Расшифровка аудио в текст

Назначение
Процесс обработки
Ключевой результат
Техническая архитектура
Процесс получения и настройки доступа
1. Использование в интерфейсе АБИС
2. Использование через API

Пример результата

Сервис автоматизированного преобразования речи в текст

> Интеллектуальный сервис, который автоматически преобразует аудиофайл в текстовый документ для чтения, поиска и анализа.

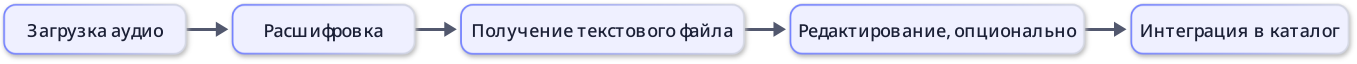
Назначение

Автоматическое создание текстовых расшифровок (транскриптов) лекций, подкастов, интервью, устных историй, конференций и видеоуроков для обеспечения полнотекстового поиска, цитирования и доступности.

- Сделать слышимое — читаемым и находимым: включить содержание аудиоматериалов в единый поисковый индекс.
- Поддержать исследования: предоставить ученым, студентам и журналистам точный текстовый материал для анализа.
- Обеспечить инклюзивность: помочь слабослышащим пользователям получить доступ к аудиоконтенту.
- Сохранить наследие: перевести в цифровой текстовый формат уникальные архивные аудиозаписи.

Повышает доступность и удобство работы с аудиовизуальными материалами библиотечного фонда, архивными записями и современным мультимедийным контентом. Открывает новые возможности для исследований, обучения и навигации по фондам, экономя время пользователей и сотрудников. Способствует сохранению и активному использованию устного наследия.

Процесс обработки



1. Ввод данных
   - Прием на вход аудио- или видеофайла в распространенных форматах (MP3, WAV, MP4, AVI и др.).
   - Проверка файла на корректность, качество звука и возможность обработки.
2. Обработка и анализ
   - Распознавание речи
     - Использование нейросетевых моделей для высокоточного распознавания русской речи, включая специальную терминологию.
     - Автоматическое определение и разделение речи разных дикторов.
   - Форматирование текста
     - Расстановка знаков препинания, заглавных букв и абзацев на основе пауз и смысловых блоков.
3. Формирование структурированного документа, готового к использованию.
   - Вывод результата
     - Формирование текстового файла
       Пользователю предоставляется готовый текстовый файл в формате (TXT с разметкой MARKDOWN).
   - Доступ к результату
     Текст доступен для скачивания, онлайн-просмотра и копирования. Возможна интеграция для отображения транскрипта в электронном каталоге.
4. Интеграция в каталог и поиск
   - Текстовая расшифровка индексируется поисковой системой платформы, позволяя находить аудио- и видеоматериалы по словам, произнесенным внутри них.
   - Ссылка на транскрипт или сам текст включается в соответствующее поле библиографической записи (856 RUSMARC).

Ключевой результат

Создание полнотекстового слоя для аудиовизуального фонда, что радикально повышает его возможности обнаружения, повышает исследовательский потенциал и доступность.

Техническая архитектура

Сервис размещен на защищенных вычислительных мощностях облачной платформы "Библиотека технологий". Обработка записей происходит на удаленных серверах, что гарантирует высокую производительность и отсутствие нагрузки

на локальную инфраструктуру библиотеки. От пользователя требуется только стабильное интернет-соединение.

Процесс получения и настройки доступа

Сервис расшифровки текстов является частью платформы "Библиотека технологий" и не требует локальной установки в библиотеке. Функциональность доступна удалённо через два интерфейса.

1. Использование в интерфейсе АБИС

> Для сотрудников

В состав платформы входит АБИС - Автоматизированная Библиотечная Информационная Система. Использование в АБИС самый простой способ для повседневной работы сотрудников. Пользователь выбирает аудиофайл включенный в каталог и нажимает кнопку обработки. В результате будет сформирован и загружен текстовый файл, внесены изменения в 856 поле.

2. Использование через API

> Для интеграции и автоматизации

Этот способ предназначен для разработчиков и администраторов, желающих встроить обработку в собственные скрипты или рабочие процессы.

Получение доступа к API Платформы "Библиотека технологий"

Пример результата

- Архивная запись выступления на радио
- Лекция