www.bytescout.com

How to find hyphenated text in PDF in VBScript and ByteScout PDF Extractor SDK

This tutorial will show how to find hyphenated text in PDF in VBScript

Find hyphenated text in PDF is easy to implement in VBScript if you use these source codes below. ByteScout PDF Extractor SDK can find hyphenated text in PDF. It can be used from VBScript. ByteScout PDF Extractor SDK is the Software Development Kit (SDK) that is designed to help developers with 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.

This code snippet below for ByteScout PDF Extractor SDK works best when you need to quickly find hyphenated text in PDF in your VBScript application. In your VBScript project or application you may simply copy & paste the code and then run your app! Detailed tutorials and documentation are available along with installed ByteScout PDF Extractor SDK if you'd like to dive deeper into the topic and the details of the API.

ByteScout PDF Extractor SDK free trial version is available on our website. VBScript and other programming languages are supported.

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

Source Code Files:

FindTextAndGetCoordinates.vbs

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

' Load sample PDF document
extractor.LoadDocumentFromFile("..\..\words-with-hyphens.pdf")

' Set the matching mode:
' 0 = WordMatchingMode.None - treats the search string as substring;
' 1 = WordMatchingMode.SmartMatch - will find the word in various forms (like Adobe Rea
' 2 = WordMatchingMode.ExactMatch - treats the search string as separate word.
extractor.WordMatchingMode = 1

' Get page count

pageCount = extractor.GetPageCount()

For i = 0 To PageCount - 1

    If extractor.Find(i, "hyphen", false) Then ' parameters are: page index, string to
        Do
            foundMessage = "Found substring 'hyphen' on page #" & CStr(i) & " at { " &
                "x = " & CStr(extractor.FoundText.Left) & "; " & _
                "y = " & CStr(extractor.FoundText.Top) & "; " & _
                "width = " & CStr(extractor.FoundText.Width) & "; " & _
                "height = " & CStr(extractor.FoundText.Height) & " }"

            elementInfo = ""

            ' Iterate through elements of the found text object
            For j = 0 to extractor.FoundText.ElementCount - 1
                Set element = extractor.FoundText.GetElement(j)
                elementInfo = elementInfo & "Element #" & CStr(j) & " at { x = " & CStr
                elementInfo = elementInfo & "Text: " & CStr(element.Text) & vbCRLF
                elementInfo = elementInfo & "Font is bold: " & CStr(element.FontIsBold)
                elementInfo = elementInfo & "Font is italic: " & CStr(element.FontIsIta
                elementInfo = elementInfo & "Font name: " & CStr(element.FontName) & vb
                elementInfo = elementInfo & "Font size: " & CStr(element.FontSize) & vb
                elementInfo = elementInfo & "Font color (as OLE_COLOR): " & CStr(elemen
            Next

            WScript.Echo foundMessage & vbCRLF & vbCRLF & elementInfo

        Loop While extractor.FindNext

    End If

Next

WScript.Echo "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

visit www.PDF.co