www.bytescout.com

How to convert PDF to JSON with PDF extractor SDK in Powershell using ByteScout PDF Suite

Learn to code in Powershell to convert PDF to JSON with PDF extractor SDK with this step-by-step tutorial

Every ByteScout tool includes simple example Powershell source codes that you can get here or in the folder with installed ByteScout product. What is ByteScout PDF Suite? It is the set that includes 6 SDK products to work with PDF from generating rich PDF reports to extracting data from PDF documents and converting them to HTML. This bundle includes PDF (Generator) SDK, PDF Renderer SDK, PDF Extractor SDK, PDF to HTML SDK, PDF Viewer SDK and PDF Generator SDK for Javascript. It can help you to convert PDF to JSON with PDF extractor SDK in your Powershell application.

This prolific sample source code in Powershell for ByteScout PDF Suite contains various functions and other necessary options you should do calling the API to convert PDF to JSON with PDF extractor SDK. Just copy and paste the code into your Powershell application's code and follow the instructions. Enjoy writing a code with ready-to-use sample Powershell codes.

You can download free trial version of ByteScout PDF Suite from our website with this and other source code samples for Powershell.

FOR MORE INFORMATION AND FREE TRIAL:

Download Free Trial SDK (on-premise version)

Read more about ByteScout PDF Suite

Explore API Documentation

Get Free Training for ByteScout PDF Suite

Get Free API key for Web API

visit www.ByteScout.com

Source Code Files:

pdf-to-json.bat

```
@echo off

if "%~1"=="" (
    echo ------------------------------------------------------
    echo Invalid parameter!
    echo ------------------------------------------------------
    echo Usage: pdf-to-json.bat folder_path
    echo Example: pdf-to-json.bat "c:\documents"
    echo ------------------------------------------------------
    if not "%NOPAUSE%"=="1" pause
    exit /b 1
)

powershell -NoProfile -ExecutionPolicy Bypass -Command "& .\pdf-to-json.ps1" "%1"
echo Script finished with errorlevel=%errorlevel%

pause
```

pdf-to-json.ps1

```
Param(
    [Parameter(Mandatory = $true)]
    [string] $InputFolder = ""
)

# Add reference to Bytescout.PDFExtractor.dll assembly
Add-Type -Path "c:\Program Files\Bytescout PDF Extractor SDK\net4.00\Bytescout.PDFExtr

# Check input folder exists
if ((Test-Path $InputFolder) -eq $false) {
    Write-Host "Target folder does not exist." -ForegroundColor Red
    exit 0
}

# Create and activate JSONExtractor instance
$jsonExtractor = New-Object Bytescout.PDFExtractor.JSONExtractor
$jsonExtractor.RegistrationName = "demo"
$jsonExtractor.RegistrationKey = "demo"

try {
    # Get PDF files from input folder
    $files = Get-ChildItem -Path $InputFolder -Recurse -Include "*.pdf"
    foreach ($file in $files) {
        Write-Host "Input file" $file.FullName
        # Construct output file name
```

```
        $jsonFileName = [System.IO.Path]::ChangeExtension($file.FullName, "json")
        Write-Host "  Output file" $jsonFileName
        # Load PDF document
        $jsonExtractor.LoadDocumentFromFile($file.FullName)
        # Disable the formatting reconstruction
        $jsonExtractor.PreserveFormattingOnTextExtraction = $false
        # Extract first page to JSON
        $jsonExtractor.SaveJSONToFile(0, $jsonFileName)
        # Reset extractor
        $jsonExtractor.Reset()
    }
}
catch {
    Write-Host $_.Exception.Message
}

$jsonExtractor.Dispose()
```

VIDEO

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

ON-PREMISE OFFLINE SDK

60 Day Free Trial or Visit ByteScout PDF Suite Home Page
Explore ByteScout PDF Suite Documentation
Explore Samples
Sign Up for ByteScout PDF Suite Online Training

ON-DEMAND REST WEB API

Get Your API Key
Explore Web API Docs
Explore Web API Samples

visit www.ByteScout.com

visit www.PDF.co