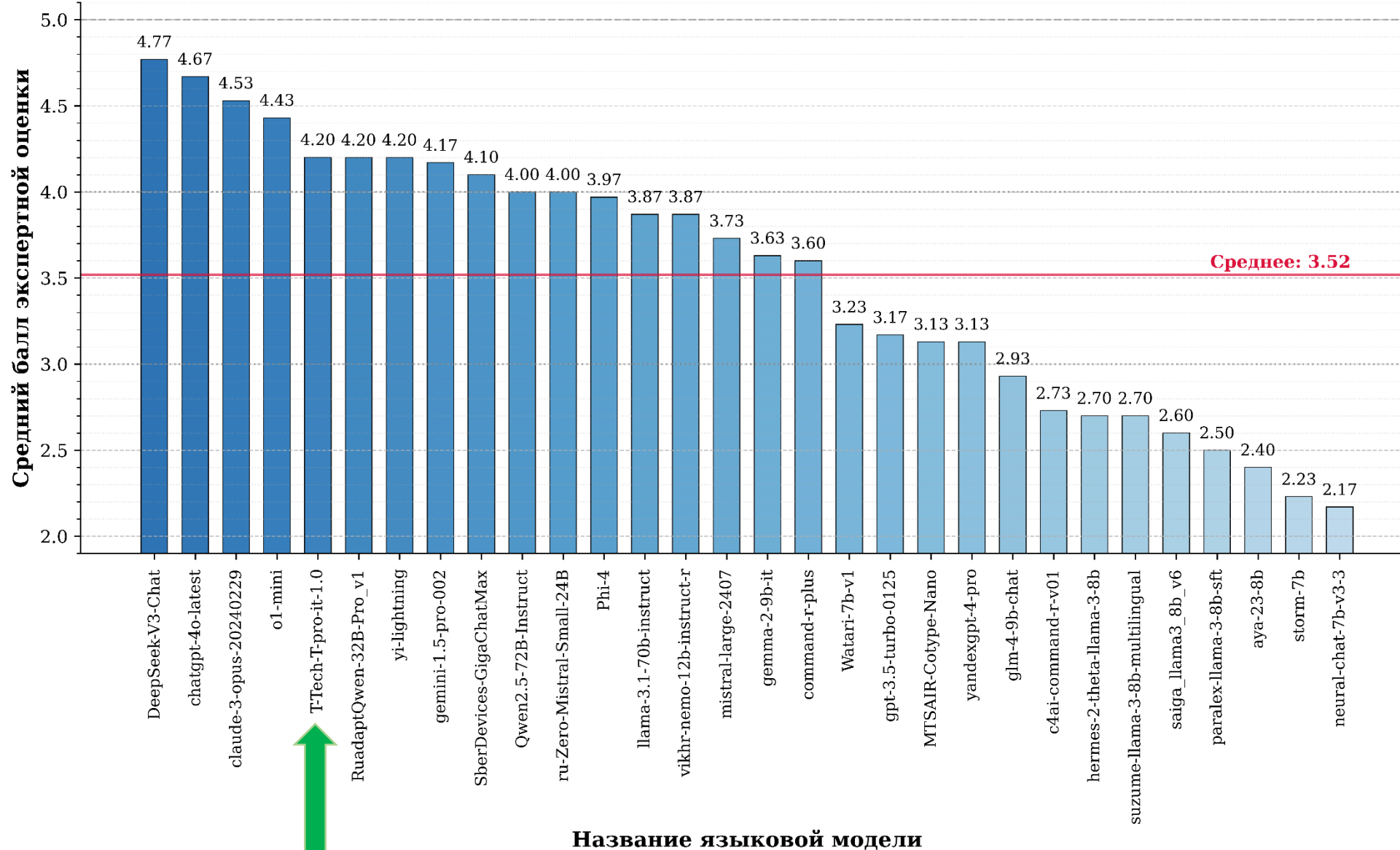
Модель	Полнота	Атомарность	Корректность	Средний балл
DeepSeek-V3-Chat	4.8	4.6	4.9	4.77
chatgpt-4o-latest	4.7	4.5	4.8	4.67
claude-3-opus-20240229	4.4	4.5	4.7	4.53
o1-mini	4.5	4.2	4.6	4.43
T-Tech-T-pro-it-1.0	4.2	4.1	4.3	4.20
RuadaptQwen-32B-Pro_v1	4.3	4.1	4.2	4.20
yi-lightning	4.3	4.1	4.2	4.20
gemini-1.5-pro-002	4.3	4.0	4.2	4.17

Проприетарные

Выбранная
модель



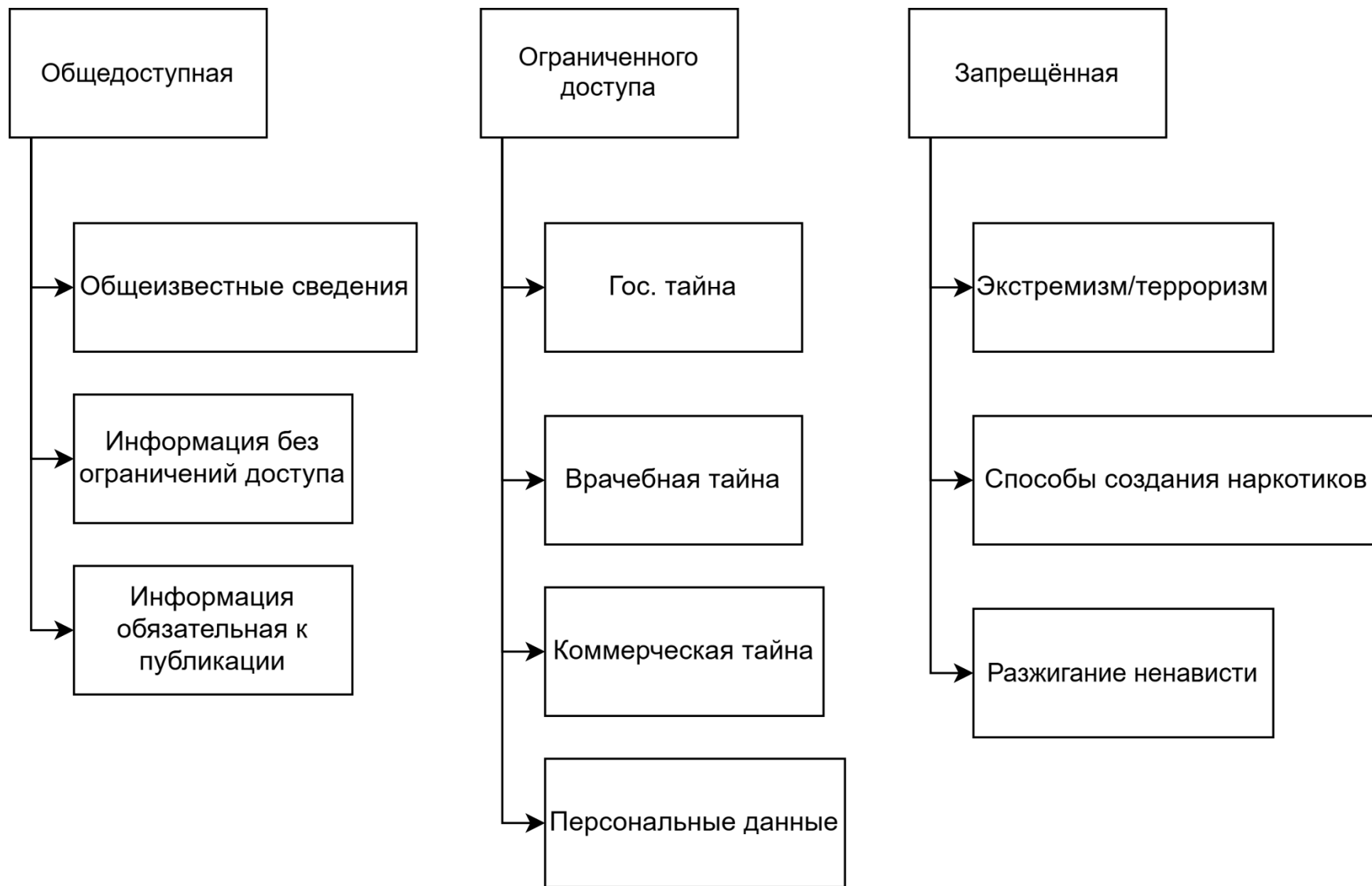
Результаты экспертной оценки (2/2)



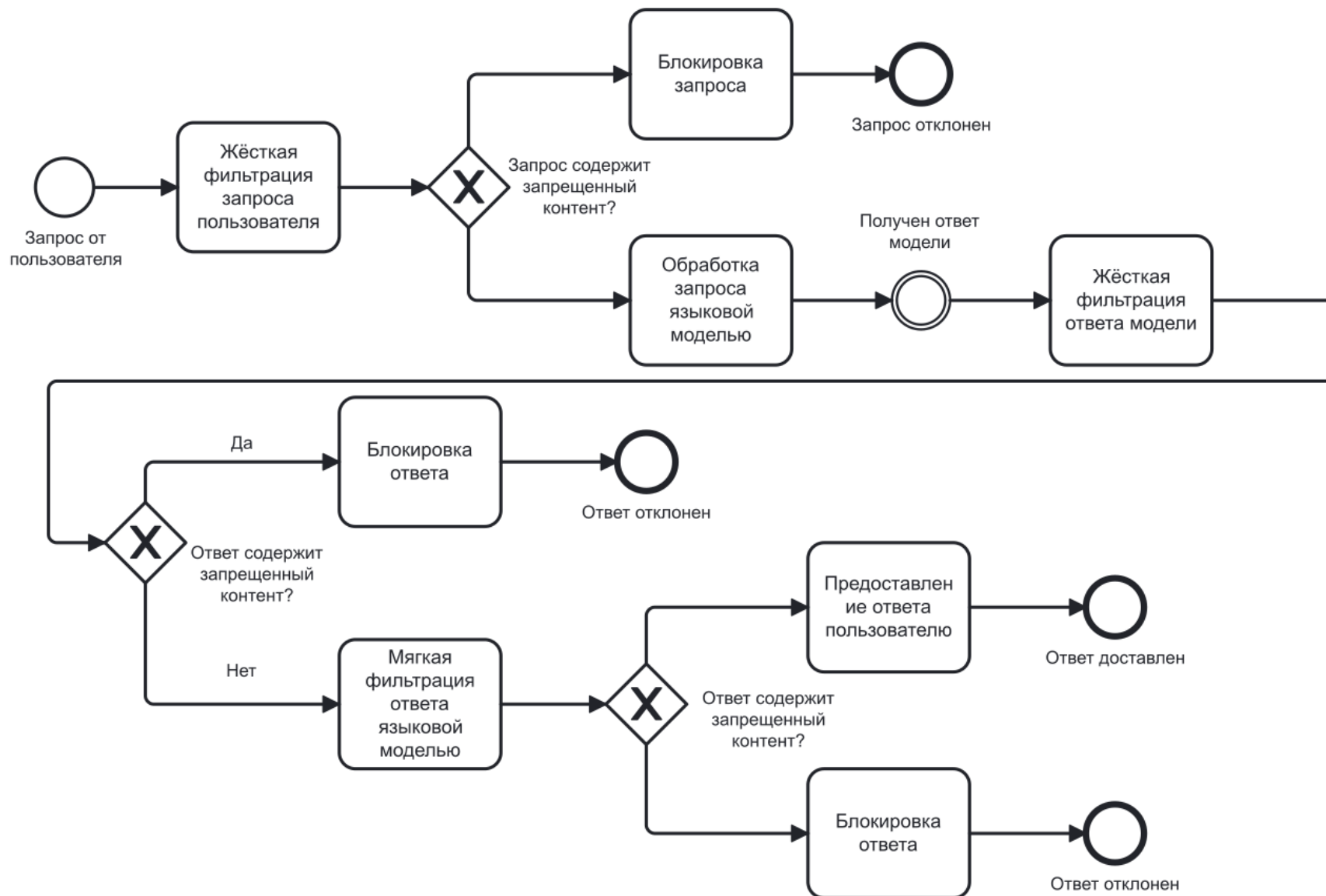
Лицензионная политика проекта

Критерий	Лицензия MIT	Лицензия Apache 2.0	Сравнение
Коммерческое использование	Разрешено	Разрешено	Одинаково
Модификация	Разрешено без ограничений	Разрешено с требованием указания изменений	Apache 2.0 строже
Распространение	Разрешено	Разрешено с дополнительными требованиями	Apache 2.0 строже
Частное использование	Разрешено	Разрешено	Одинаково
Указание изменений	Не требуется явно	Требуется явное указание	Apache 2.0 строже
Защита авторских прав	Требуется включения уведомления	Требуется сохранения всех уведомлений	Apache 2.0 строже
Патентные права	Нет явного предоставления	Явное предоставление патентных прав	Apache 2.0 строже
Ограничение торговых марок	Отсутствует	Прямо запрещает использование	Apache 2.0 строже
Распространение производных	Разрешено без особых условий	Разрешено с дополнительными условиями	Apache 2.0 строже

Ключевые виды информации по ФЗ-149



Многоуровневая система фильтрации запроса



Матрица распределения ответственности

Правовой аспект	Ответственность разработчика	Ответственность пользователя
Обновление баз фильтрации	Предоставляет технический механизм обновления	Обязан регулярно обновлять списки запрещенных слов согласно ст. 10.1 ФЗ-149
Обработка персональных данных	-	Несет полную ответственность по ФЗ-152 и ст. 13.11 КоАП РФ
Фильтрация запрещенного контента	Реализует базовые механизмы фильтрации, обновление баз фильтрации	Отвечает за настройку и эффективность фильтров
Соблюдение лицензий	Обеспечивает правильное лицензирование (Apache 2.0)	Соблюдает условия лицензии при использовании
Верификация результатов	Не отвечает за проверку выдаваемой информации	Обязан верифицировать результаты на соответствие законодательству РФ
Учет отраслевых ограничений	Не отвечает за специфические отраслевые требования	Полностью отвечает за соблюдение отраслевых норм

Легенда:

Полная ответственность

Отсутствие ответственности

Заключение

В результате выполнения работы разработано программное решение для разбиения сложных вопросов на простые вопросы с использованием больших языковых моделей.

Выполнены следующие задачи:

1. Проанализированы существующие методы декомпозиции вопросов.
2. Разработан алгоритм декомпозиции на основе больших языковых моделей.
3. Реализовано программное решение.
4. Проведено исследование эффективности разработанного решения.
5. Проанализированы правовые аспекты программной реализации и выбрана лицензия.