**Заказная техническая документация для IT-бизнеса**

Прокаев Арсений

Ведущий технический писатель

WhatsApp/Telegram +7 (919) 688-46-48

[skyperfect2@mail.ru](mailto:skyperfect2@mail.ru.)

|  |  |  |
| --- | --- | --- |
| УТВЕРЖДАЮ |  | УТВЕРЖДАЮ |
| **Заказчик:**  Должность  **ООО «Автосистем»** |  | **Подрядчик:** |
|  |  |  |
| \_\_\_\_\_\_\_\_\_\_\_\_\_\_ Фамилия И.О. |  | \_**Прокаев А.Д.\_\_** |
| «\_\_\_» \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ 2025 г. |  | «\_07\_» \_\_\_\_\_\_08\_\_\_\_\_\_\_\_\_\_2025 г. |

Веб-сервис формирования субтитров для видео «SubtitleGenius»

Версия 1.0

Дата составления: 07.08.2025

Содержание

[1. Введение 4](#_Toc205424881)

[1.1 Назначение системы 4](#_Toc205424882)

[2. Техническая архитектура 4](#_Toc205424883)

[3. Технологический стек 4](#_Toc205424884)

[3.1 Клиентская часть (Frontend) 4](#_Toc205424885)

[3.2 Серверная часть (Backend) 5](#_Toc205424886)

[3.3 Инфраструктура 6](#_Toc205424887)

[4. Ключевые модули системы 7](#_Toc205424888)

[4.1 Загрузка видео 7](#_Toc205424889)

[4.2 Транскрипция и диаризация 7](#_Toc205424890)

[4.3 Редактор субтитров 7](#_Toc205424891)

[4.4 Экспорт 7](#_Toc205424892)

[5. Особенности реализации 8](#_Toc205424893)

[5.1 Обработка больших видео 8](#_Toc205424894)

[5.2 Безопасность 8](#_Toc205424895)

[5.3 Логирование 8](#_Toc205424896)

[6. Развертывание 9](#_Toc205424897)

2) требования к способам и средствам связи для информационного обмена между компонентами системы;

5) требования по диагностированию системы;

# 1. Введение

## 1.1 Назначение системы

SubtitleGenius — интерактивный веб-сервис для автоматической генерации, редактирования и экспорта субтитров в форматах SRT, VTT, TXT.  
Цели:

* Автоматизация создания субтитров с использованием нейросетевых технологий (Whisper AI).
* Поддержка мультиязычности (русский/английский).
* Обеспечение пользовательского опыта через интуитивный интерфейс.

## 2. Техническая архитектура

Система построена по клиент-серверной схеме:

[Клиент (React)] ↔ [REST API (Flask)] ↔ [Сервисы обработки (Whisper)]

Компоненты:

* Фронтенд: Single Page Application на React.
* Бэкенд: Микросервис на Flask (Python).
* Обработка видео:

1. Транскрибация через Whisper.
2. Диаризация (определение спикеров).

* Хранение данных: Временные файлы на сервере (очистка после обработки).

## 3. Технологический стек

## 3.1 Клиентская часть (Frontend)

* React 18: Ядро интерфейса.
* Redux Toolkit: Управление состоянием.
* CSS-модули: Стилизация компонентов.
* Библиотеки:

1. react-player для воспроизведения видео.
2. axios для запросов к API.

Пример компонента выбора языка представлен на рисунке 1.



Рисунок 1 – Компонент выбора языка

## 3.2 Серверная часть (Backend)

* Python 3.10 + Flask: REST API.
* Celery: Фоновая обработка задач.
* FFmpeg: Извлечение аудио из видео.
* Whisper: Модели base/medium для транскрипции.

Пример функции извлечения аудио с улучшением качества представлен на рисунке 2.



Рисунок 2 – Функция извлечения аудио с улучшением качества

## 3.3 Инфраструктура

* + Docker: Контейнеризация сервисов.
  + Redis: Брокер задач для Celery.
* Nginx: Реверс-прокси и балансировщик.

# 4. Ключевые модули системы

## 4.1 Загрузка видео

* Поддерживаемые форматы: MP4, AVI, MOV, MKV.
* Макс. размер: 5 ГБ.
* Валидация файлов через сигнатуры (бинарные заголовки).

## 4.2 Транскрипция и диаризация

* Whisper: Определение языка, транскрипция в текст с таймкодами.
* Диаризация: Разделение текста по спикерам через алгоритм кластеризации.

Логика обработки:

1. Извлечение аудио с частотой 16 kHz (оптимально для Whisper).
2. Сегментация длинных видео (>10 мин) для экономии памяти.
3. Перевод субтитров через интеграцию с Yandex Translate API (опционально).

## 4.3 Редактор субтитров

* Интерфейс: Синхронизированная временная шкала видео и текста.
* Возможности:
  1. Ручная корректировка текста.
  2. Сдвиг таймкодов.
  3. Объединение/разделение сегментов.

## 4.4 Экспорт

* Форматы: SRT (стандарт), VTT (для веба), TXT (чистый текст).
* Кодировка: UTF-8.

## 5. Особенности реализации

## 5.1 Обработка больших видео

* Видео >1 ГБ обрабатываются частями через Celery-задачи.
* Прогресс отображается клиенту через WebSocket.

## 5.2 Безопасность

* JWT-аутентификация для API.
* Изоляция пользовательских данных через временные каталоги.
* Проверка файлов на наличие вредоносного кода (ClamAV).

## 5.3 Логирование

Пример расширенной настройки логирования представлен на рисунке 3.

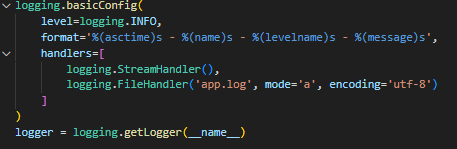


Рисунок 3 – Расширенная настройка логирования

## 6. Развертывание

Требования к серверу:

* CPU: 4+ ядра (рекомендуется x86\_64).
* RAM: 8+ ГБ.
* Диск: 50+ ГБ (SSD/NVMe).