

How to find email addresses in PDF using regex in VB.NET and ByteScout PDF Extractor SDK

Tutorial on how to find email addresses in PDF using regex in VB.NET

ByteScout tutorials are designed to explain the code for both VB.NET beginners and advanced programmers. ByteScout PDF Extractor SDK can find email addresses in PDF using regex. It can be used from VB.NET. 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.

VB.NET code samples for VB.NET developers help to speed up coding of your application when using ByteScout PDF Extractor SDK. This VB.NET sample code is all you need for your app. Just copy and paste the code, add references (if needs to) and you are all set! Implementing VB.NET application typically includes multiple stages of the software development so even if the functionality works please test it with your data and the production environment.

Free trial version of ByteScout PDF Extractor SDK is available on our website. Documentation and source code samples are included.

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:

Program.vb

```
Imports Bytescout.PDFExtractor

Module Program

    Sub Main()

        Try

            ' Create Bytescout.PDFExtractor.TextExtractor instance
            Using extractor As TextExtractor = New TextExtractor()
                extractor.RegistrationName = "demo"
                extractor.RegistrationKey = "demo"

                ' Load sample PDF document
                extractor.LoadDocumentFromFile("samplePDF_EmailAddress.pdf")

                ' Enable the regular expression
                extractor.RegexSearch = True

                Dim pageCount As Integer = extractor.GetPageCount()

                ' Search through pages
                For i As Integer = 0 To pageCount - 1
                    ' Search Email addresses
                    Dim regexPattern As String = "\b[A-Za-z0-9._%+-]+@[A-Za-z0-9.-]+\.[A-Z|a-z]{2,}\b"
                    ' See the complete regular expressions reference at https://msdn.microsoft.com/en-us/library/2185e26c.aspx

                    ' Search each page for the pattern
                    If extractor.Find(i, regexPattern, False) Then

                        Do
                            ' Iterate through each element in the found text
                            For Each element As ISearchResultElement In extractor.FoundText(i)
                                Console.WriteLine("Found Email Addresses: " & element.Text)
                            Next
                        Loop While extractor.FindNext()

                    End If
                Next
            End Using

            Catch ex As Exception
                Console.WriteLine("Error: " & ex.Message)
            End Try

            Console.WriteLine()
            Console.WriteLine("Press enter key to continue...")
            Console.ReadLine()

        End Sub

    End Module
```

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