x.1 Project Development

Using a multi-phase approach, the project development phase aims at converting text from PDFs into audio books. This methodical procedure encompasses several steps, each leveraging various Python libraries. Specifically, we have utilized PyPDF2 for PDF parsing, Pillow for image processing, Pytesseract for OCR (image to text) functionality, and potentially pyttsx3 for text-to-speech synthesis. The project commences with the selection of an input file, which can either be a PDF or an image. Based on this file format, appropriate preprocessing, and OCR text extraction techniques such as converting it to grayscale or enhancing its quality are applied to the input file. A pivotal aspect of our approach is text recognition, which extends beyond conventional text to include recognition of handwritten text, tables, diagrams, and programming language code. These additional features are intended to be seamlessly integrated into our system. Following text extractor, text-to-speech synthesis will be employed to convert the extracted text into audio format. The final output will be stored as an MP3 audio book file. The workflow of the project is designed to ensure user-friendly access to audio books, thereby promoting inclusivity and diversity by enhancing accessibility to various types of textual content. Fig x shows a flowchart which illustrates detailed streamlined process of converting text from PDFs or images into audio books, enhancing accessibility and diversity in accessing textual content.

A diagram of a software flow

Description automatically generated

Fig x.1 Flowchart of audio-book system with different components

x. 2 Project Testing