

API ДЛЯ РАБОТЫ С ПОДКАСТАМИ

Лысяков И.С., студент

Белорусский государственный университет информатики и радиоэлектроники
г. Минск, Республика Беларусь

Куликов С.С. – канд. техн. наук, доцент

В современную эпоху развития цифровых технологий появились различные способы получения информации. Пользователи всё чаще выбирают цифровые платформы для потребления образовательного, развлекательного и новостного контента. Одним из активно развивающихся форматов являются подкасты, позволяющие получать различную информацию в удобной аудиоформе и сочетать её с повседневной деятельностью. Сам рынок подкастинга демонстрирует устойчивый рост [1]. Потому была поставлена цель создать веб-приложение для работы с подкастами, где можно будет настраивать всё под себя, удобно управлять контентом и использовать новые технологии.

Основной задачей приложения является создание качественного сервиса, охватывающего как можно большее количество пользователей, от рядовых слушателей до меломанов и людей с ограничениями слуха. Основной функционал, обеспечивающий успех приложения, представляет следующее:

- 1) поиск и открытие новых шоу;
- 2) поиск по названиям; возможность подписки на обновления (автоматическая загрузка новых эпизодов);
- 3) управление скоростью воспроизведения и таймер сна;
- 4) создание плейлистов и очереди воспроизведения;
- 5) синхронизация прогресса прослушивания между устройствами.

Однако в условиях высокой конкуренции и наличия достаточно большого количества готовых решений, этот функционал является необходимым минимумом для вхождения на рынок. Для того, чтобы приложение выделялось на фоне конкурентов и предоставляло уникальный опыт взаимодействия с ним, нужно разработать новые и уникальные решения.

В связи с вышеописанным, было принято решение, что приложение должно:

- 1) проводить транскрибацию эпизодов (перевод из аудио в текстовый формат);
- 2) формировать краткий пересказ сути эпизодов;
- 3) перевод содержимого эпизодов на иностранные языки.

Все три пункта подразумевают собой использование современных технологий – искусственного интеллекта. Этот подход полностью соответствует глобальным трендам развития аудиоформатов, где искусственный интеллект и мультязычность становятся ключевыми драйверами [2].

Транскрибация осуществляется следующим образом: исходный аудиофрагмент разбивается на короткие сегменты продолжительностью 5–10 секунд с перекрытием. Каждый сегмент подвергается обработке моделью распознавания речи. На следующем этапе проводится постобработка, включающая удаление слов-паразитов, корректировку терминологии с использованием специализированного словаря и сегментацию текста на предложения по паузам. В случае участия нескольких спикеров применяется диаризация, которая позволяет отделять голоса по кластерам. В результате получается структурированный текст с таймкодами и отметками о говорящих.

Формирование краткого пересказа происходит путем исключения из транскрибированного текста стандартных служебных фрагментов, таких как приветствия и рекламные вставки. Оставшийся материал делится на смысловые блоки, после чего модель экстрактивной суммаризации оценивает предложения на основе частоты ключевых слов, их позиции и семантической релевантности к описанию выпуска. Отобранные предложения передаются генеративной модели, которая создаёт связный реферат с логическими переходами. Перевод на иностранные языки реализуется через разбиение транскрибированного текста на предложения с сохранением таймкодирования. Имена, названия и сленговые выражения помещаются в особый словарь исключений. Основной текст переводится с помощью нейросетевой модели, после чего выполняется постредактирование: унификация терминов, адаптация идиоматических выражений и перестановка фраз в соответствии с синтаксисом целевого языка. Итоговый перевод выводится синхронно с аудио в формате субтитров и доступен для экспорта. Таким образом, разрабатываемое веб-приложение для работы с подкастами объединяет базовый функционал с тремя инновационными AI-решениями. Это позволяет не только удовлетворить потребности широкой аудитории слушателей, но и обеспечить доступность для людей с нарушениями слуха, а также предложить уникальный опыт взаимодействия, выделяющий сервис среди аналогичных продуктов на рынке.

Список использованных источников:

1. Fortune Business Insights. Podcasting Market Size, Share & Industry Analysis, 2026-2034. – URL: <https://www.fortunebusinessinsights.com/podcasting-market-107055>
2. Шумилова К. Тренды подкастов: что ждет аудиоформаты в ближайший год // РБК Тренды. – 2025. – 30 сент. – URL: <https://trends.rbc.ru/trends/social/6842b1ac9a79478e09233558>