ТЕХНИЧЕСКОЕ ЗАДАНИЕ

По Курсовому проекту:

«Сервис для транскрибации лекций с глоссарием терминов»

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  | | |  |  |  |  |
|  | | |  |  |  |  |
|  | | |  |  |  |  |
|  | | |  |  |  |  |
|  | | |  |  |  |  |
|  | | |  |  |  |  |
| Разработчики:  Розов С.А.  Александров С.В | | |  |  |  |  |
|  |  |
| **Утверждено:**  Начальник управления … ООО «…» | | |  | \_\_\_\_\_\_\_\_\_\_ |  | …………….. |
| (должность полностью) | | |  | (подпись) |  | (Ф.И.О) |

**СОДЕРЖАНИЕ**

[1.](#_gjdgxs) Общие сведения 3

[1.1.](#_30j0zll) Наименование проекта 3

[1.2.](#_1fob9te) Список принятых сокращений 3

[2.](#_3znysh7) Назначения, цели и сроки проекта 3

[2.1.](#_2et92p0) Назначение проекта 3

[2.2.](#_tyjcwt) Цели проекта 3

[2.3.](#_3dy6vkm) Задачи проекта 4

[2.3.1.](#_1t3h5sf) Результаты проекта 4

[2.4.](#_4d34og8) Требования к промежуточной отчетности 5

[2.5.](#_2s8eyo1) Сроки проекта 5

[3.](#_17dp8vu) Требования к документированию 5

# **Общие сведения**

## **Наименование проекта**

Курсовой Проект по теме "Сервис для транскрибации лекций с глоссарием терминов"

## **Список принятых сокращений**

|  |  |
| --- | --- |
| Термин | Расшифровка/Определение |
| Компания | Группа компаний … |
| Заказчик | ООО «…» |
| Транскрибация | Процесс преобразования устной речи (аудио/видео) в текстовый формат |
| Глоссарий | Словарь специализированных терминов с их определениями |
| Шумоподавление | Очистка аудио от фоновых шумов для улучшения качества распознавания |
| Конфиденциальность | Обеспечение защиты данных пользователей и их неразглашение |

# **Назначения, цели и сроки проекта**

## **Назначение проекта**

Автоматическая расшифровка лекций с выделением сложных терминов, формированием структурированного конспекта и глоссария для удобства обучения и преподавания.

## **Цели проекта**

| Цель проекта | Метрика |
| --- | --- |
| * Обеспечить точность транскрибации ≥ 90% для академического контента. * Автоматически определять и выносить в глоссарий ключевые термины. * Поддерживать экспорт материалов в форматы PDF, DOCX и Markdown. | 1. Скорость обработки: ≤ 1 минута на 10 минут аудио. 2. Точность распознавания терминов: ≥ 85%. 3. Возможность ручной корректировки текста и глоссария. |

## **Задачи проекта**

В рамках проекта требуется выполнить следующие задачи:

**2.3.1. Командой проекта:**

* Сбор и обработка данных;
* Идентификация проблемных зон;
* Идентификация причин возникновения заторов;
* Подготовка рекомендаций по устранению проблем в каждой точке перегрузки. Оценка стоимости каждой рекомендации. Оценка объема устранения заторов.

**2.3.2. Внутренним исполнителем:**

* Предоставление доступа к данным с муниципальных датчиков и систем видеонаблюдения
* Экспертный анализ наработок команды;
* Обзор и сравнительный анализ существующих opensource алгоритмов преобразования по разработанным критериям;
* Реализация и адаптация наилучшего из существующих алгоритмов, либо разработка собственного, в случае отсутствия подходящих по критериям;
* Тестирование алгоритма на данных Заказчика, до этого никак не задействованных в создании алгоритма;
* Подготовка рекомендаций по дальнейшей доработке алгоритма;
* Подготовка итоговых материалов по проекту (отчет, презентация).

## **Результаты проекта**

Результаты проекта:

1. Программный код алгоритма

##### 2.1. Результаты анализа и предобработки данных

* Датасет:
  + Собраны аудиозаписи лекций (5–20 часов) из разных областей (технические, гуманитарные).
  + Текстовые расшифровки для обучения/тестирования (ручная разметка части данных).
* Предобработка:
  + Шумоподавление (LibROSA, PyAudioFilter).
  + Сегментация аудио по тишине (VAD – Voice Activity Detection).
  + Нормализация текста (удаление стоп-слов, лемматизация – SpaCy, NLTK).

##### 2.2. Метрики качества работы алгоритмов

* Транскрибация:
  + WER (Word Error Rate): 12% (Whisper-large) против 18% (Wav2Vec 2.0).
  + Скорость обработки: 1x (реальное время) для Whisper-small, 0.5x для Whisper-large.
* Глоссарий:
  + Точность извлечения терминов (F1-score): 0.78 (KeyBERT + ручная постобработка).
  + Полнота определений: 65% (автоматически) → 85% с доработкой эвристик.

##### 2.4. Границы применимости

* Шумы в аудио → рост WER до 25%.
* Специфические термины (медицина, юриспруденция) → требуют кастомного словаря.
* Длинные лекции (>1 часа) → необходимость сегментации.

##### 2.5. Рекомендации по доработке

* Добавить адаптацию модели под домен (fine-tuning на учебных лекциях).
* Улучшить постобработку терминов (контекстный анализ для глоссария).
* Оптимизировать скорость (квантование модели, GPU-ускорение).

##### 2.6. Бэклог на следующий этап

1. Интеграция с облачными API (Google Speech-to-Text для сравнения).
2. Разработка веб-интерфейса (Flask/Streamlit).
3. Добавление поддержки форматов видео (извлечение аудио).
4. Механизм исправления ошибок пользователем (crowdsourcing).

##### 2.7. Демо-ролик работы алгоритма

* Видео (3–5 мин) с примерами:
  + Загрузка аудио → транскрибация → выделение терминов → глоссарий.
  + Сравнение ручной и автоматической расшифровки.

1. Заключения экспертов от бизнеса;
2. Презентация с выводами по итогу проекта, рекомендациями и заключением по дальнейшему развитию проекта.

## **Требования к промежуточной отчетности**

В рамках проекта осуществляется взаимодействие между представителями Заказчика и Исполнителя путем оговоренной еженедельной конференцсвязи для демонстрации результатов еженедельного спринта в формате презентации. Вопросы, возникающие в ходе реализации проекта, уточняются между указанными лицами по мере возникновения.

## **Сроки проекта**

Результаты по проекту необходимо передать и согласовать с Заказчиком в срок до **03.06.2025 г**.

# **Требования к документированию**

По результатам реализации НИОКР-проекта необходимо предоставить:

1. Отчет, в формате .*doсx*;
2. Демо-ролик, в формате .*avi*;
3. Презентация об итогах реализации проекта, в формате .*pptx*.