



Балтийский государственный
технический университет
«ВОЕНМЕХ»
им. Д.Ф. Устинова

Выпускная квалификационная работа бакалавра на тему «Разработка API для автоматизации сбора, структурирования, анализа и фильтрации данных социальных сетей»

Самохин Аркадий Владимирович

Студент группы И591

Руководитель: Ананченко И.В., к.т.н, доцент



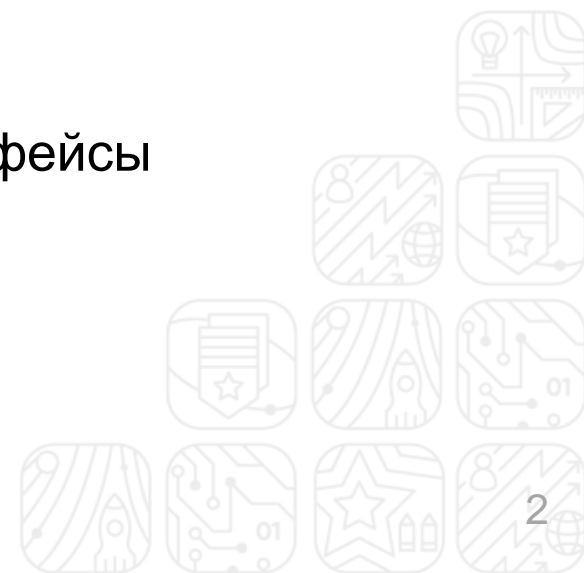


Цель и задачи ВКР

Цель работы – разработать API для автоматизации сбора, структурирования, анализа и фильтрации данных социальных сетей.

Задачи:

- проанализировать структуру социальной сети ВКонтакте и технологии проектирования API
- провести анализ существующих аналогов
- сформировать требования к программному продукту
- спроектировать приложение, пользовательский и программный интерфейсы
- выбрать и обосновать средства реализации
- выполнить программную реализацию
- произвести тестирование программного продукта



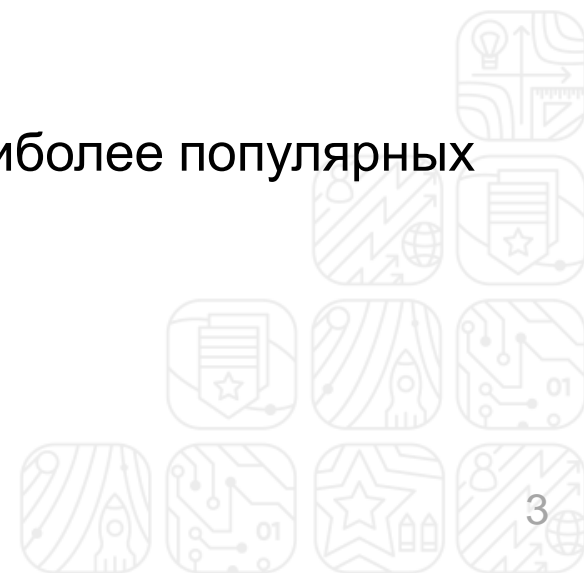


Анализ предметной области

Социальные сети представляют собой онлайн-платформы, которые позволяют пользователям создавать профили, устанавливать связи с другими пользователями и обмениваться информацией.

Используются не только как средство общения и поддержания связей, но и важный инструмент для бизнеса, который используется для продвижения продуктов и услуг, увеличения охвата и узнаваемости бренда.

Используемая социальная сеть - ВКонтакте, являющаяся одной из наиболее популярных в России и близлежащих странах.





Анализ структуры социальной сети ВКонтakte

API ВКонтakte представляет собой интерфейс, позволяющий сторонним приложениям и сервисам получать данные помощью HTTP-запросов к серверу. HTTP-запрос состоит из нескольких частей: протокола соединения, адреса API-сервиса, названия метода, соответствующий определенному разделу, и параметров запроса, которые являются входными данными.

Выходные данные имеют текстовый тип в формате JSON.

Для выполнения запроса необходим ключ-доступа (токен).





SMM-парсер – инструмент, использующийся в сфере социального медиа-маркетинга для автоматизации работы с данными из социальных медиа платформ для поиска целевой аудитории.

- сбор данных
- структурирование данных
- анализ данных
- фильтрация данных





Анализ технологий проектирования API

API – набор протоколов и правил, предоставляющий возможности взаимодействия программных продуктов друг с другом.

Существуют следующие стили архитектурного проектирования API:

- REST
- RPC
- Websocket
- SOAP
- GraphQL





Сравнение аналогов

Таблица сравнения аналогов

Название	Авторизация через социальную сеть	Русская локализ ация	Графическое отображение данных	Расширен ные условия филътрац ии	Анализ различного контента	Создание облачных списков
OkTarget	+	+	-+	-	-	-
Clever Target	+	+	-	-	+	+
PromoPult	-	+	-	-	-	-



Функциональные требования

- авторизация через социальную сеть
- получение результатов работы задач
- создание/удаление/поиск задач по сбору, анализу и фильтрации данных
- создание/удаление/поиск облачных списков
- добавление/удаление данных облачных списков







Шаблоны проектирования

Шаблоны проектирования, также называемые паттернами проектирования - это повторяемые решения проблем проектирования, которые могут быть применены для создания гибких, расширяемых и переиспользуемых систем. Шаблоны проектирования описывают общие проблемы и предлагают готовые решения для этих проблем.

Используемые шаблоны:

- MVC
- Очередь сообщений





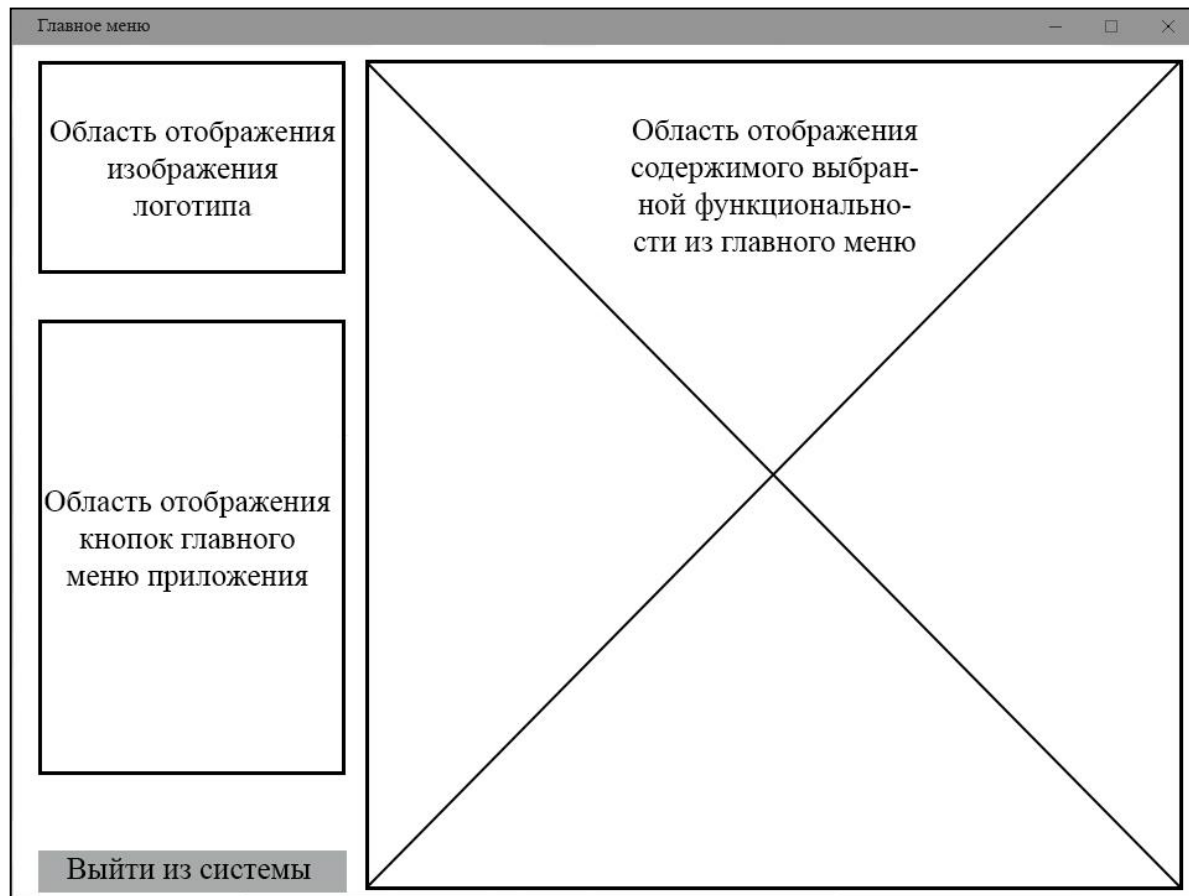
Проектирование клиентской части

Настольное ПО - это программное обеспечение, которое устанавливается и выполняется на локальной машине (компьютер, ноутбук), работающей под операционной системой, такой как Windows, macOS или Linux.

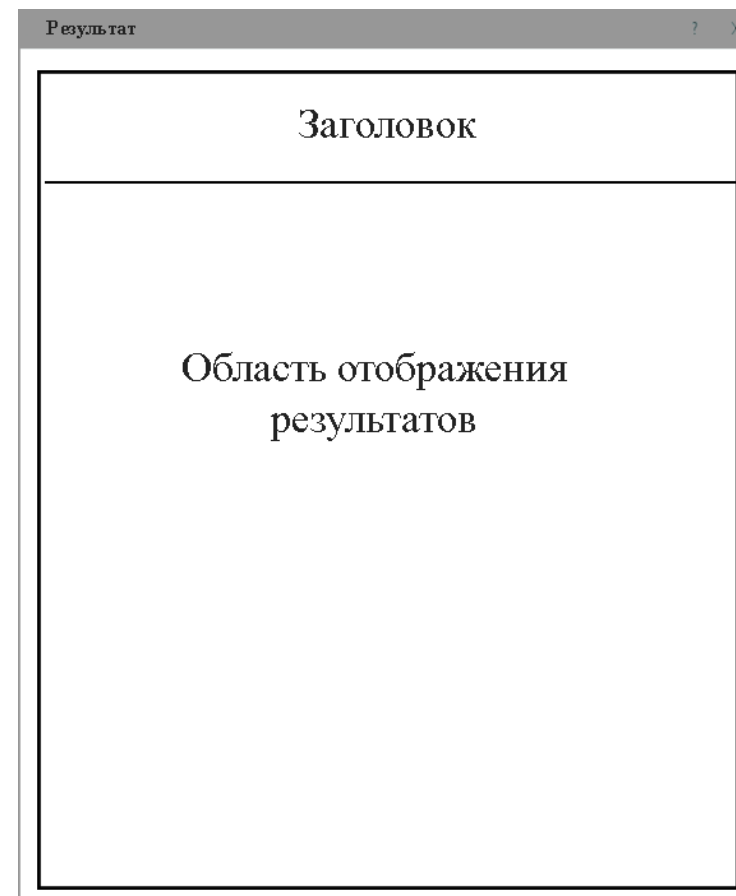




Прототипирование пользовательского интерфейса



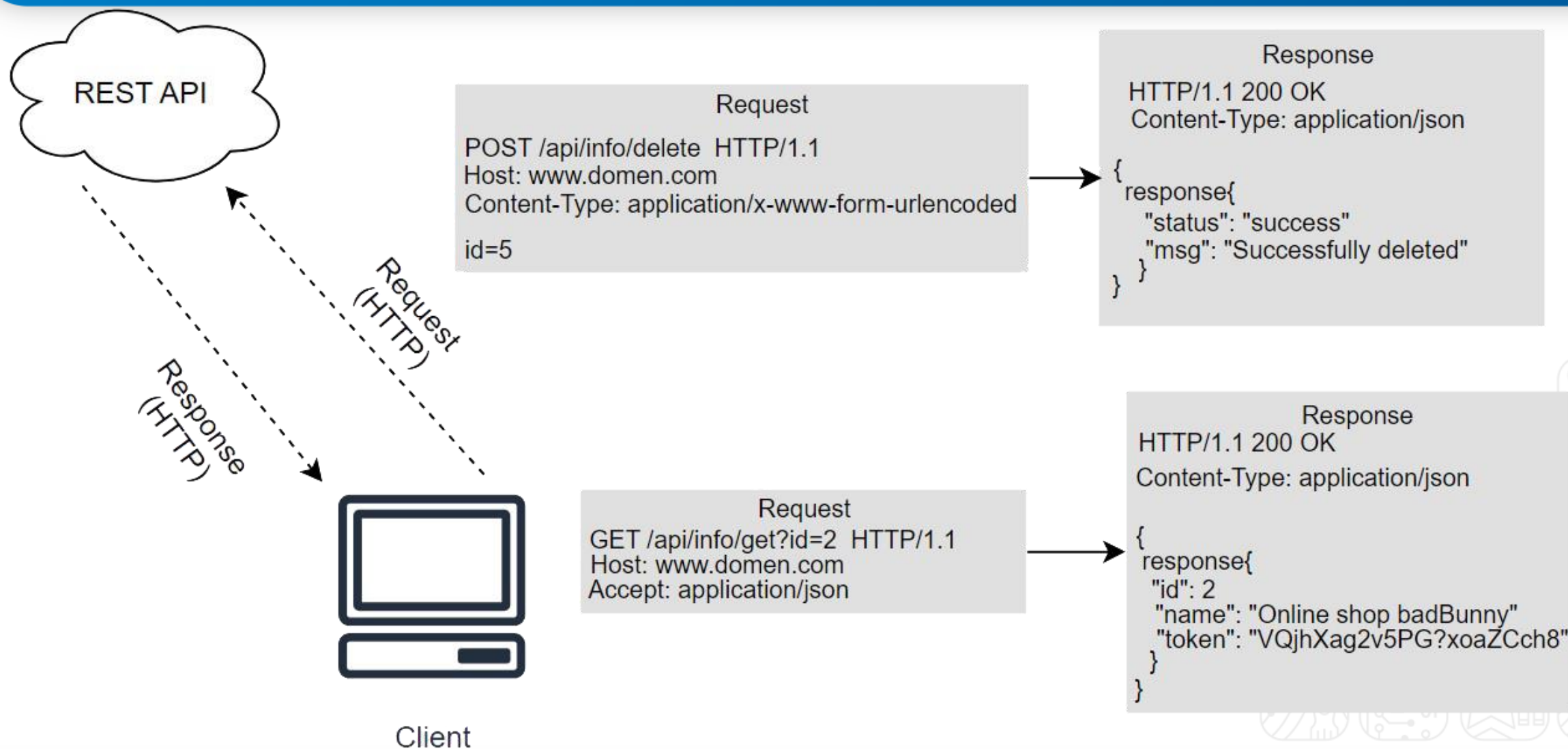
Главное окно



Окно отображения результата



Проектирование программного интерфейса





Проектирование программного интерфейса

Используются HTTP-запросы с необходимыми параметрами и токеном доступа, отсылаемые на URL-ссылки.

Входные и выходные данные в формате JSON.

<http://5.63.155.164/api/parser/analysis/audits/>

Адрес сервер

Обращение
к API

Обращение
к парсеру

Функции парсера:
-analysis (анализ)
-search (поиск)
-filter (фильтрация)

Контент:

- audits (аудитория активная)
- groups (сообщества)
- posts (посты)
- пользователи (users)

<http://5.63.155.164/api/auth/url/>

Адрес сервер

Обращение
к API

Обращение к
авторизации

Функции:
-url (получение ссылки)
-code (получение кода)
-login (авторизация)
-logout (выход)



Проектирование программного интерфейса

```
{  
  "title": "Поиск групп по запчастям",  
  "keyword_words": "запчасти",  
  "count_limit": 40,  
  "search_status": 1  
}
```

JSON для сбора

```
{  
  "title": "Поиск групп по запчастям",  
  "task_id": 4,  
  "keyword_words": "запчасти",  
  "description_words": "авто машины",  
  "count_limit": 40,  
  "age_filter": {  
    "14-18": {  
      "operator": "lt",  
      "percent": 15  
    },  
    "18-24": {  
      "operator": ",",  
      "percent":  
    },  
    "24-30": {  
      "operator": "gt",  
      "percent": 30  
    },  
    "30-35": {  
      "operator": ",",  
      "percent":  
    },  
    "35-45": {  
      "operator": "lt",  
      "percent": 8  
    },  
    "45-50": {  
      "operator": ",",  
      "percent":  
    }  
  }  
}
```

JSON для фильтрации

```
{  
  "title": "Анализ аудитории списка",  
  "list_id": 2,  
}
```

JSON для анализа





Средства разработки

Средства разработки клиентской части:

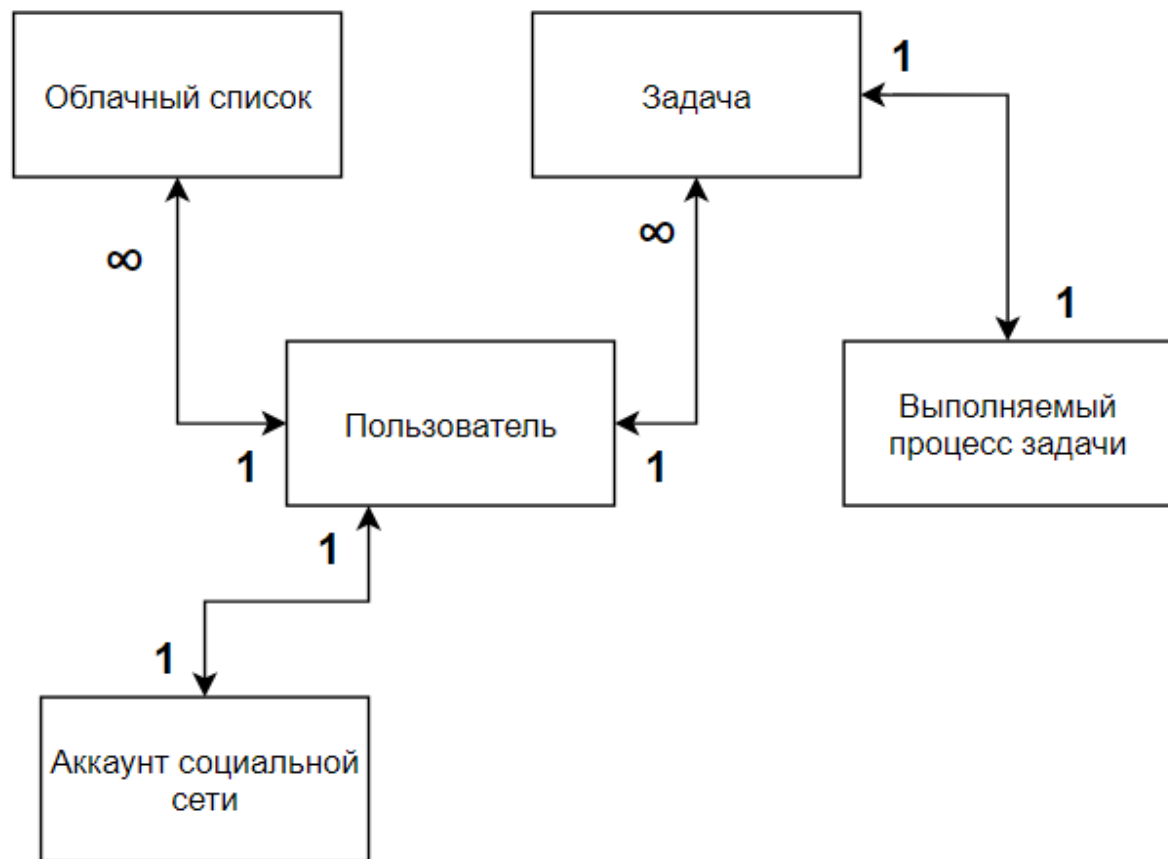
- язык программирования C++
- фреймворк Qt

Средства разработки серверной части:

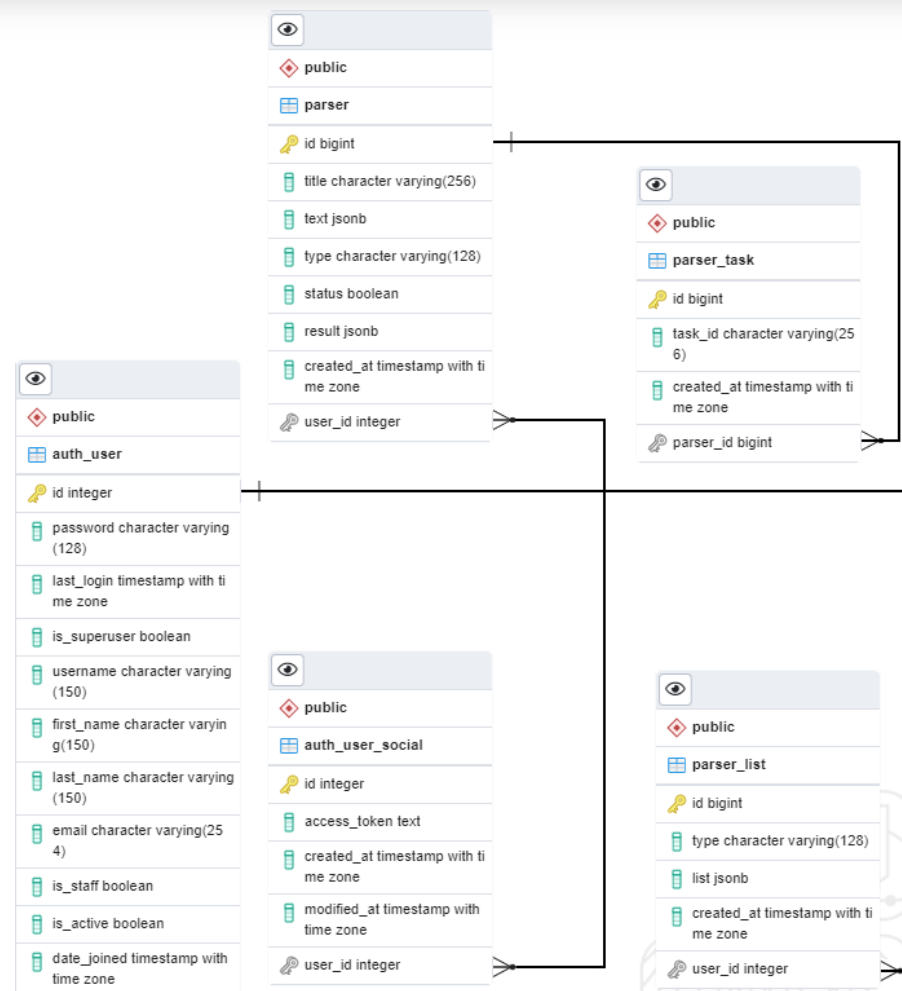
- язык программирования Python
- фреймворк Django
- система управления базой данных PostgreSQL
- брокер сообщений Redis



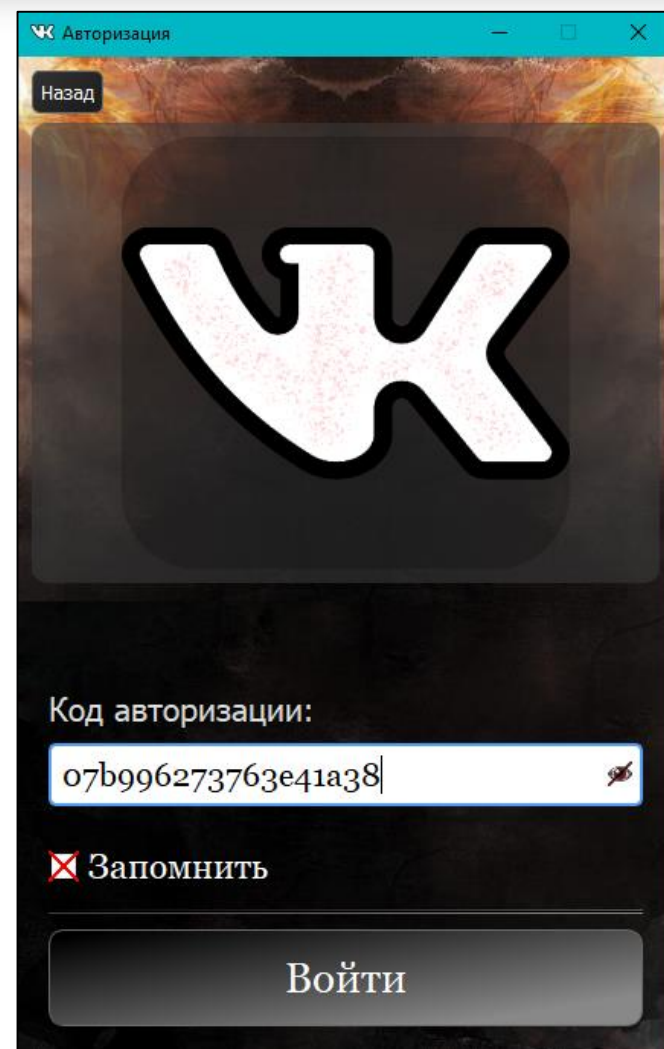
База данных



Инфологическая модель




Даталогическая модель





Реализация приложения

Приложение



Парсер

Задания

Списки

Сообщества

Поиск сообществ

Фильтрация сообществ

Анализ сообществ

Пользователи

Аудитория

Посты

Выйти из системы

ПОИСК СООБЩЕСТВ

Название задачи

Питомники в городе "Москва"

Ключевые слова

питомники Москва

Минус слова

Введите минус слова через запятую

Подписчики

от 1000 до Верхняя граница

Где искать

☐ Поиск в статусе

☒ Поиск в описании

☐ Включенные сообщения

Тип сообщества

☒ Любой

☐ Группа

☐ Публичная

☐ Мероприятие

Ограничение количества результатов


20

Поиск

Результат

НАЙДЕННЫЕ СООБЩЕСТВА

Количество: 20




Москва и МО - потеряшки | Пропала собака, кошка

Москва: пропала собака, кошка

Подписчики: 22735

Публичная страница [vk 32880564](#)



Питомник Розебук. Купить саженцы роз

Розебук. Интернет-магазин питомника подмосковных саженцев роз от производителя с доставкой курьером и почтой

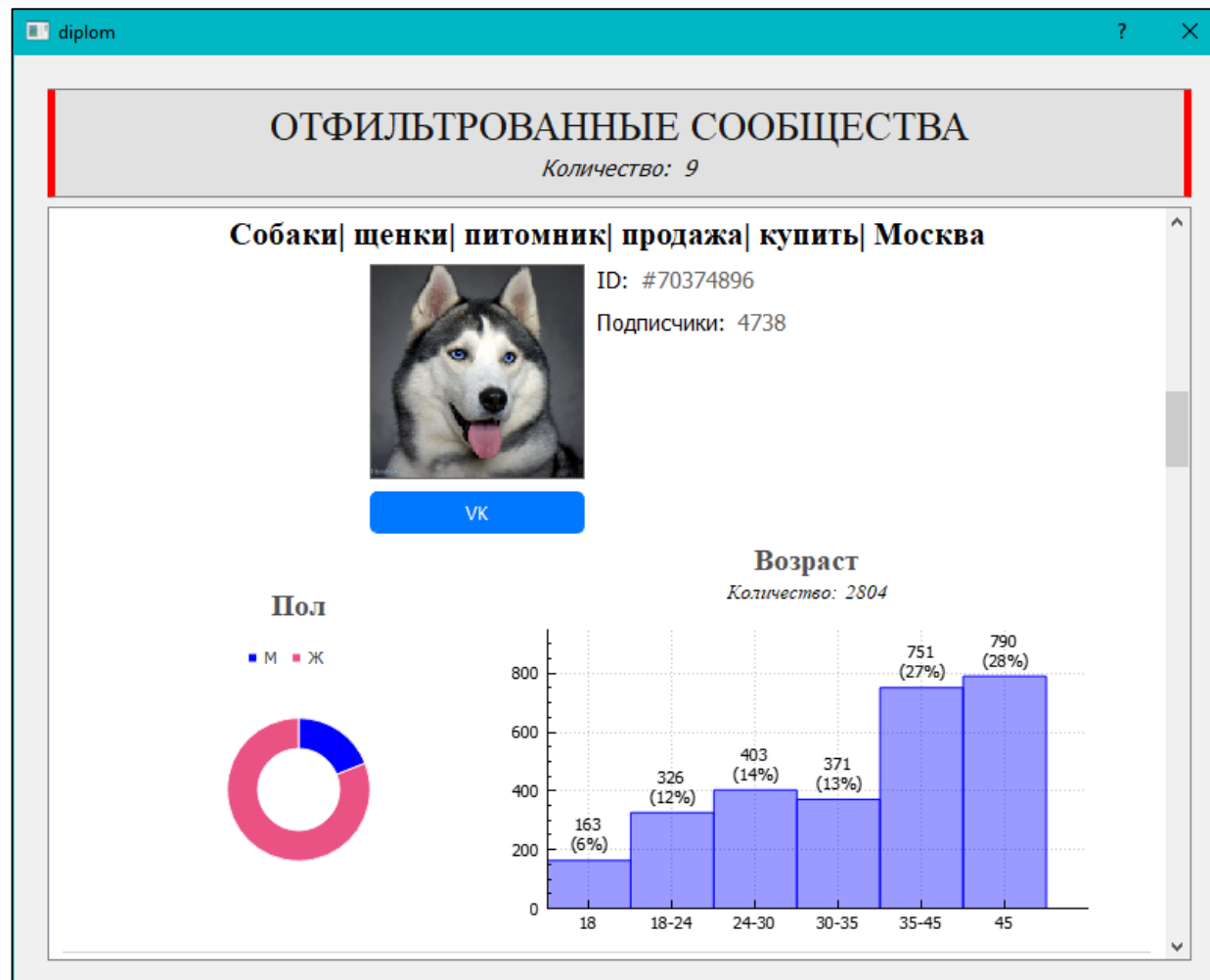
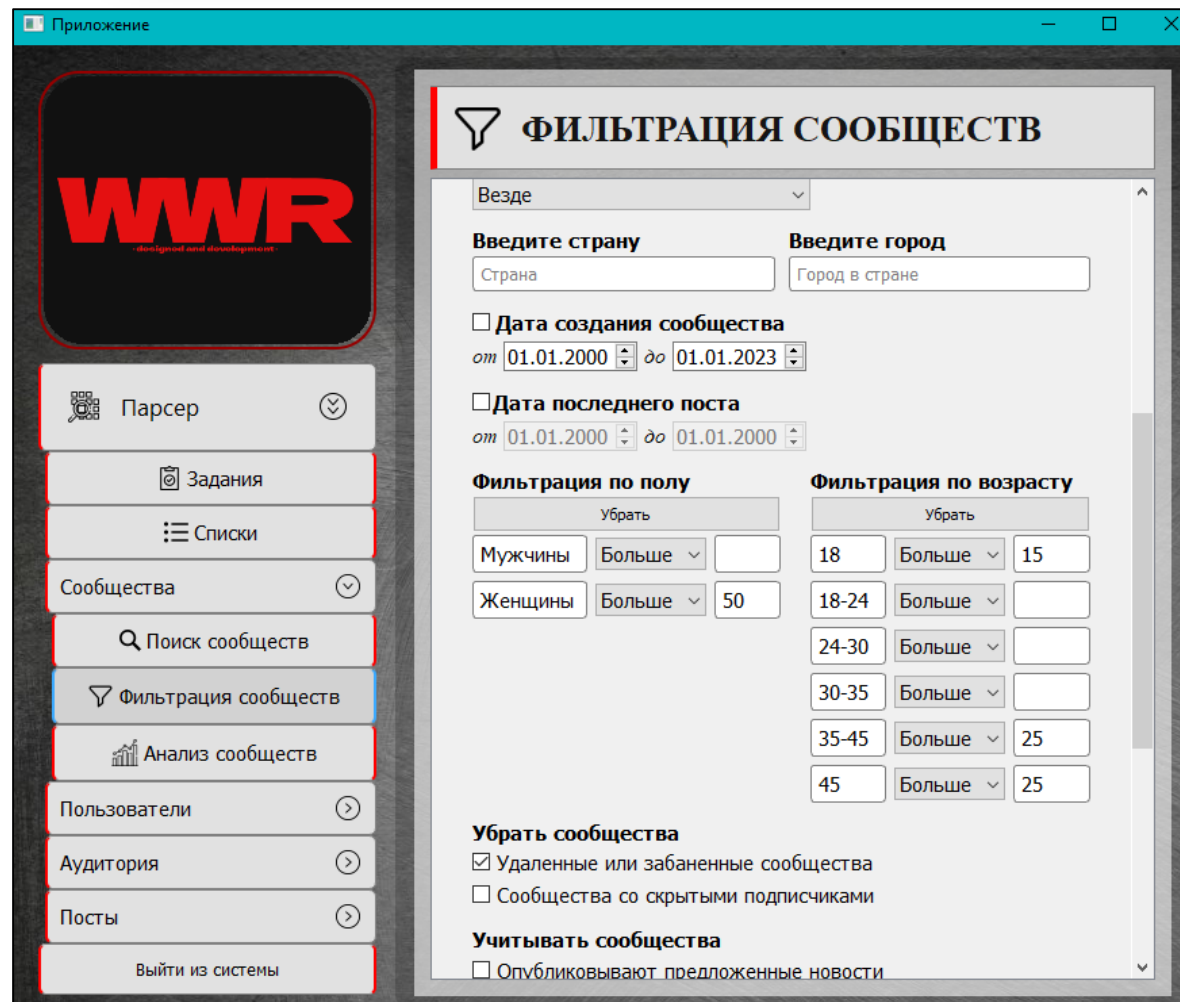
Подписчики: 3298

Группа [vk 116433916](#)

19

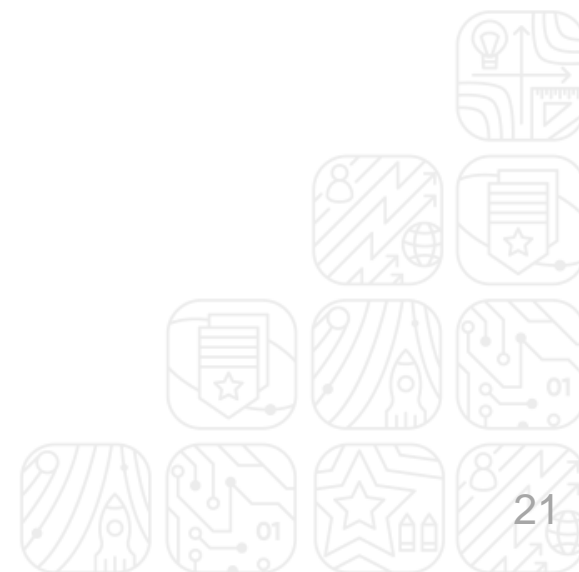
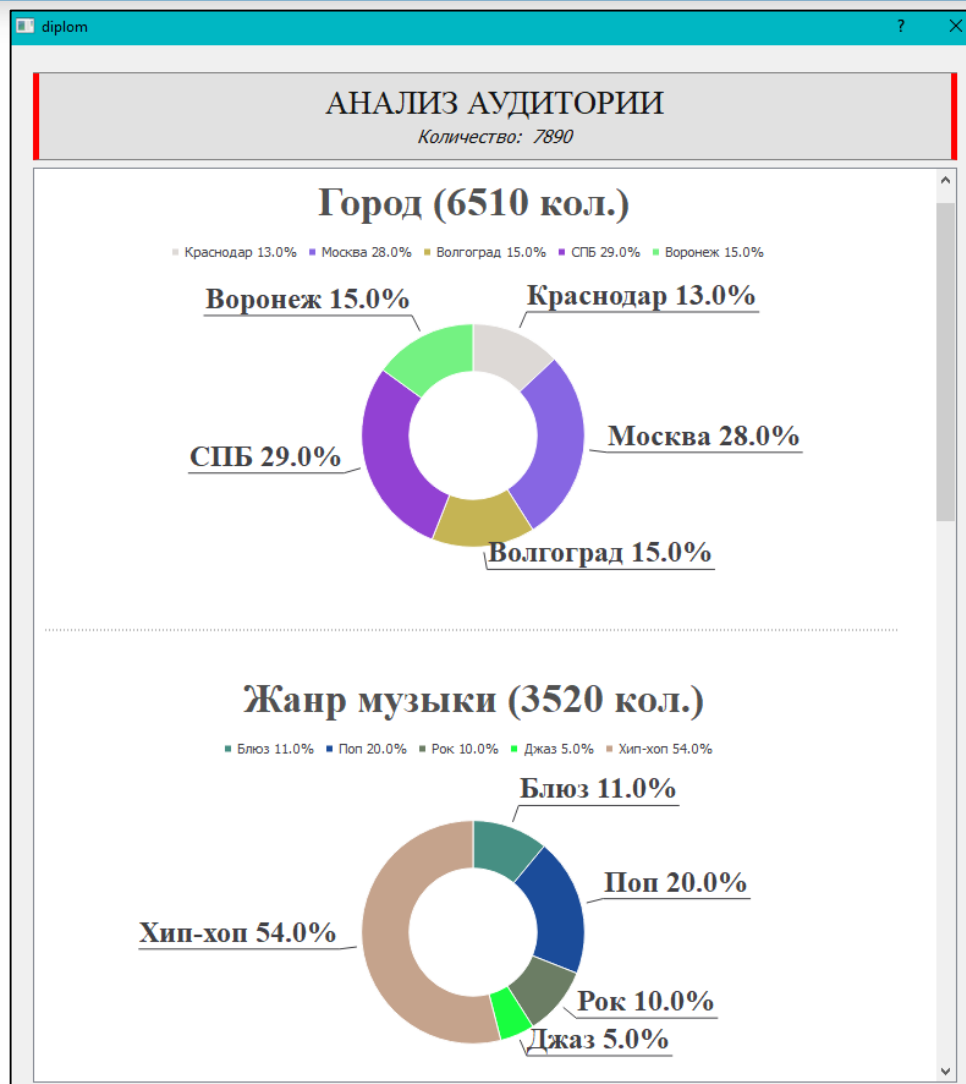


Реализация приложения





Реализация приложения





Функциональное тестирование

Описание	Описание действий	Ожидаемый результат	Выполнение теста
Авторизация через социальную сеть	<ol style="list-style-type: none">1. Открытие окна авторизации2. Нажатие на кнопку «Авторизоваться»3. Разрешение получения доступа к аккаунту4. Ввод кода в окно приложения5. Нажатие кнопки «Войти»	Получение токена доступа от приложения и открытие окна с главным меню	Выполнен
Создание списка	<ol style="list-style-type: none">1. Открытие главного меню2. Нажатие на кнопку «Списки»3. Нажатие на кнопку «Создать список»4. Ввод данных о списке5. Нажатие кнопки «Создать»	Появление созданного списка в списке списков	Выполнен



Функциональное тестирование

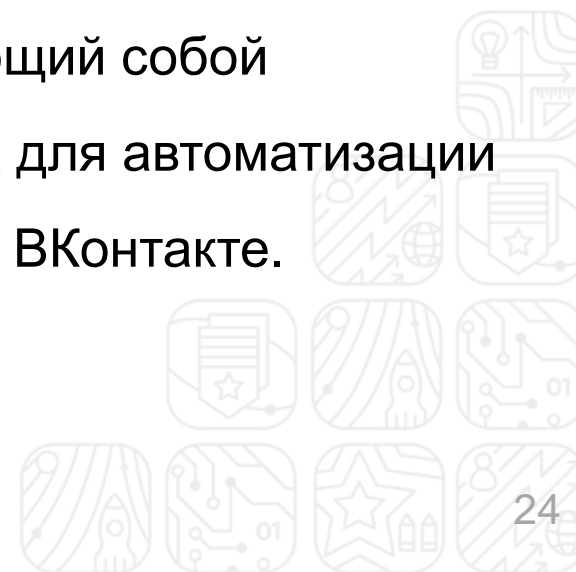
Описание	Описание действий	Ожидаемый результат	Выполнение теста
Создание задачи поиска данных	<ol style="list-style-type: none">1. Открытие главного меню2. Нажатие на кнопку «Парсер»3. Нажатие на кнопку «Поиск сообществ»4. Ввод параметров задачи5. Нажатие на кнопку «Поиск»	Появление созданной задачи в списке задач с последующим ее выполнением	Выполнен
Получение результатов выполненной задачи	<ol style="list-style-type: none">1. Открытие главного меню2. Нажатие на кнопку «Парсер»3. Нажатие на кнопку «Задания»4. Нажатие на кнопку «Результат» напротив необходимой задачи	Появление окна с отображением результата выполненной задачи	Выполнен



Заключение

В ходе выполнения выпускной квалификационной работы был произведен анализ структуры социальной сети ВКонтакте, технологий проектирования API и существующих аналогов, сформированы требования к программному продукту, спроектировано приложение, пользовательский и программный интерфейсы, были выбраны средства реализации, а также произведено тестирование приложения.

В результате выполнения работы было разработано API, представляющий собой инструмент, основанный на принципах и функциональности SMM-парсера для автоматизации сбора, структурирования, анализа и фильтрации данных социальной сети ВКонтакте.





Балтийский государственный
технический университет
«ВОЕНМЕХ»
им. Д.Ф. Устинова

Спасибо за внимание!

