

How to find text in PDF using regex in VBScript with ByteScout PDF Extractor SDK

The tutorial below will demonstrate how to find text in PDF using regex in VBScript

Learn how to find text in PDF using regex in VBScript with this source code sample. ByteScout PDF Extractor SDK is the SDK is designed to help developers with pdf tables and pdf data extraction from unstructured documents like pdf, tiff, scans, images, scanned and electronic forms. The library is powered by OCR, computer vision and AI to provide unique functionality like table detection, automatic table structure extraction, data restoration, data restructuring and reconstruction. Supports PDF, TIFF, PNG, JPG images as input and can output CSV, XML, JSON formatted data. Includes full set of utilities like pdf splitter, pdf merger, searchable pdf maker and other utilities. It can find text in PDF using regex in VBScript.

VBScript code samples for VBScript developers help to speed up coding of your application when using ByteScout PDF Extractor SDK. Just copy and paste the code into your VBScript application's code and follow the instruction. Use of ByteScout PDF Extractor SDK in VBScript is also explained in the documentation included along with the product.

You can download free trial version of ByteScout PDF Extractor SDK from our website to see and try many others source code samples for VBScript.

FOR MORE INFORMATION AND FREE TRIAL:

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

[Read more about ByteScout PDF Extractor SDK](#)

[Explore API Documentation](#)

[Get Free Training for ByteScout PDF Extractor SDK](#)

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

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

Source Code Files:

FindTextUsingRegex.vbs

```
' Create Bytescout.PDFExtractor.TextExtractor object
Set extractor = CreateObject("Bytescout.PDFExtractor.TextExtractor")
extractor.RegistrationName = "demo"
extractor.RegistrationKey = "demo"

' Load sample PDF document
extractor.LoadDocumentFromFile("../..\Invoice.pdf")

extractor.RegexSearch = True ' Turn on the regex search
pattern = "[0-9]{2}/[0-9]{2}/[0-9]{4}" ' Search dates in format 'mm/dd/yyyy'

' Get page count
pageCount = extractor.GetPageCount()

For i = 0 to PageCount - 1
    If extractor.Find(i, pattern, false) Then ' Parameters are: page index, string to
        Do
            extractedString = extractor.FoundText.Text
            MsgBox "Found match on page #" & CStr(i) & ": " & extractedString
            extractor.ResetExtractionArea()
        Loop While extractor.FindNext
    End If
Next

MsgBox "Done"

Set extractor = Nothing
```

VIDEO

https://www.youtube.com/watch?v=s28W3_KMraU

ON-PREMISE OFFLINE SDK

[60 Day Free Trial](#) or [Visit ByteScout PDF Extractor SDK Home Page](#)
[Explore ByteScout PDF Extractor SDK Documentation](#)
[Explore Samples](#)

[Sign Up for ByteScout PDF Extractor SDK 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 \[www.PDF.co\]\(http://www.PDF.co\)](#)

[\[www.bytescout.com\]\(http://www.bytescout.com\)](#)