



Министерство науки и высшего образования Российской Федерации

Федеральное государственное автономное образовательное
учреждение высшего образования
ТОМСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ СИСТЕМ
УПРАВЛЕНИЯ И РАДИОЭЛЕКТРОНИКИ (ТУСУР)

Кафедра автоматизированных систем управления (АСУ)

Программный комплекс интеллектуального поиска в корпоративных документах

Выполнил: студент гр. 431-3
Бекиш Егор
Павлович

Руководитель:
доцент кафедры АСУ, к.т.н.
Суханов Александр
Яковлевич

Томск 2025

Цель и задачи

Цель работы

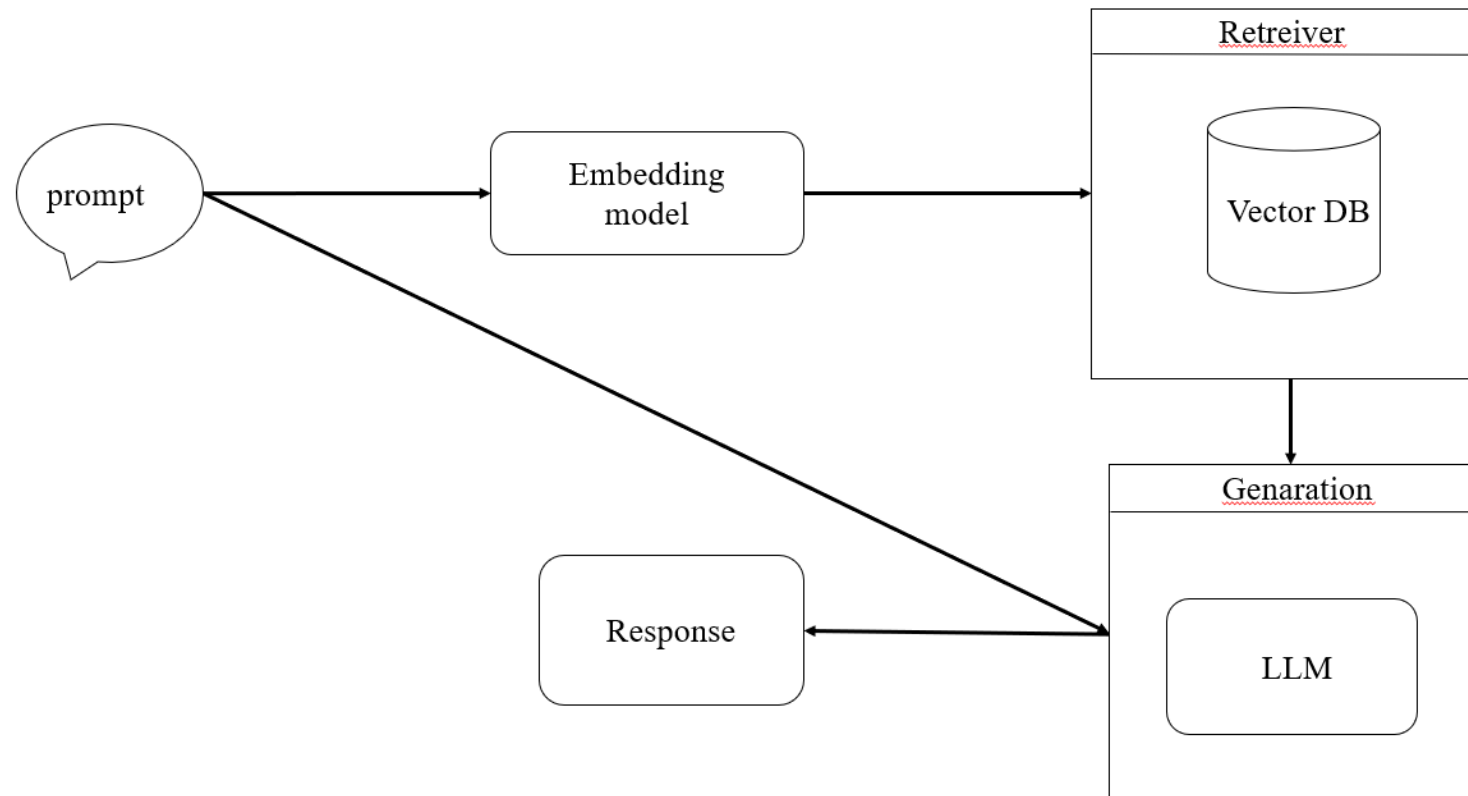
разработать программный комплекс, который сокращает время доступа сотрудников к корпоративной информации за счет обработки запросов к внутренним документам предприятия, формата txt, pdf, docx, doc, pptx, ppt, на естественном языке и средствами ИИ.

Задачи:

1. Анализ необходимых требований;
2. Реализация функциональных возможностей и выработка спецификаций;
3. Проектирование;
4. Реализация системы;
5. Тестирование.

Обзор предметной области

- Выборка необходимых документов и данных
- Развитие векторного представления слов
- Виды языковых моделей



Обзор предметной области

- Выборка необходимых документов и данных
- Развитие векторного представления слов
- Виды языковых моделей

Обзор предметной области

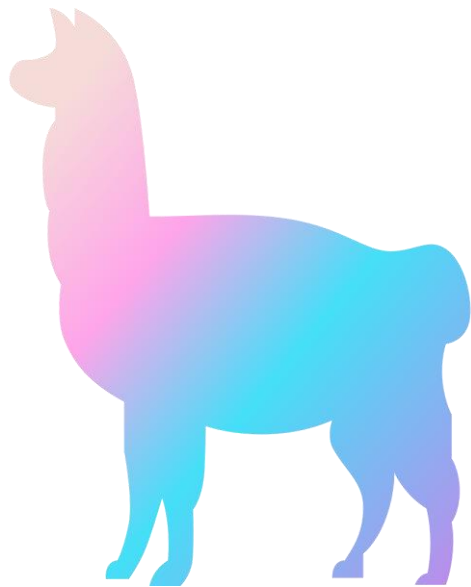
- Выборка необходимых документов и данных
- Развитие векторного представления слов
- **Виды языковых моделей**

Модель	Вычислительная сложность	Длина контекста L	Пропускная способность
Трансформер	L^2	10^3	x
Mamba	L	10^6	$5x$

- IBM Watson
- Azure Cognitive Services от Microsoft
- Google Cloud AI
- Amazon Lex
- OpenAI GPT
- Hugging Face Chat
- DeepSeek

Функциональные требования

Языковые модели



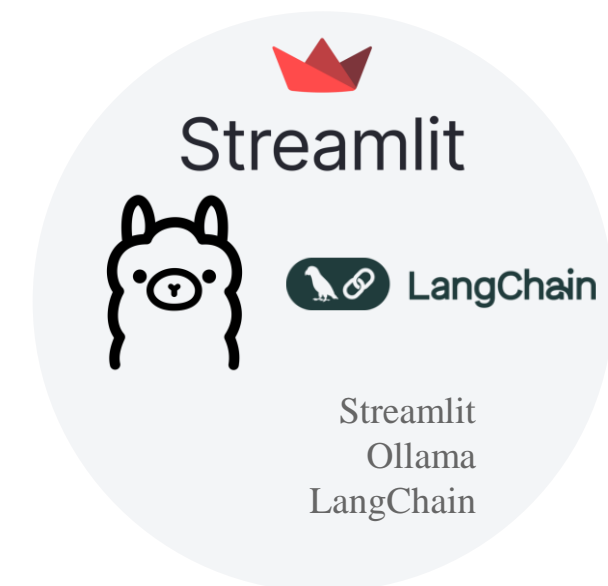
Функциональные требования

Инструменты реализации

1. IDE, ЯП, Менеджер зависимостей

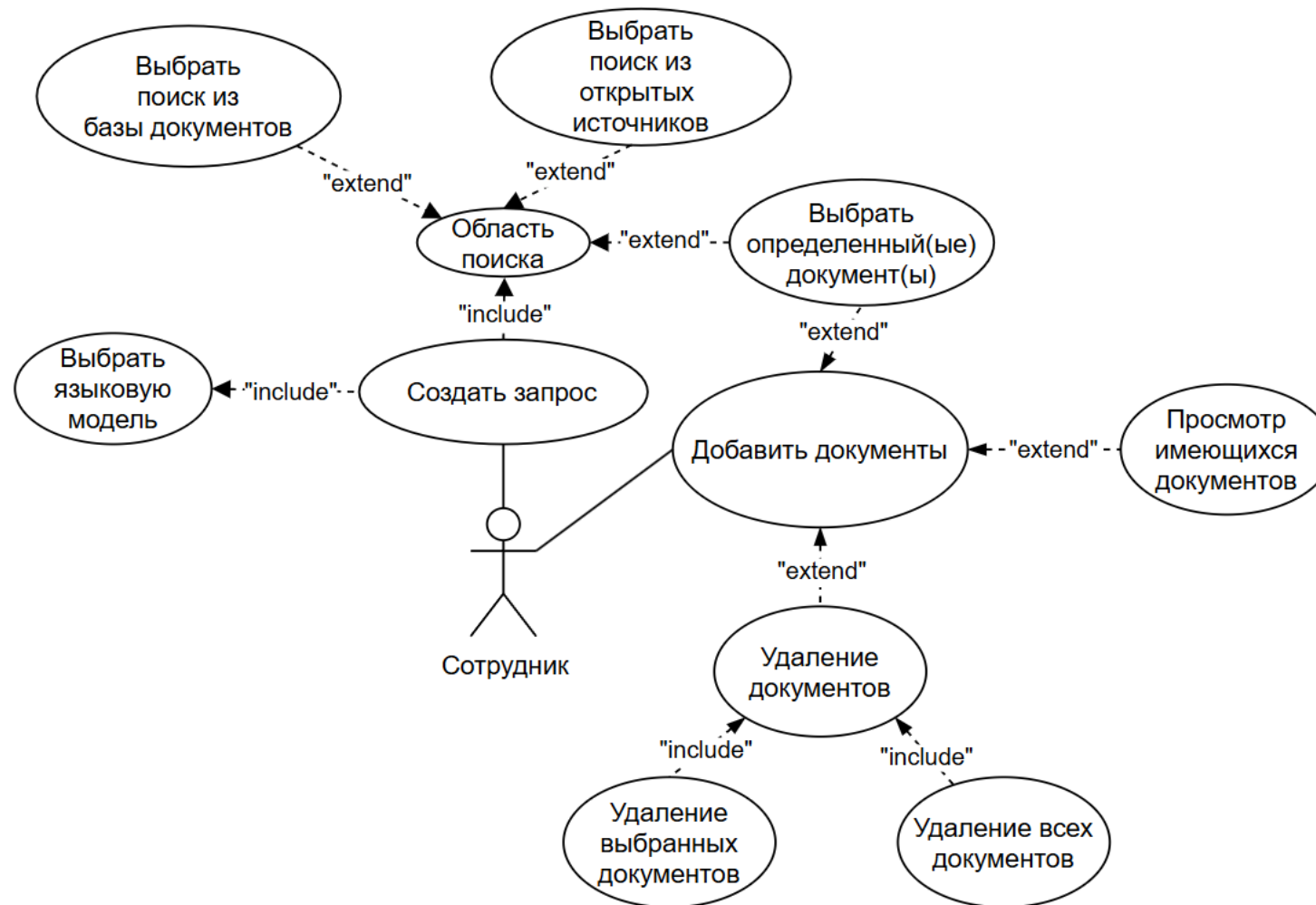


2. Библиотеки



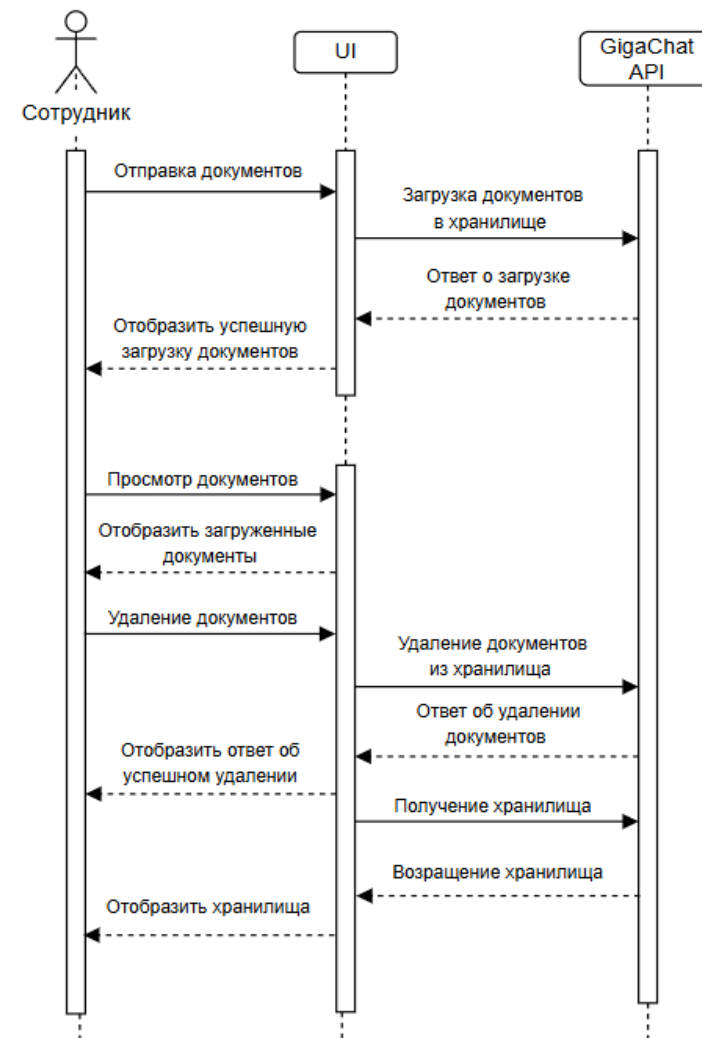
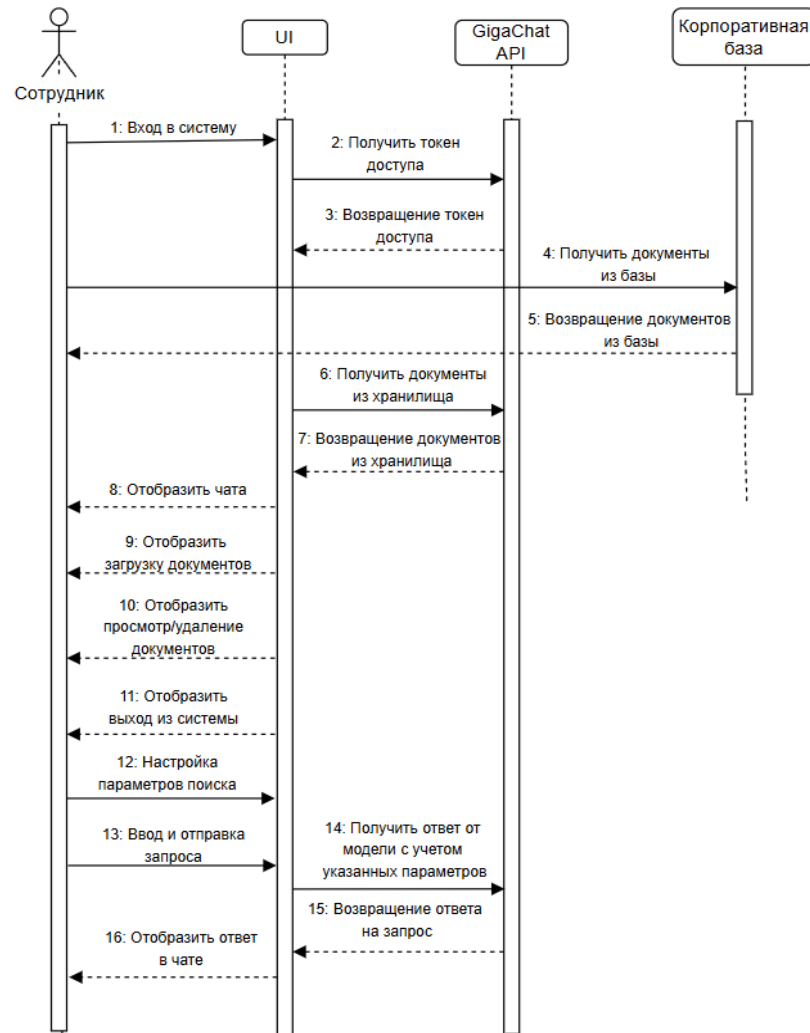
Проектирование

Диаграмма прецедентов

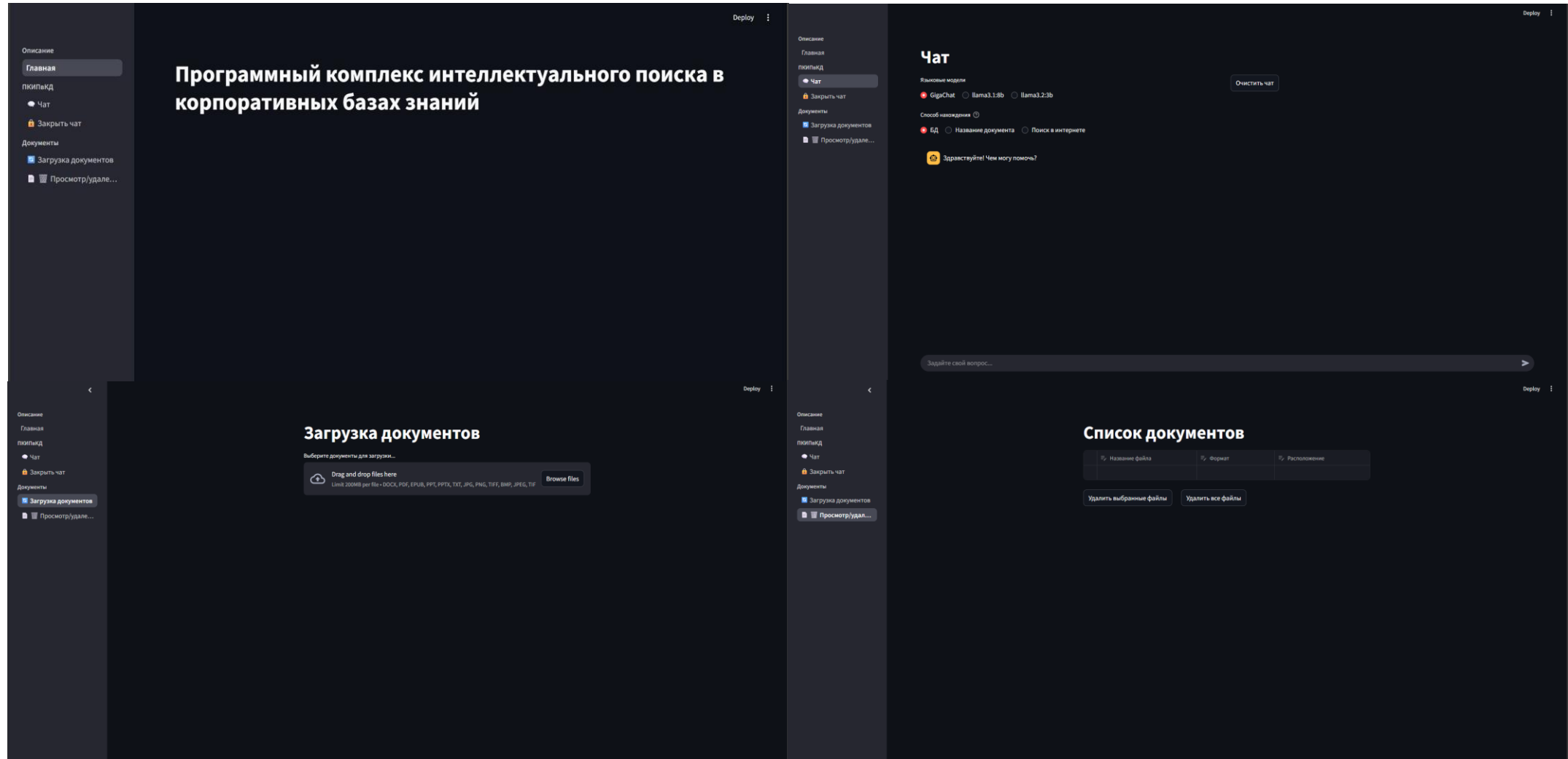


Проектирование

Диаграмма последовательности

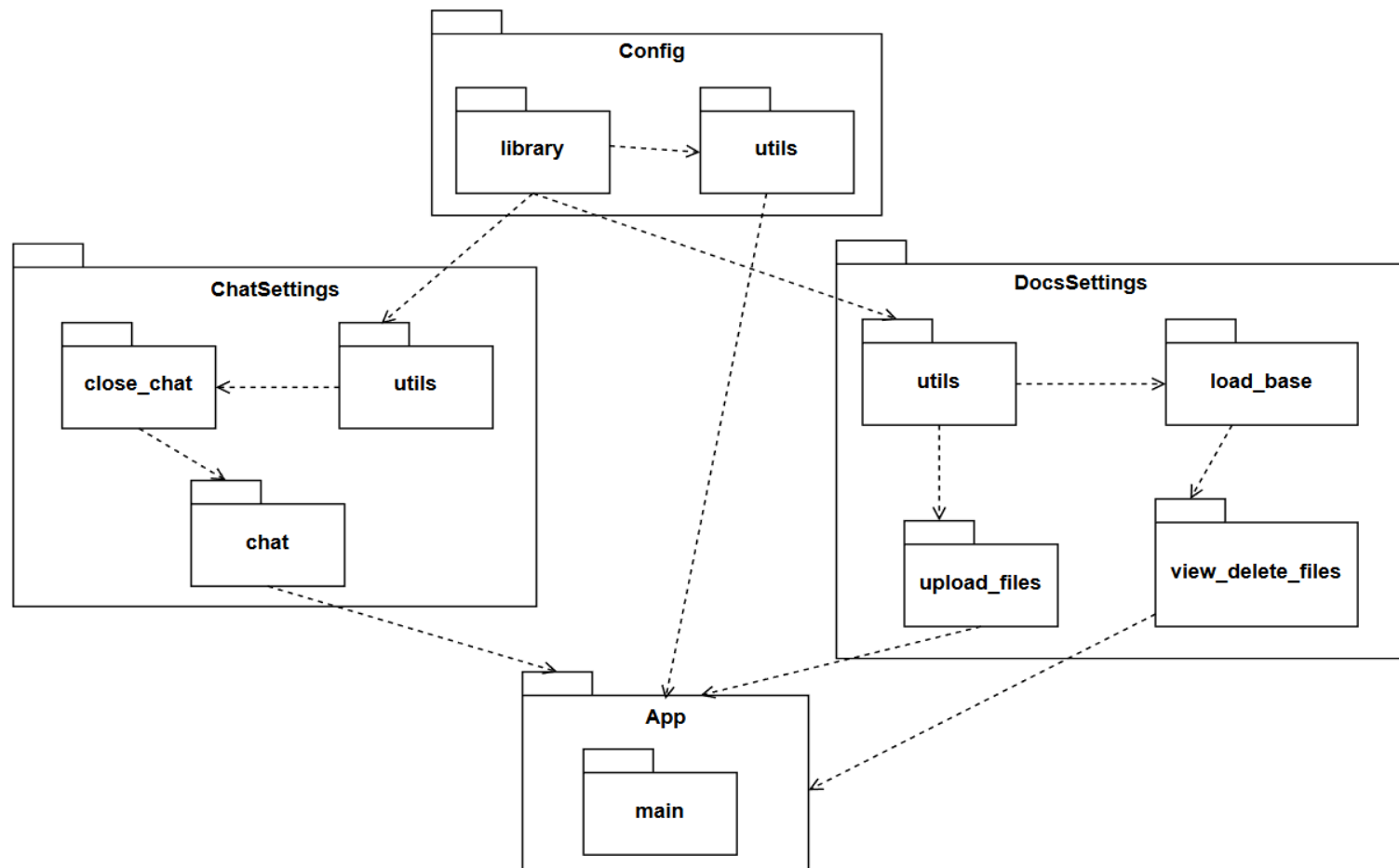


Проектирование Макеты интерфейса

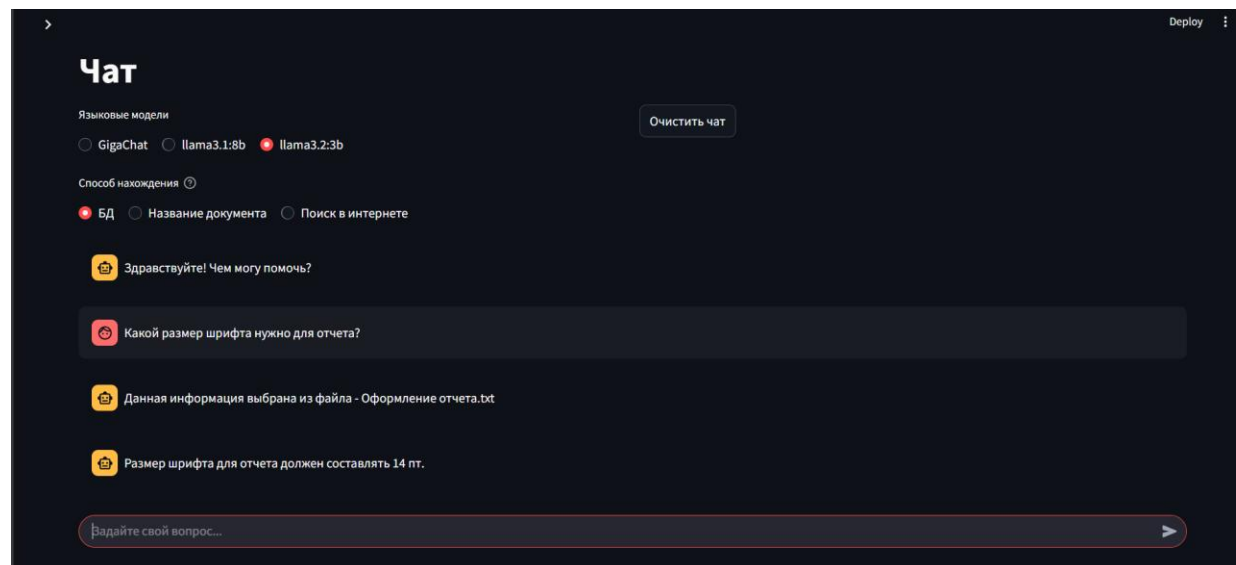
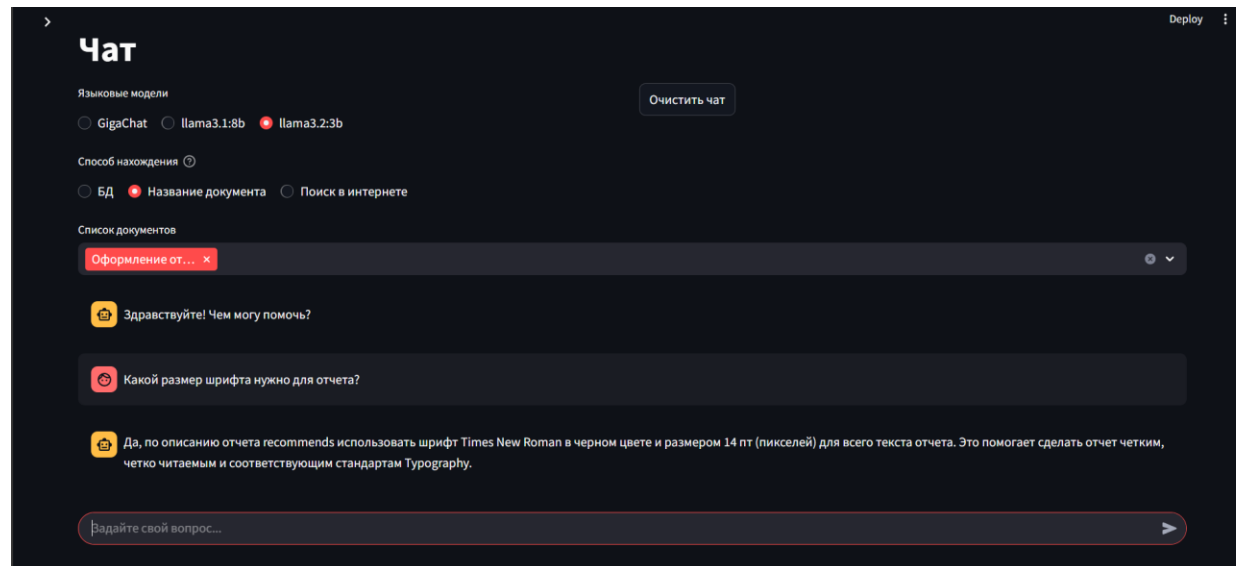


Реализация

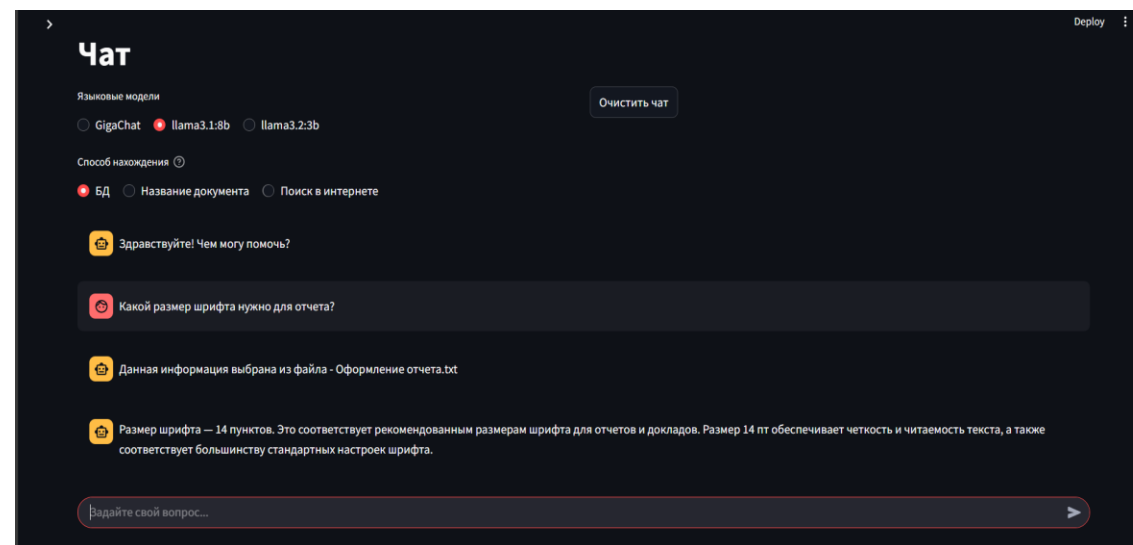
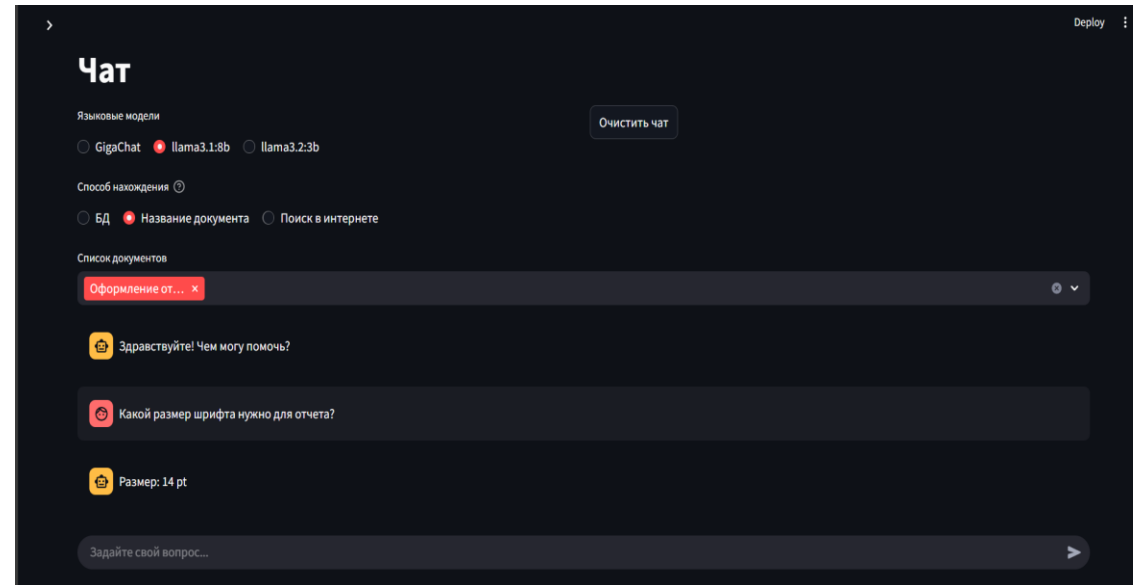
Диаграмма пакетов



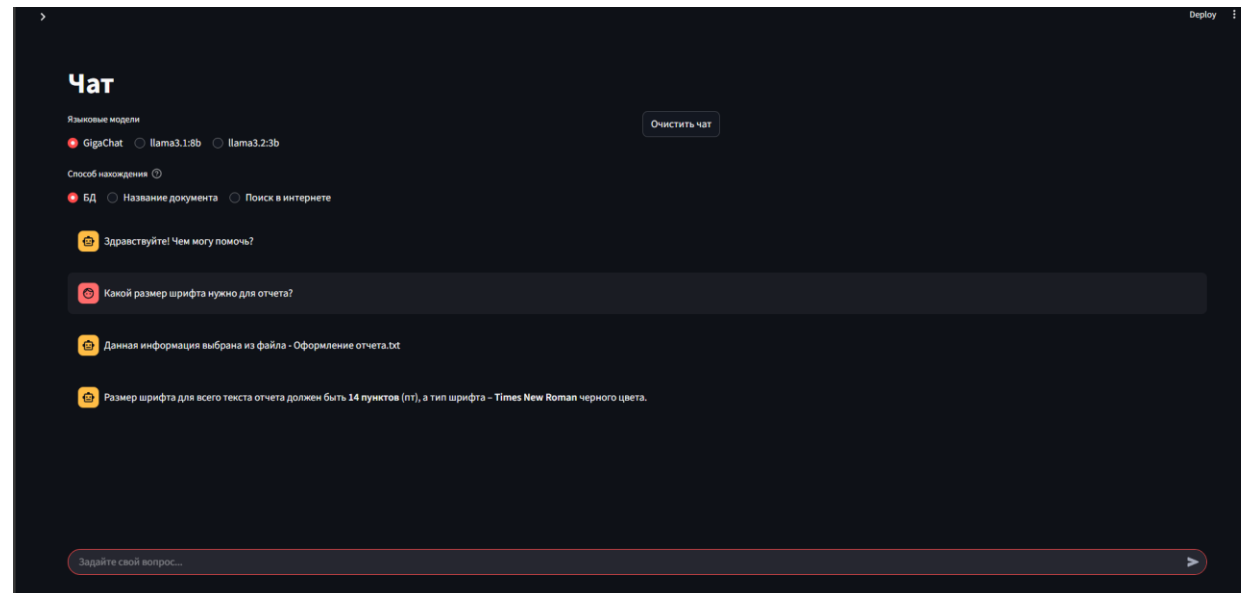
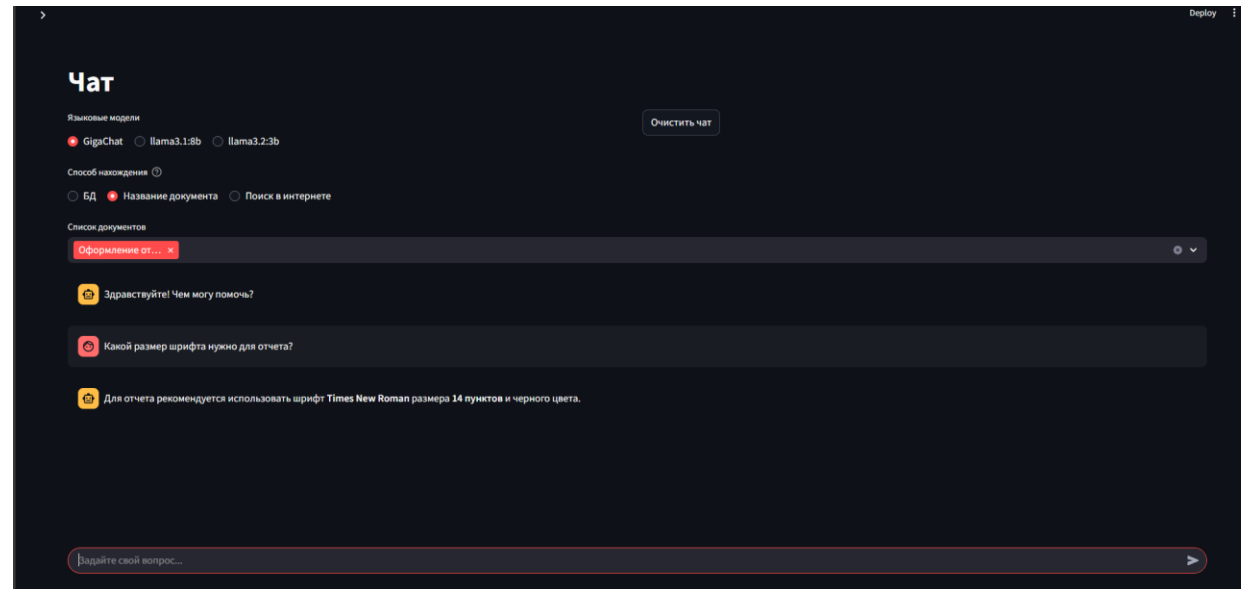
Тестирование



Тестирование



Тестирование



Заключение

В ходе выполнения работы был разработан программный комплекс интеллектуального поиска в корпоративных документах, который позволяет повысить эффективность работы с информационными данными, что поможет ускорить бизнес-процессы.

Дальнейшее развитие программного комплекса планируется в рамках магистерской диссертации.

По теме бакалаврской работы была подготовлена статья:

- Бекиш Е.П. Система интеллектуального поиска в корпоративных базах знаний // XIII Региональная научно-практическая конференция «НАУКА И ПРАКТИКА: ПРОЕКТНАЯ ДЕЯТЕЛЬНОСТЬ – ОТ ИДЕИ ДО ВНЕДРЕНИЯ – 2024». Принято к публикации.