

#### 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

# 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.

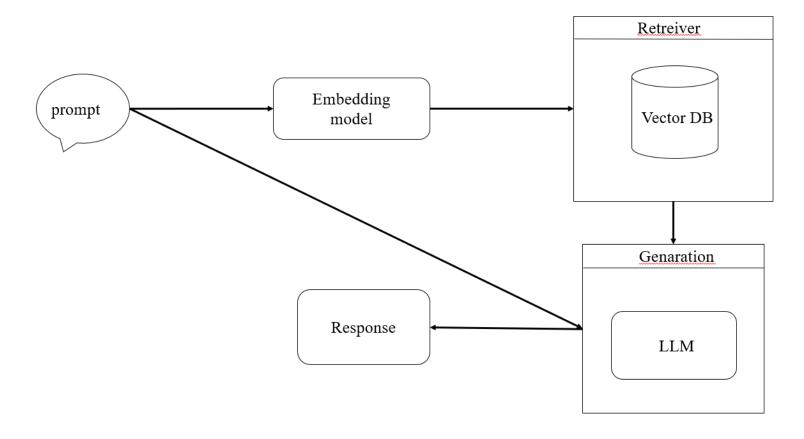
#### Задачи:

- 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 <sup>3</sup>	x
Mamba	L	$10^{6}$	5 <i>x</i>

### Обзор аналогов

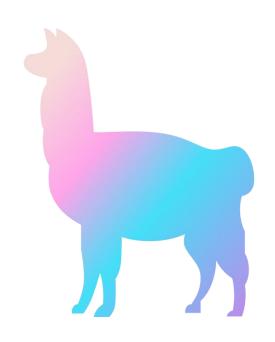


- IBM Watson
- Azure Cognitive Services or Microsoft
- Google Cloud AI
- Amazon Lex
- OpenAI GPT
- Hugging Face Chat
- DeepSeek

## Функциональные требования

TYCYP
TUSUR UNIVERSITY

Языковые модели



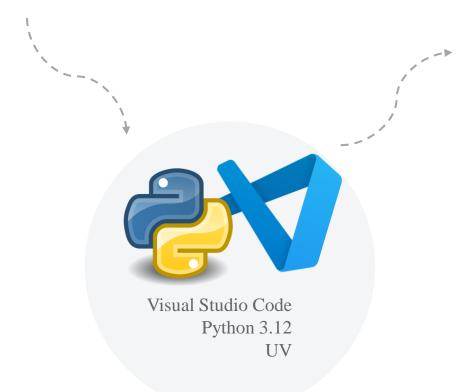


### Функциональные требования

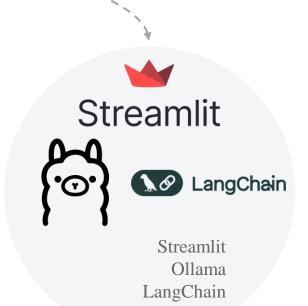
#### TYCYP TUSUR UNIVERSITY

#### Инструменты реализации

#### 1. IDE, ЯП, Менеджер зависимостей



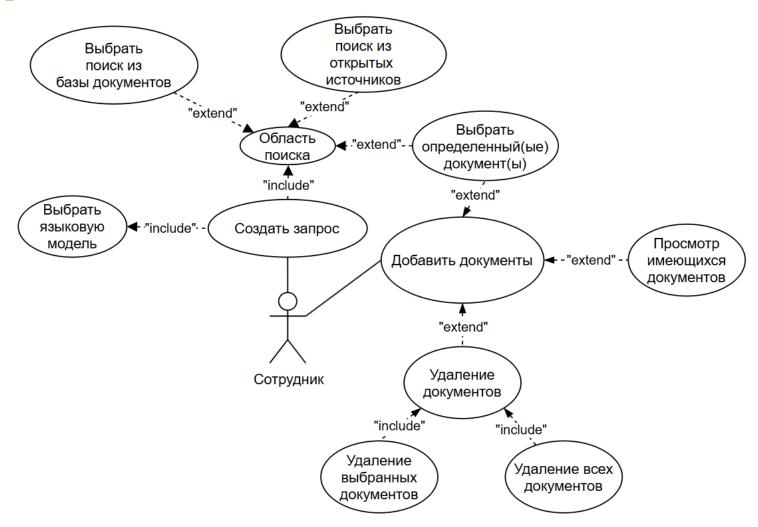
2. Библиотеки



## Проектирование

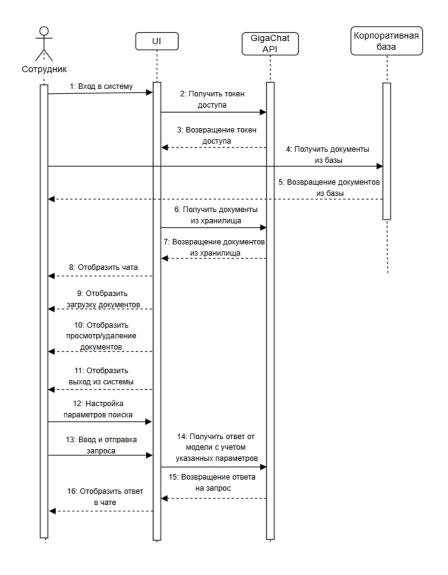
#### Диаграмма прецендентов



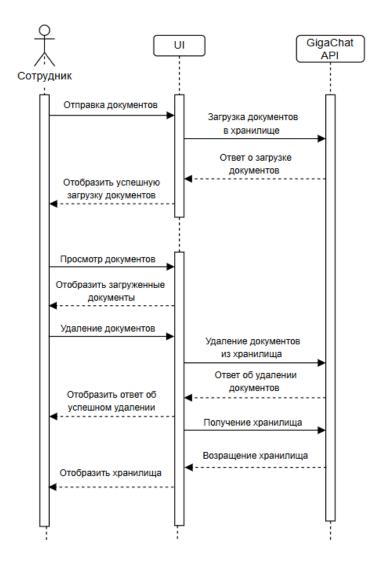


### Проектирование

#### Диаграмма последовательности



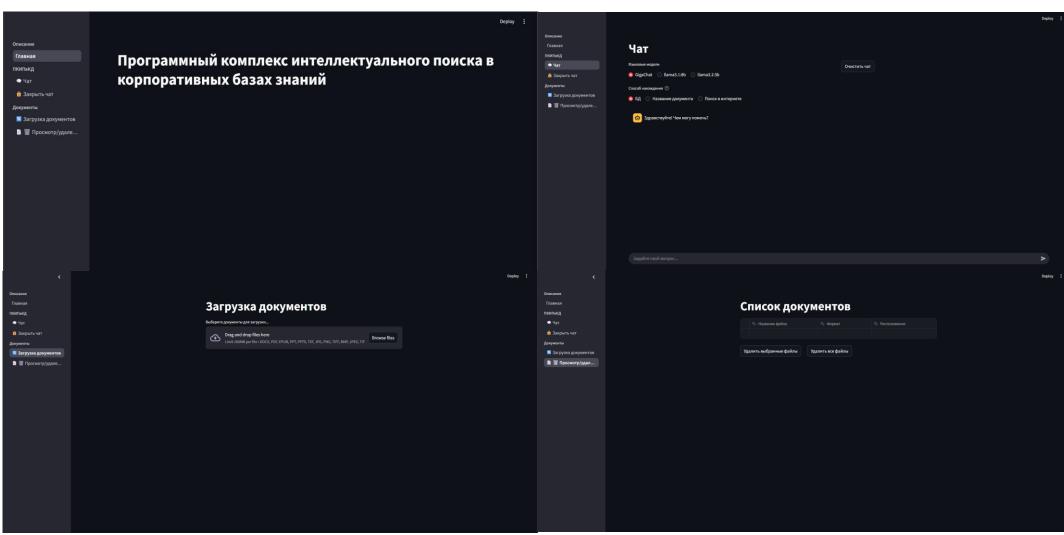




### Проектирование

#### Макеты интерфейса

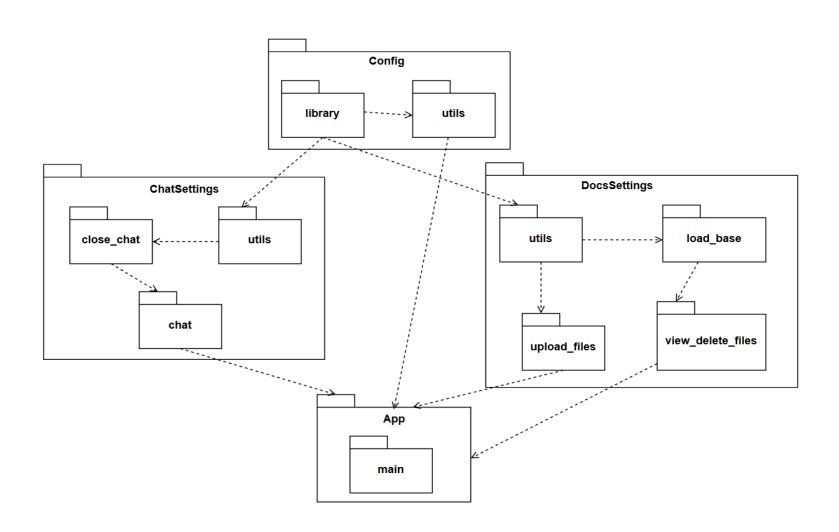




### Реализация

#### Диаграмма пакетов





# Тестирование

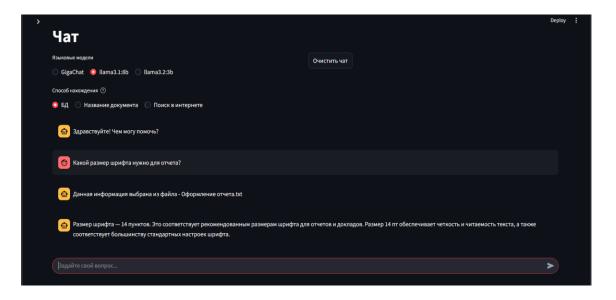


		De	ploy :
	Чат		- 1
	Языковые модели	Очистить чат	
	○ GigaChat ○ llama3.1:8b		
	Способ нахождения ③		
	БД		
	Список документов		
	Оформление от ×		
	Здравствуйте! Чем могу помочь?		
	Какой размер шрифта нужно для отчета?		
	Да, по описанию отчета recommends использовать шрифт Times New Roman в черном цв четко читаемым и соответствующим стандартам Typography.	оте и размером 14 пт (пикселей) для всего текста отчета. Это помогает сделать отчет четким,	
	Вадайте свой вопрос	>	)
			ĺ
- >		De	ploy :
>		De	ploy :
<b>,</b>	Чат	De	ploy :
,	<b>Чат</b> языковые модели	Очистить чат	ploy :
,			ploy :
,	Языковые модели		ploy :
>	Языковые модели  GigaChat Ulama3.1:8b Ilama3.2:3b		ploy :
>	Языковые модели  GigaChat □ Ilama3.1:8b  Ilama3.2:3b  Способ нахождения ③		ploy :
,	Языковые модели  GigaChat ☐ Ilama3.1:8b		ploy :
,	Языковые модели		ploy :
,	Языковые модели		pioy [

# Тестирование



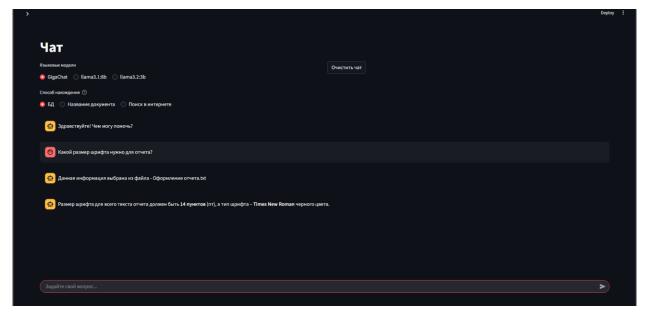
>	Deplo	y <b>:</b>
Чат		
Языковые модели	Очистить чат	
◯ GigaChat		
Способ нахождения ③		
○ БД		
Список документов		
Оформление от х	0 🗸	
<b>②</b> Здравствуйте! Чем могу помочь?		
Какой размер шрифта нужно для отчета?		
<mark> </mark>		
Задайте свой вопрос	>	



# **Тестировани**е



<b>,</b>		Deploy :
Чат		
Языковые модели	Очистить чат	
⑤ GigaChat ○ Ilama3.1:8b ○ Ilama3.2:3b		
Способ нахождения ①		
○ БД		
Список документов		
Оформление от х		
Здравствуйте! Чем могу помочь?		
Какой размер шрифта нужно для отчета?		
Вадайте свой вопрос		>
	<u> </u>	



#### 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.