www.bytescout.com

How to find US address in PDF with regex in C# and ByteScout PDF Extractor SDK

Tutorial on how to find US address in PDF with regex in C#

Learn how to find US address in PDF with regex in C# with this source code sample. ByteScout PDF Extractor SDK: the SDK that helps developers to extract data from unstructured documents, pdf, images, scanned and electronic forms. Includes AI functions like automatic table detection, automatic table extraction and restructuring, text recognition and text restoration from pdf and scanned documents. Includes PDF to CSV, PDF to XML, PDF to JSON, PDF to searchable PDF functions as well as methods for low level data extraction. It can find US address in PDF with regex in C#.

The SDK samples like this one below explain how to quickly make your application do find US address in PDF with regex in C# with the help of ByteScout PDF Extractor SDK. In your C# project or application you may simply copy & paste the code and then run your app! Enjoy writing a code with ready-to-use sample codes in C#.

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

Source Code Files:

Program.cs

```
using Bytescout.PDFExtractor;
using System;

namespace FindUsAddressRegex
{
    class Program
    {
        static void Main(string[] args)
        {
            try
            {
                // Create Bytescout.PDFExtractor.TextExtractor instance
                using (TextExtractor extractor = new TextExtractor())
                {
                    extractor.RegistrationName = "demo";
                    extractor.RegistrationKey = "demo";

                    // Load sample PDF document
                    extractor.LoadDocumentFromFile("samplePDF_Address.pdf");

                    extractor.RegexSearch = true; // Enable the regular expressions

                    int pageCount = extractor.GetPageCount();

                    // Search through pages
                    for (int i = 0; i < pageCount; i++)
                    {
                        // Search Address
                        string regexPattern = @"((\w+[ ,])+ ){2}([a-zA-Z]){2}[ , ] (\d-
                        // See the complete regular expressions reference at https://ms

                        // Search each page for the pattern
                        if (extractor.Find(i, regexPattern, false))
                        {
                            do
                            {
                                // Iterate through each element in the found text
                                foreach (ISearchResultElement element in extractor.Four
                                {
                                    Console.WriteLine("Found Address: " + element.Text
                                }
                            }
                            while (extractor.FindNext());
                        }
                    }
                }
            }
            catch (Exception ex)
            {
                Console.WriteLine("Error: " + ex.Message);
            }

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

```
        }
    }
}
```

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