TED UNIVERSITY

CMPE 491 / SENG 491 Senior Project

<<BooTunes>>

Project Specifications Report

Fall 2024

Team Members:

Mehlika Eroğlu, 58717180232, Software Engineering

Melih Aydın, 28180576276, Computer Engineering

Elif Alptekin, 31376233198, Computer Engineering

Batuhan Dalkılıç, 11491297756, Computer Engineering

Supervisor: Venera Adanova

Jury Members:

Gökçe Nur Yılmaz

Eren Ulu

Firat AKBA

1. Introduction

1.1 Description

Our project BooTunes aims to make the reading experience unique for users by combining text-based sentiment analysis with personalized music and artificial intelligence-generated visuals. The system will be developed as a web and mobile application that allows users to upload books in PDF format and access pre-loaded books. As the user reads the book, the app will analyze the emotional tone of each page using natural language processing (NLP) techniques and play relevant music from predefined genres that aligns with the perceived emotion. Additionally, the project will include AI-based text-to-image models to create visualizations of each page, enriching the reading process with engaging visuals.

The main goal is to enable readers to emotionally connect with the content and visualize the content creatively by creating a dynamic and immersive environment that focuses readers' attention while reading a book. This project solves the problem of monotonous reading by combining audio, text and images, transforming static text into interactive and sensory-rich experiences.

This project will also promote accessibility and cultural interaction by offering a personalized approach to reading for young audiences who prefer engaging experiences during their reading experience. The goal is to reduce readers' cognitive fatigue, increase engagement, and strengthen the emotional connection with the content.

1.2 Constraints

- 1.2.1. Economic Constraints: The project requires access to music libraries, AI rendering tools, and software licenses, which may impose financial constraints depending on available resources. In addition, project's backend, including databases and storage for user-uploaded PDFs and AI-generated visuals, requires reliable servers.
- 1.2.2. Social Constraints: BooTunes can enhance social interactions by providing a new way to enjoy and discuss books, especially among readers interested in multimedia integration. Additionally, images and music recommendations generated by AI must be compatible with users' cultural expectations and avoid content that may be perceived as offensive or inappropriate.
- 1.2.3. Environmental Constraints: The digital nature of this project minimizes environmental impact by reducing the need for physical materials. However, if scaled, the energy consumption of AI processing can be quite high.
- 1.2.4. Ethical Constraints: The ethical use of content created by artificial intelligence and the control of these uses, music rights and responsible data use as a result of the recording of personal data must be strictly adhered to.
- 1.2.5. Manufacturability Constraints: The project primarily revolves around software development. Its manufacturability depends on cloud computing resources, AI libraries, and device compatibility.
- 1.2.6. Political Constraints: There are minimal political constraints that may set for this

project, but content selection for visual and musical components must comply with copyright and cultural sensitivity regulations.

- 1.2.7. Health and Safety Constraints: Prolonged use of multimedia features (dynamic music and visuals) may cause problems such as eye strain or overexposure and requires user-friendly design with adjustable settings.
- 1.2.8. Sustainability Constraints: BooTunes offers long-term sustainability with its digital-only approach; however, ongoing server maintenance for AI processing may impact energy usage.

1.3 Professional and Ethical Issues

The project must adhere to established codes of ethics in software development. Compliance with the ACM and IEEE Code of Ethics ensures that all software developed prioritizes user privacy, transparency, and accessibility. We will take extra care with the privacy and security of user data as the system will collect personal information such as reading habits and emotional responses and must comply with data protection laws such as GDPR. Ethical rules regarding artificial intelligence must be followed, including responsible use and minimizing bias in emotional recognition algorithms. We must ensure transparency about how AI models are trained and actively work to minimize such biases. Additionally, copyright laws must be strictly adhered to for music libraries and rendering tools. Finally, our group will uphold the principles outlined in the ACM Code of Ethics and the IEEE Code of Ethics by maintaining integrity, fairness, and respect for users throughout the lifecycle of the project.

2. Requirements

2.1. Functional Requirements

- PDF Upload and Processing: Users must be able to upload PDF files through the application. The project must recognize and process the text of each page in real time for analysis.
- Emotion Detection System: The text on each page will be analyzed to determine the dominant emotion using an AI model. The results must be accurate and dynamically updated as the user moves to the next page.
- Music Recommendation: The system will recommend and play music based on perceived emotions using a music API. Music selection should actively adapt to the emotional changes on each page, and music selection should be made according to the prevailing emotion of the page. More than one piece of music should not be played throughout the page in order not to disrupt the reader's focus.
- Visualizing Text: The application will use the necessary tools to create visuals based on the page content. Images should fit the theme of the page and increase user engagement. In addition, necessary tools will be provided to analyze text and visual harmony.
- User Interface (UI/UX): The front-end built with React will seamlessly display PDFs, audio players, and AI-generated visuals. Users should be able to pause, skip, or replay music and navigate pages easily. Necessary licenses and development environments will be arranged for this user-friendly designed platform.

- Backend and Database: The backend will manage API calls, store user preferences, and process data securely. The database to be created will store user-uploaded files and

logs for analytics.

- User Feedback Mechanism: We will be developing a provisioning method in addition to system validation with a feature that allows users to rate music/visuals and provide feedback on the overall experience.

2.2. Non-Functional Requirements

- Performance: One of our goals is to provide minimum delay between emotion detection and visual rendering to ensure seamless system interaction.
- Security and Privacy: Protecting user data, including reading behavior and preferences, under GDPR and other privacy standards is our priority.
- Compatibility: We will build the system to work on both web and mobile platforms, supporting a variety of screen sizes and operating systems.
- Accessibility: The interface should be suitable for a wide range of users, including dark mode.
- Maintenance and Updates: The project should allow for future improvements and AI model updates without significant disruption.

3. References

1) ACM Code of Ethics and Professional Conduct - https://www.acm.org/code-of-ethics
2) The Software Engineering Code of Ethics, IEEE Computer Society - https://www.computer.org/education/code-of-ethics
3) DALL-E information and API usage - https://openai.com/index/dall-e/
4) Overview of text-to-image models - https://byby.dev/ai-text-to-image-models
5) GoEmotions Dataset for Emotion Detection - https://github.com/google-research/google-research/tree/master/goemotions
6) Spotify API Documentation - https://developer.spotify.com/documentation/web-api