

БЕЛОРУССКИЙ НАЦИОНАЛЬНЫЙ ТЕХНИЧЕСКИЙ УНИВЕРСИТЕТ
Факультет информационных технологий и робототехники
Кафедра «Программное обеспечение информационных систем и технологий»

ДОПУЩЕН К ЗАЩИТЕ

Старший преподаватель

Н.В. Воюш
(подпись)

Н.В. Воюш
(инициалы и фамилия)

« 03 » 06 2026 г.

РАСЧЕТНО-ПОЯСНИТЕЛЬНАЯ ЗАПИСКА
ДИПЛОМНОГО ПРОЕКТА

«Приложение для транскрибации видео и создания субтитров»

Специальность 1-40 05 01 «Информационные системы и технологии (по направлениям)»

Направление специальности 1-40 05 01-01 «Информационные системы и технологии (в проектировании и производстве)»

Обучающийся

группы 10702322
(номер)

Е.И. Белуш
(подпись, дата)

Руководитель

А.Н. Марков
(подпись, дата)

Консультанты:

по разделу «Компьютерное проектирование»

А.Н. Марков
(подпись, дата)

по разделу «Охрана труда»

М.Л. Калиниченко
(подпись, дата)

по разделу «Экономика»

Е.А. Хвитько
(подпись, дата)

Ответственный за нормоконтроль

Н.В. Романюк
(подпись, дата)

Объем проекта:

расчетно-пояснительная записка — 81 страниц;

графическая часть — 10 листов;

магнитные (цифровые) носители — 1 единиц.

Минск 2026

РЕФЕРАТ

ПРИЛОЖЕНИЕ ДЛЯ ТРАНСКРИБАЦИИ ВИДЕО, СОЗДАНИЕ СУБТИТРОВ, ЛОКАЛЬНАЯ ОБРАБОТКА, РАСПОЗНАВАНИЕ РЕЧИ, КОРРЕКЦИЯ ТЕКСТА, ПЕРЕВОД СУБТИТРОВ

Объектом исследования является процесс автоматической транскрибации видеофайлов и формирования субтитров в полностью локальной вычислительной среде.

Цель проекта – разработка приложения для транскрибации видео и создания субтитров, обеспечивающего полный цикл обработки в локальном режиме без обращения к облачным сервисам.

В процессе проектирования выполнены следующие исследования (или разработки):

- 1) Спроектирована модульная архитектура приложения;
- 2) Интегрирована модель распознавания речи;
- 3) Реализован модуль коррекции текста на базе языковой модели с настраиваемыми режимами обработки;
- 4) Проведено функциональное тестирование системы.

Элементами научной новизны (практической значимости) полученных результатов являются локальная архитектура, возможность работы без подключения к интернету, применение современных моделей машинного обучения для достижения высокого качества субтитров.

Областью возможного практического применения являются организации, работающие с конфиденциальными видеоматериалами: юридические фирмы, банки, государственные учреждения и образовательные организации.

В ходе дипломного проектирования прошли апробацию такие предложения, как совместное применение моделей распознавания речи и обработки текста.

Результатами внедрения могут стать сокращение времени на создание субтитров, обеспечение конфиденциальности обрабатываемых материалов и независимость от интернет-соединения.

Студент-дипломник подтверждает, что приведенный в дипломном проекте и расчетно-аналитический материал объективно отражает состояние разрабатываемого объекта, все заимствованные из литературных и других источников теоретические и методологические положения сопровождаются ссылками на их авторов.

Дипломный проект: 81 с., 30 рис., 8 табл., 18 источников, 1 прил.

СПИСОК ИСПОЛЬЗОВАННЫХ ИСТОЧНИКОВ

- 1 Прохоренок, Н. А. PyQt. Создание оконных приложений на Python / Н. А. Прохоренок, В. А. Дронов. – СПб.: БХВ-Петербург, 2021. – 544 с.
- 2 Саммерфилд, М. Python на практике / М. Саммерфилд. – М.: ДМК Пресс, 2015. – 338 с.
- 3 Шифман, П. А. Сравнительный анализ систем автоматического распознавания речи / П. А. Шифман, А. В. Кузнецов // Программные системы: теория и приложения. – 2021. – № 3. – С. 45–58.
- 4 Иванов, А. В. Локальное развёртывание больших языковых моделей: инструменты и практические аспекты / А. В. Иванов, Д. С. Петров // Программные продукты и системы. – 2024. – № 2. – С. 33–40.
- 5 Qt Documentation [Электронный ресурс]. – Режим доступа: <https://doc.qt.io/qt-6/>.
- 6 FFmpeg Documentation [Электронный ресурс]. – Режим доступа: <https://ffmpeg.org/documentation.html>.
- 7 Tkinter Documentation [Электронный ресурс]. – Режим доступа: <https://docs.python.org/3/library/tkinter.html>
- 8 PyTorch Documentation [Электронный ресурс]. – Режим доступа: <https://pytorch.org/docs/stable/>.
- 9 OpenAI Whisper [Электронный ресурс]. – Режим доступа: <https://openai.com/research/whisper>.
- 10 Python Package Index (PyPI) [Электронный ресурс]. – Режим доступа: <https://pypi.org/>.
- 11 Librosa Documentation [Электронный ресурс]. – Режим доступа: <https://librosa.org/doc/latest/>
- 12 Kaldi ASR [Электронный ресурс]. – Режим доступа: <https://kaldi-asr.org/>
- 13 Молодяков, С. А. Применение функций FFmpeg в мультимедийных приложениях (100 примеров на Python) : монография / С. А. Молодяков. – СПб.: ПОЛИТЕХ-ПРЕСС, 2023. – 512 с.
- 14 Михнин, А. Искусство общения с AI: Мастерство создания инструкций для языковых моделей / Алексей Михнин. – [Б. м.] : Издательские решения, 2024. – 333 с.
- 15 Немачук, Е. П. Использование нейронной сети Whisper для распознавания и конспектирования речи / Е. П. Немачук, В. В. Доманский // Цифровые инфокоммуникационные технологии : сборник научных трудов. – Ростов-на-Дону : РГУПС, 2023. – С. 75–79.
- 16 Лазаренков, А. М. Охрана труда и пожарная безопасность: учебное пособие / А. М. Лазаренков, Ю.Н. Фасевич. – Минск: ИВЦ Минфина, 2025. – 636 с.

17 Лазаренков А.М. Охрана труда. Учебно-практическое пособие по расчетам в охране труда: электронное пособие / А.М. Лазаренков, Т.П. Кот, Е.В. Мордик, Л.П. Филянович. – Минск: Регистр. номер БНТУ/МТФ 35-42.2018. Зарегистрировано 04.05.2018. – 11,7 усл.эл.л.

18 Вершина Г.А. Охрана труда: учебник / Г.А. Вершина, А. М. Лазаренков, Мусаев М.Н. – Минск: ИВЦ Минфина, 2022. – 584 с.