

How to convert PDF to text from uploaded file for PDF to text API in Python using PDF.co Web API

How to convert PDF to text from uploaded file for PDF to text API in Python: Step By Step Instructions

We regularly create and update our sample code library so you may quickly learn PDF to text API and the step-by-step process in Python. PDF to text API in Python can be applied with PDF.co Web API. PDF.co Web API is the flexible Web API that includes full set of functions from e-signature requests to data extraction, OCR, images recognition, pdf splitting and pdf splitting. Can also generate barcodes and read barcodes from images, scans and pdf.

Use the code displayed below in your application to save a lot of time on writing and testing code. This Python sample code can be used by copying and pasting into your project. Once done, just compile your project and click Run. Want to see how it works with your data then code testing will allow the function to be tested and work properly.

Our website provides free trial version of PDF.co Web API that gives source code samples to assist with your Python project.

FOR MORE INFORMATION AND FREE TRIAL:

[Download Free Trial SDK \(on-premise version\)](#)

[Read more about PDF.co Web API](#)

[Explore API Documentation](#)

[Get Free Training for PDF.co Web API](#)

[Get Free API key for Web API](#)

[visit www.ByteScout.com](http://www.ByteScout.com)

Source Code Files:

ConvertPdfToTextFromUploadedFile.py

```
import os
import requests # pip install requests

# The authentication key (API Key).
# Get your own by registering at https://app.pdf.co/documentation/api
API_KEY = "*****"

# Base URL for PDF.co Web API requests
BASE_URL = "https://api.pdf.co/v1"

# Source PDF file
SourceFile = ".\sample.pdf"
# Comma-separated list of page indices (or ranges) to process. Leave empty for all pages. Example: '0,2-5,7-'.
Pages = ""
# PDF document password. Leave empty for unprotected documents.
Password = ""
# Destination TXT file name
DestinationFile = ".\result.txt"

def main(args = None):
    uploadedFileUrl = uploadFile(SourceFile)
    if (uploadedFileUrl != None):
        convertPdfToText(uploadedFileUrl, DestinationFile)

def convertPdfToText(uploadedFileUrl, destinationFile):
    """Converts PDF To Text using PDF.co Web API"""

    # Prepare URL for 'PDF To Text' API request
    url = "{}/pdf/convert/to/text?name={}&password={}&pages={}&url={}".format(
        BASE_URL,
        os.path.basename(destinationFile),
        Password,
        Pages,
        uploadedFileUrl
    )

    # Execute request and get response as JSON
    response = requests.get(url, headers={ "x-api-key": API_KEY, "content-type": "application/octet-stream" })
    if (response.status_code == 200):
        json = response.json()

        if json["error"] == False:
            # Get URL of result file
            resultFileUrl = json["url"]
            # Download result file
            r = requests.get(resultFileUrl, stream=True)
            if (r.status_code == 200):
                with open(destinationFile, 'wb') as file:
                    for chunk in r:
                        file.write(chunk)
                print(f"Result file saved as \"{destinationFile}\" file.")
            else:
                print(f"Request error: {response.status_code} {response.reason}")
        else:
            # Show service reported error
            print(json["message"])
        else:
            print(f"Request error: {response.status_code} {response.reason}")

def uploadFile(fileName):
    """Uploads file to the cloud"""

    # 1. RETRIEVE PRESIGNED URL TO UPLOAD FILE.

    # Prepare URL for 'Get Presigned URL' API request
    url = "{}/file/upload/get-presigned-url?contenttype=application/octet-stream&name={}".format(
        BASE_URL, os.path.basename(fileName))

    # Execute request and get response as JSON
    response = requests.get(url, headers={ "x-api-key": API_KEY })
```

```
if (response.status_code == 200):
    json = response.json()

    if json["error"] == False:
        # URL to use for file upload
        uploadUrl = json["presignedUrl"]
        # URL for future reference
        uploadedFileUrl = json["url"]

        # 2. UPLOAD FILE TO CLOUD.
        with open(fileName, 'rb') as file:
            requests.put(uploadUrl, data=file, headers={"x-api-key": API_KEY, "content-type": "application/octet-stream"})

        return uploadedFileUrl
    else:
        # Show service reported error
        print(json["message"])
else:
    print(f"Request error: {response.status_code} {response.reason}")

return None

if __name__ == '__main__':
    main()
```

VIDEO

<https://www.youtube.com/watch?v=NEwNs2b9YN8>

ON-PREMISE OFFLINE SDK

[60 Day Free Trial](#) or [Visit PDF.co Web API Home Page](#)

[Explore PDF.co Web API Documentation](#)

[Explore Samples](#)

[Sign Up for PDF.co Web API Online Training](#)

ON-DEMAND REST WEB API

[Get Your API Key](#)

[Explore Web API Docs](#)

[Explore Web API Samples](#)

[visit www.ByteScout.com](http://www.ByteScout.com)

[visit www.PDF.co](http://www.PDF.co)

