Project Proposal: AI-Powered Interactive Reading Software

Tim Xu

Abstract

This proposal aims to develop an AI-powered reading application designed for an interactive and immersive reading experience. The software will read eBooks aloud and respond to user voice commands, providing translations, definitions, and content summaries. Utilizing text-to-speech, voice recognition, and AI integration, the goal is to enhance reading accessibility and support for non-native speakers, as well as multitaskers. This project will combine my expertise in software development with my passion for audio technology.

Introduction

The concept for this project emerged from personal experiences as a non-native English speaker struggling to maintain reading continuity, especially while multitasking. The app will bridge the current gap in eBook technology by providing an intelligent, voice-interactive solution that reads content aloud and answers real-time user queries. The primary objective is to create a seamless, user-friendly reading platform that encourages deeper content engagement.

Background

Existing eBook formats, such as EPUB, offer static text displays and limited interaction, which can disrupt reading flow when users need translations, definitions, or additional context. Current text-to-speech tools, while capable of reading text aloud, are restricted to basic functionality and

do not allow users to engage with the content in meaningful ways, such as asking questions or receiving summaries.

This project aims to address these limitations by incorporating advanced technologies, including text-to-speech APIs for natural-sounding narration, speech recognition for seamless voice command processing, and AI-driven content analysis for dynamic responses. By combining these tools, the software will transform traditional reading into an interactive experience, where users can effortlessly access definitions, translations, and contextual insights without interrupting their engagement with the content.

Drawing inspiration from platforms like audiobooks, which emphasize audio accessibility, and smart assistants like Alexa, which excel in interactivity, this project integrates the best of both worlds. However, it focuses uniquely on enhancing the reading experience, making it an innovative solution for those who want deeper engagement with their texts.

Proposed Work

The project will be developed in several phases:

1. **Book Processing**
    - Convert multiple formats (PDF, EPUB, TXT) into a standard text format.
    - Integrate a text-to-speech API for read-aloud functionality.
2. **Voice Recognition**
    - Capture and process voice commands using a speech recognition API.
3. **AI Interaction**

- Implement AI processing for intelligent responses to user queries.

4. **User Interface Development**

   - Design a cross-platform mobile app using React Native.

5. **Testing and Refinement**

   - Ensure smooth interaction between the system components and user experience.

Specific Tasks

- Develop a backend system for file handling and conversion.
- Implement speech-to-text capabilities for voice command processing.
- Integrate an AI API for generating contextual answers.
- Build a user-friendly UI with multitasking-optimized features.

Rationale

This interactivity ensures that reading is no longer a linear, passive activity but an adaptable and immersive experience tailored to individual needs. The project's emphasis on accessibility also aligns with the growing demand for tools that support diverse learning styles, including auditory and interactive learners.

Additionally, by combining cutting-edge AI technologies with user-focused design, this project has the potential to benefit a wide audience, including:

- Busy professionals who want to stay informed while multitasking.

- Students and researchers seeking quick access to supplementary content like summaries or definitions.
- People with disabilities or visual impairments who require enhanced reading support.

Timeline (Dec 2024 – Apr 2025)

- **Dec 2024**: Complete initial research and finalize APIs used in the project.
- **Jan 2025**: Develop backend for file conversion and integrate text-to-speech.
- **Feb 2025**: Implement voice recognition and AI response functionalities.
- **Mar 2025**: Develop the user interface and connect all components.
- **Apr 2025**: Conduct testing, debugging, and prepare for the final presentation.