

A Synopsis on

“PDF to Audio Converter with Language Translation”

Submitted By

Name of Student : Devashri Nandkumar Bhosale

Roll No : A19

Department of Artificial Intelligence

**G.H.Raisoni College of Engineering and Management
Wagholi, Pune – 412207**

2021-22

Contents

1. Abstract & Technical Keywords

The PDF to Audio Converter project provides an alternative way to access the PDF books for blind, lazy, readers and others. Using this PDF to Audio Converter the user will be able to listen to his/her favorite PDF and can do their daily routine. The application can be used to read any PDF which has page numbers. The following application can be used to convert text from PDF to audio using Python predefined libraries.

Keywords: Python, pygame, User-Interface, Text to Speech, converter , glob , PySimpleGUI , fitz , pytesseract , os

2. Introduction

The proposed PDF to audio converter is intended to do the task of reading a pdf which will eliminate the efforts of printing the bulky Braille script books & manual recording of normal textbooks. The system is provided with a language translator, where one can choose and add code of their preferred language.

Text-to-speech and related read audio tools are being widely implemented in an attempt to assist students' reading comprehension skills. PDF to the audio system is a screen reader application designed and constructed for an effective audio communication system. PDFs were designed to present and exchange documents reliably, PDFs are an open standard document format used globally, maintained by the International Organization for Standardization (ISO). The document format is one of the most convenient methods for electronic communication, and also for the exchange of information. Hence, there is a need to make it more accessible to readers on-screen through audio. The PDF to the audio system will power text on screens to read aloud (speak) with support for many languages. The PDF to Audio Converter project provides an alternative to access the PDF books for the blind, lazy, readers, and others. Using this PDF to Audio Converter the user will be able to

listen to his\her favorite PDF and can do their daily routine. The following application can be used to convert text from PDF to audio using Python predefined libraries.

In the current busy routine people do not have time to take a book and spend time reading it, instead, everyone needs alternative access to read the content. If a person is traveling, he\she cannot read a book, instead of reading, they can listen to it. Reading stories or essays or any text can be arduous however an audiobook would make the task easy, by reading the text. However, an audio reading of the text is convenient and does not require much concentration as reading requires. When a person tends to read a book, it requires to invest his/her time in reading. Whereas the audiobook makes the task easy, and the user can perform their own task as well as listening to the audio. In this project, we have implemented a simple PDF to audio converter using python. When we compare with the current features present in a normal audiobook converter, they convert PDF texts (or images) into speech, and they have volume controls with single voice conversion (either male or female). Only a single choice is given to the user in case of voice modification. They provide the play and pause options. The speed of voice is always fixed.

While seeing the Audiobook converter which we have built, it converts the PDF text (or images) into speech with differences in the same PDF document. Voice conversions are possible. It contains both the male and female voice modifying features, which helps the user to change with a single click on the button. The rate of speed of voices can also be changed (fast, normal, and slow) for better clarifications.

The user can also save, and modify the PDF document, while the audio of the converted PDF is running in the background. This exclusive feature reduces the time consumption of writing notes and reduces paper consumption.

3. Literature Review/Related work

Sr. No.	Name	Method	Observations
1.	Dorian Miller., 2003	Microsoft's TTS engine. Microsoft Speech Application Programming Interface (SAPI), LAME software is used	The software is developed by python programming language and third-party components .
2.	Mentor Hamiti, Agni Dika., 2010	They have used Acoustic files of special letters and basic units albanian language [6].	In this paper, they presented that generating speech from written text in Albanian language, which helps the user those who have difficulties with their eye vision
3.	Itunuoluwa Isewon, AJelili Oyelade, Olufunke Oladipupo., 2014	Natural Language Processing (NLP) and Digital Signal Processing (DSP) technology are used	The software is called TextToSpeech Robot, it is a simple application that has text to speech functionality. The system was developed using Java programming language. Java programming language is used because it is an independent platform

4.	S. Venkateswarlu et. al, 2016	Optical Character Recognition (OCR) and Text to Speech Synthesizer (TTS) in Raspberry pi are used.	The present paper has introduced a real time cost beneficial technique. It has the concept of optical character recognition and text to speech synthesizer. It basically has two modules, image processing module and voice processing .
5.	Mr. Manohar M.,2020	Python predefined libraries are used .	This application can read the text from a PDF that the user has selected which has page numbers .

4. Proposed Work and Objectives

In this current busy routine people do not find time to read a book, or to convert the PDF file into an MP3 player using third-party applications or web applications. Even I have a directory at which I store pdf books that I plan on reading, but I never do. So, I thought hey, why do not I make them audiobooks and listen to them while I do something else ! In this system, we are developing a GUI application using python to convert the PDF file into audio format and read it out to the user. The application is more user-friendly as it does not require any audio file or MP3 player. The user will have to select the PDF file which the user wants to listen to.

The problem Statement of this project is:

To create a PDF to Audiobook Converter using different Tkinter and python files, functions, and definitions.

The main packages used in this audiobook converter are PyMuPDF,Pytesseract, fitz and Pygame.

PyMuPDF is a Python binding for MuPDF – a lightweight PDF, XPS, and E-book viewer, renderer, and toolkit, which is maintained and developed by Artifex Software, Inc

MuPDF can access files in PDF, XPS, OpenXPS, CBZ, EPUB and FB2 (e-books) formats, and it is known for its top performance and high rendering quality.

MuPDF stands out among all similar products for its top rendering capability and unsurpassed processing speed. At the same time, its “light weight” makes it an excellent choice for platforms where resources are typically limited, like smartphones.

Python-tesseract is a wrapper for Google’s Tesseract-OCR Engine. It is also useful as a stand-alone invocation script to tesseract, as it can read all image types supported by the Pillow and Leptonica imaging libraries, including jpeg, png, gif, bmp, tiff, and others. Additionally, if used as a script, Python-tesseract will print the recognized text instead of writing it to a file.

The pygame library is an open-source module for the Python programming language specifically intended to help you make games and other multimedia applications. Built on top of the highly portable SDL (Simple DirectMedia Layer) development library, pygame can run across many platforms and operating systems.

By using the pygame module, you can control the logic and graphics of your games without worrying about the backend complexities required for working with video and audio.

Googletrans is a **free** and **unlimited** python library that implemented Google Translate API. This uses the Google Translate Ajax API to make calls to such methods as detect and translate.

PDF to Audio Converter is a GUI application containing play, pause(buttons), and label to display text which allows the user to select the PDF files and the user has to click on the play button in order to extract text from the PDF file and will read the text. The application has been developed in a way that until the speaker reads out the extracted text the player cannot be paused. The GUI is also provided with a label to display the text, the text will only be displayed only after reading the extracted text.

With PDF being the most used document format globally, there is a need to convert the text in PDF formats into Audio signals. These can be utilized for various purposes, e.g., in the educational system, car navigation, announcements in railway stations, response services in telecommunications, and email reading. Furthermore, people with vision disabilities cannot view or read PDF files and this is a major setback. This research addresses the problems in converting PDF text into speech. One is how to improve the naturalness of synthetic speech in PDF-based text into an Audio system.

The Workflow of the project is:

- In this PDF to Audio Converter the user needs to select any PDF file from the desired location by pressing the open pdf.
- After selecting the PDF file, we have to select the Range of pages we want to convert in the audio.
- After selecting the Range, we have to select in which language we want the pdf to be converted .
- After pressing enter key ,Program will display the desired output.
- To exit the program, we press the cancel button.

The **objective** of our project is:

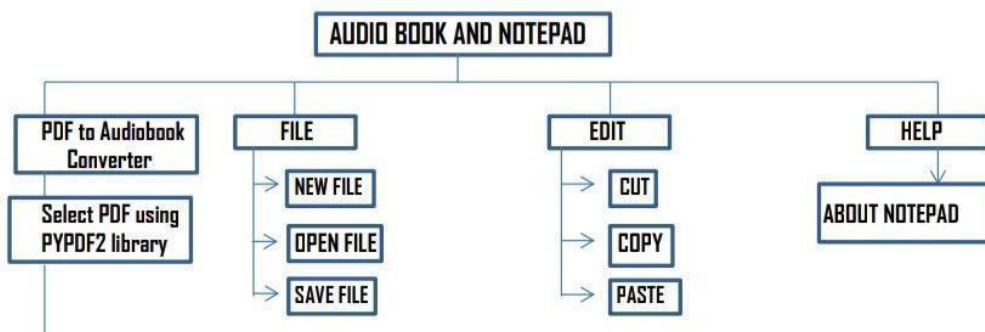
In the current generation students, researchers, authors don't find time to read a book on an electronic device as that might strain their eyes and might face other issues (headache, itchiness in the eye). So, to overcome those problems we have designed an application that extracts the text from the selected PDF and reads it out to the user.

Audiobooks allow students to hear explicit sounds of letters and letter patterns that form words. Audiobooks also help students engage in text and gain exposure to more words, ultimately improving vocabulary, comprehension, and critical thinking skills.

In this project, we present a simple way to combine different Python libraries for creating an Audiobook that takes a PDF file path as input and reads the text in the PDF file to the user via audio. Python programming language was used to create this project

5. Mathematical Model

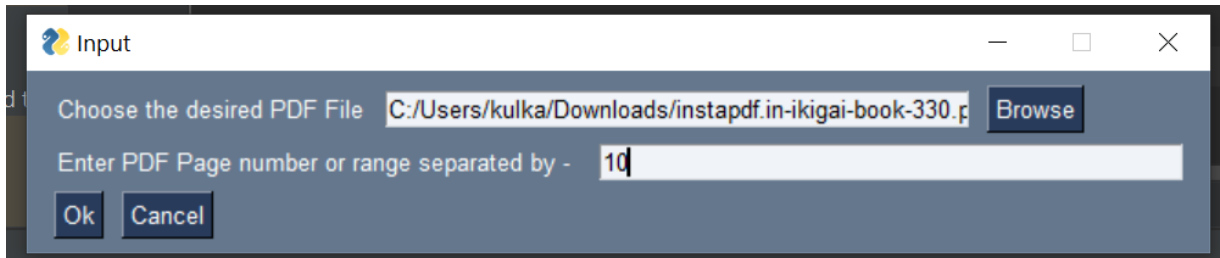
Proposed system - Block Diagram



Algorithm:

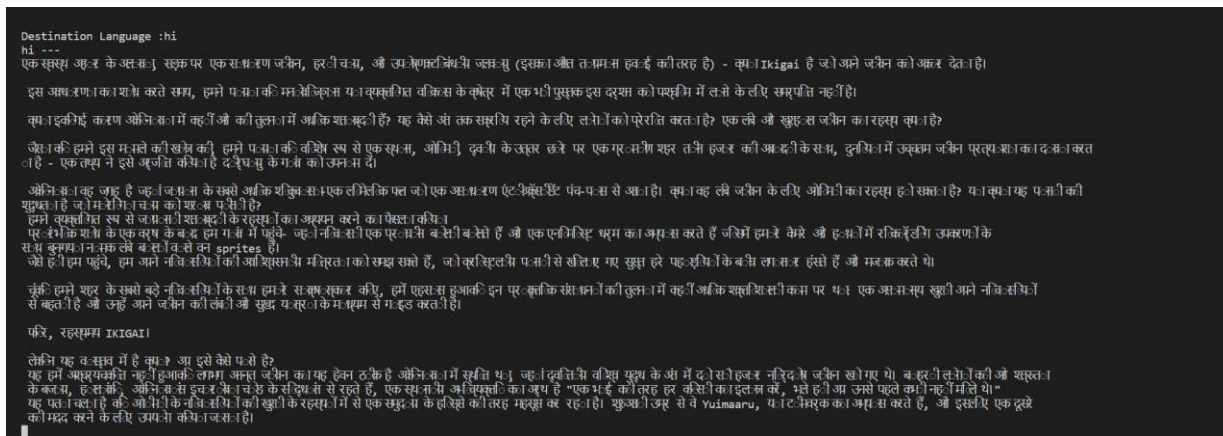
1. Take the PDF file and convert each page into image using PyMuPDF library.
2. Take the image(s) and scan the text in the image using Pytesseract OCR library.
3. Use Google Text to Speech (gTTS) library to convert text to audio file.
4. Get the Pygame mixer to play the audio file loud.
5. Convert text in selected language and create a .txt file for storing the converted text.

6. Desired Implications

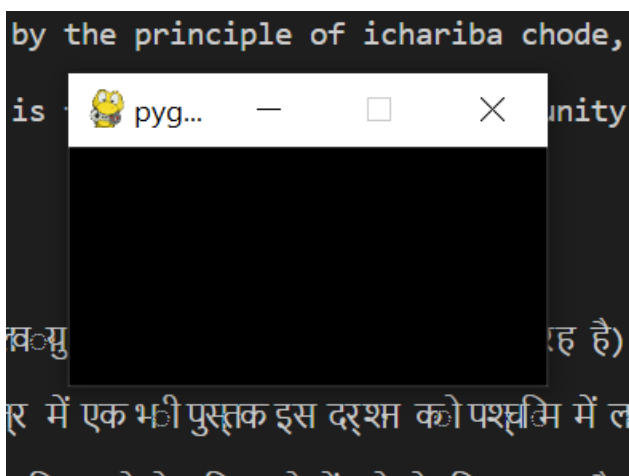


SCREENSHOT OF PDF TO AUDIO CONVERTER INTERFACE

TRANSLATED TEXT :



AUDIO FILE GENERATED :



7. References: (as per IEEE format)

- AJINKYA DOMALE; BHIMSEN PADALKAR; RAJ PAREKH; M.A. JOSHI
PRINTED BOOK TO AUDIO BOOK CONVERTER FOR VISUALLY IMPAIRED
<https://ieeexplore.ieee.org/document/6757125>
- Bernard Conan Pobiak-ADJUSTABLE ACCESS ELECTRONIC BOOKS, The Reader Project of Washington, D.C., 1992 IEEE
- R. M. K. Sinha, "A Syntactic pattern analysis system and its application to Devanagari script recognitio," Ph.D. Thesis, Dept. Elect. Eng., Indian Institute of Technology, Kanpur, India, 1973.
- R. M. K. Sinha and H. Mahabala, "Machine recognition of Devanagari script" IEEE Trans. Syst. Man Cybern., vol. 9, no. 8, pp. 435-441, Aug.1979.
- S. Palit and B. B. Chaudhuri, "A feature-based scheme for the machine recognition of printed Devanagari script" in Pattern Recognition, Image Processing and Computer Vision, P. P. Das and B. N. Chatterjee, Eds. New Delhi, India: Narosa Publishing House, 1995, pp. 163-168.
- P. V. S. Rao and R. B. Thosar, "A programming system for studies in speech synthesis," IEEE Trans. Acoust., Speech and Signal Processing, vol. 22, no. 3, pp. 217-225, 1974.
- B. Yegnanarayana, S. Rajendran, V. R. Ramachandran, and A. S. Madhukumar, "Significance of knowledge sources for a text-to-speech system for Indian languages," Sadhana, pp. 147-169,1994
- S. S. Agrawal, "Synthesizing Hindi speech using Klsyn and Hlsyn for natural sounding," in Workshop on Spoken Language Processing, TIFR, Mumbai, 2003.
- X. A. Furtado and A. Sen, "Synthesis of unlimited speech in Indian languages using formant-based rules," Sadhana, vol. 21, no. 3, pp. 345-362, 1996.
- Kishore Prahallad, E.Naresh Kumar, Venkatesh Keri, S.Rajendran, Alan W Black," The IIIT-H Indic Speech Databases
- Jan Mennens "Optical recognition of Braille writing", IEEE 1993. pp 428-431.
- Rajamani M. Kulkarni; Preetam R. Kalburgimath An Intelligent Text Reader based on Python [An Intelligent Text Reader based on Python | IEEE Conference Publication | IEEE Xplore](#)
- Design and Implementation of Text To Speech Conversion for Visually Impaired People Itunuoluwa Isewon*, Jelili Oyelade, Olufunke Oladipupo [Design and Implementation of Text To Speech Conversion for Visually Impaired People \(core.ac.uk\)](#)
- Chithra Selvaraj, Bhalaji Natarajan Enhanced portable text to speech converter for visually impaired.
 - [\(PDF\) Enhanced portable text to speech converter for visually impaired \(researchgate.net\)](#)

Coordinator Name
Mr.Sharad Adsure Sir

Prof.Rachna Sable
H.O.D

Class Teacher:
Ganesh Kadam Sir