

www.bytescout.com

delete text from PDF in Python using ByteScout Cloud API Server

What is ByteScout Cloud API Server? It is API server that is ready to use and can be installed and deployed in less than 30 minutes on your own Windows server or server in a cloud. It can save data and files on your local server-based file storage or in Amazon AWS S3 storage. Data is processed solely on the API server and is powered by ByteScout engine, no cloud services or Internet connection is required for data processing..

FOR MORE INFORMATION AND FREE TRIAL:

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

[Read more about **ByteScout Cloud API Server**](#)

[Explore API Documentation](#)

[Get Free Training for ByteScout Cloud API Server](#)

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

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

Source Code Files:

DeletePdfTextFromUrlAsynchronously.py

```

""" Cloud API asynchronous "PDF To Text" job example.
    Allows to avoid timeout errors when processing huge or scanned PDF documents.
"""
import os
import requests # pip install requests
import time
import datetime

# Please NOTE: In this sample we're assuming Cloud Api Server is hosted at "https://lo
# If it's not then please replace this with with your hosting url.

# 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://localhost"

# Direct URL of source PDF file.
SourceFileURL = "https://bytescout-com.s3.amazonaws.com/files/demo-files/cloud-api/pdf
# PDF document password. Leave empty for unprotected documents.
Password = ""
# Destination PDF file name
DestinationFile = ".\\result.pdf"
# (!) Make asynchronous job
Async = True

def main(args = None):
    deleteTextFromPdf(SourceFileURL, DestinationFile)

def deleteTextFromPdf(uploadedFileUrl, destinationFile):
    """Delete Text from PDF using PDF.co Web API"""

    # Prepare requests params as JSON
    # See documentation: https://apidocs.pdf.co
    parameters = {}
    parameters["async"] = Async
    parameters["name"] = os.path.basename(destinationFile)
    parameters["password"] = Password
    parameters["url"] = uploadedFileUrl
    parameters["searchString"] = "conspicuous"

    # Prepare URL for 'Delete Text from PDF' API request
    url = "{}pdf/edit/delete-text".format(BASE_URL)

    # Execute request and get response as JSON
    response = requests.post(url, data=parameters, headers={ "x-api-key": API_KEY })
    if (response.status_code == 200):
        json = response.json()

        if json["error"] == False:
            # Asynchronous job ID
            jobId = json["jobId"]
            # URL of the result file
            resultFileUrl = json["url"]

            # Check the job status in a loop.

```

```

# If you don't want to pause the main thread you can rework the code
# to use a separate thread for the status checking and completion.
while True:
    status = checkJobStatus(jobId) # Possible statuses: "working", "failed"

    # Display timestamp and status (for demo purposes)
    print(datetime.datetime.now().strftime("%H:%M:%S") + ": " + status)

    if status == "success":
        # 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}")
            break
    elif status == "working":
        # Pause for a few seconds
        time.sleep(3)
    else:
        print(status)
        break

else:
    # Show service reported error
    print(json["message"])

else:
    print(f"Request error: {response.status_code} {response.reason}")

def checkJobStatus(jobId):
    """Checks server job status"""

    url = f"{BASE_URL}/job/check?jobid={jobId}"

    response = requests.get(url, headers={ "x-api-key": API_KEY })
    if (response.status_code == 200):
        json = response.json()
        return json["status"]
    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 ByteScout Cloud API Server Home Page](#)
[Explore ByteScout Cloud API Server Documentation](#)
[Explore Samples](#)
[Sign Up for ByteScout Cloud API Server 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)

www.bytescout.com