

AI PDF-to-Audio Platform — Requirements Document

1. Core Purpose

The project is a web-based platform that allows users to upload books or PDF documents and automatically convert them into natural, human-sounding audio using advanced AI voice technology. The main audience includes authors, educators, students, and audiobook enthusiasts who prefer listening instead of reading. In addition to serving end users, the platform will include a developer API so that other software companies or SaaS applications can integrate the same PDF-to-audio conversion system within their own platforms.

2. Core Concept

The core component of this platform is an artificial intelligence voice engine capable of producing speech that sounds expressive, realistic, and human-like. Unlike traditional robotic text-to-speech systems, this AI generates audio with emotion, tone, and pacing similar to professional voice narration. The system is designed to work in two modes: a local mode that can run on a GPU-powered laptop for testing and development without external costs, and a cloud mode that provides scalable performance for production use.

3. Key Features

The platform will include a complete user system with registration, login, and profile management through email verification. Each user will have access to a personal dashboard where they can upload PDF files, track their conversions, and download or play the generated audio. The system will support realistic AI-generated narration, multiple language options, and a user-friendly interface for managing subscriptions and usage. A developer API will be provided for external integrations, and an admin panel will allow management of users, files, and overall system activity.

4. PDF-to-Audio Conversion

Users will be able to upload PDF files up to thirty megabytes in size. The system will automatically extract the text and process it chapter by chapter for smoother narration and faster conversion. Once processed, the text will be converted into lifelike audio narration that users can either stream directly or download as an MP3 or WAV file. The voice engine will support various voice types, accents, and playback speeds. Both the original PDF and the resulting audio file will be stored securely, and users will be able to monitor the progress of each conversion from start to finish.

5. Subscription Plans

The platform will offer two subscription tiers.

The **Free Plan** will cost nothing and allow up to five conversions per month, with each file limited to one hundred and fifty pages. It will use the standard realistic AI voice and give users a chance to experience the quality of the voice engine by converting entire chapters or shorter books.

The **Pro Plan** will be priced at thirty-nine dollars per month and allow up to fifty conversions per month, with each file limited to five hundred pages. This plan will include access to the premium human-like voice library, faster processing speed, and priority in the conversion queue. Unused conversions will reset monthly, and accounts that are not renewed will automatically revert to the Free Plan.

6. AI Voice Engine

The AI voice engine is the core of this platform and is responsible for generating the realistic and expressive speech that defines the product. It will produce audio that closely resembles a real human narrator, maintaining clarity, emotion, and natural rhythm. Users will be able to adjust the tone, emotion, and speed of the voice. The engine will include a preview feature, allowing users to hear a short sample before generating the full conversion. It can be deployed locally on a GPU laptop for internal testing or run in the cloud for live production use.

7. Developer API

The platform will include a REST API that enables external developers and businesses to integrate the PDF-to-audio conversion feature into their own applications. Through this API, developers can upload PDF files, check conversion status, retrieve the final audio output, and monitor their usage limits. Each developer will receive a secure API key for authentication. Usage quotas and access levels will depend on the user's subscription plan, and the system will maintain detailed logs for monitoring, analytics, and billing transparency.

8. Admin Panel

An administrative interface will be included to manage the overall platform. Administrators will be able to view and edit user accounts, manage subscription plans, monitor conversion activity, and track system performance. The panel will also allow

administrators to delete old or unnecessary files, review error logs, and monitor analytics such as conversion rates and API usage. This section ensures that the platform remains secure, efficient, and easy to manage as it scales.

9. Notifications

The platform will include automated notifications through email and the user dashboard. Emails will be sent for account verification, conversion completion, password resets, and subscription updates. Dashboard notifications will provide real-time feedback on the status of uploads, payments, and conversions, ensuring that users are always aware of their account activity.

10. Database and Storage

All user information, subscription details, and usage statistics will be securely stored in a cloud-based database. Metadata for uploaded PDFs and generated audio files will also be recorded. API requests and conversion activity will be logged for analytics and reporting purposes. Both uploaded files and generated audio will be encrypted during transfer and at rest to maintain data security and privacy.

11. Security

The platform will use full HTTPS encryption for all communication between users, the server, and the API. Authentication will be token-based to ensure secure access. Rate limiting will be implemented to prevent system abuse, and strict access control will protect sensitive data. Files that are inactive or expired will be automatically deleted to optimize storage and ensure compliance with data protection standards.

12. User Interface Pages

The interface will include several key sections: a home page describing the platform and showcasing a demo, authentication pages for login and signup, a user dashboard for managing uploads and conversions, a pricing page for subscription details, an API documentation page for developers, and a help and support page containing frequently asked questions and a contact form.

13. Analytics and Monitoring

The system will track key metrics such as daily and monthly conversions, API usage, active users, and system performance. Administrators will be able to view summaries and usage patterns through the admin panel. The data will be used to improve the AI engine's performance and optimize user experience.

14. Optional Enhancements

Future updates may include a selection of narrator voices and languages, a “Book Mode” that automatically divides and converts large documents chapter by chapter, integration with cloud drives such as Google Drive and Dropbox, a referral and affiliate system to encourage user growth, and optional dark or light themes for the user interface.

Summary

This platform is designed to make audiobook creation simple, accessible, and affordable by combining PDF processing with AI-powered natural voice narration. It aims to serve both individual users and businesses through its easy-to-use dashboard and developer API. The inclusion of both local GPU testing and cloud deployment provides flexibility for development and scalability for production, making this a versatile solution for AI-driven audiobook generation.