

# File Conversion

Spoken Tutorial Project

<https://spoken-tutorial.org>

National Mission on Education through ICT

Sai Sathwik Matury

Domain: Dr. T. Subbulakshmi

VIT Chennai

9 August 2024



# Learning Objectives

In this tutorial, we will learn how to



# Learning Objectives

In this tutorial, we will learn how to

- **Convert JPG images to PNG format**



# Learning Objectives

In this tutorial, we will learn how to

- Convert JPG images to PNG format
- Extract text from PDF file and convert to audio



# Learning Objectives

In this tutorial, we will learn how to

- Convert JPG images to PNG format
- Extract text from PDF file and convert to audio
- Convert a PDF file to DOCX format



# System Requirements

**To record this tutorial, I am using**



# System Requirements

To record this tutorial, I am using

- **Ubuntu Linux OS 22.04**



# System Requirements

To record this tutorial, I am using

- Ubuntu Linux OS 22.04
- Python 3.12.3





# Pre-requisites

- You must have basic knowledge of using Linux Terminal and Python



# Pre-requisites

- You must have basic knowledge of using Linux Terminal and Python
- For pre-requisite Linux and Python Tutorials, please visit <https://www.spoken-tutorial.org/>



# Pre-requisites

- You must have basic knowledge of using Linux Terminal and Python
- For pre-requisite Linux and Python Tutorials, please visit <https://www.spoken-tutorial.org/>
- Python libraries required for automation must be installed



# Code Files

- The files used in this tutorial are provided in the Code files link



# Code Files

- The files used in this tutorial are provided in the Code files link
- Please download and extract the files



# Code Files

- The files used in this tutorial are provided in the Code files link
- Please download and extract the files
- Make a copy and then use them while practicing



# File Conversion

- **File conversion is the process of transforming a file from one format to another**



# File Conversion

- File conversion is the process of transforming a file from one format to another
- This can involve changing the file type, the structure of its contents, or both





# File Conversion

- This makes the file compatible with different applications and user requirements



# File Conversion - Libraries



# File Conversion - Libraries

- **PyPDF2 to handle PDF manipulation**



# File Conversion - Libraries

- **PyPDF2 to handle PDF manipulation**
- **Pyttsx3 to help with extracting content**



# File Conversion - Libraries

- **PyPDF2 to handle PDF manipulation**
- **Pyttsx3 to help with extracting content**
- **Pdfplumber to convert the text to speech**



# File Conversion - Libraries

- **PIL for image processing and manipulation**



# File Conversion - Libraries

- **PIL for image processing and manipulation**
- **pdf2docx to convert PDF files to DOCX format**



# File Conversion - Libraries

- **PIL for image processing and manipulation**
- **pdf2docx to convert PDF files to DOCX format**
- **os handles file manipulation and operating system interactions**





# Install - espeak package



# Install - espeak package

- `sudo apt-get update`
- `sudo apt-get install espeak`



# Summary

In this tutorial, we have learnt to

- Convert JPG images to PNG format
- Extract text from PDF file and convert to audio
- Convert a PDF file to DOCX format



# Assignment

**As an assignment, please do the following:**

- **Take a GIF and convert it to an image in PNG format**
- **Take a PDF of your favorite book and convert it to an audio book**



# About the Spoken Tutorial Project

- Watch the video available at [http://spoken-tutorial.org/What\\_is\\_a\\_Spoken\\_Tutorial](http://spoken-tutorial.org/What_is_a_Spoken_Tutorial)
- It summarises the Spoken Tutorial project
- If you do not have good bandwidth, you can download and watch it



# Spoken Tutorial Workshops

## The Spoken Tutorial Project Team

- Conducts workshops using spoken tutorials
- Gives certificates to those who pass an online test
- For more details, please write to [contact@spoken-tutorial.org](mailto:contact@spoken-tutorial.org)



# Answers for THIS spoken tutorial

- Questions in THIS Spoken Tutorial
- Visit <https://forums.spoken-tutorial.org>
- Choose the minute and second where you have the question
- Explain your question briefly
- The Spoken Tutorial project will ensure an answer
- You will have to register to ask questions



# FOSSEE Forum

- For any general or technical questions on **Python for Automation**, visit the FOSSEE forum and post your question <https://forums.fossee.in/>





# Acknowledgement

- The Spoken Tutorial project was established by Ministry of Education, Govt. of India



# Thank you

- This is Sai Sathwik, a FOSSEE Semester Long Intern 2024, IIT Bombay

