How to find keyword and extract page in PDF in C# and ByteScout PDF Extractor SDK

The tutorial shows how to find keyword and extract page in PDF in C#

The sample shows steps and algorithm of how to find keyword and extract page in PDF and how to make it work in your C# application. What is ByteScout PDF Extractor SDK? It 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. It can help you to find keyword and extract page in PDF in your C# application.

This code snippet below for ByteScout PDF Extractor SDK works best when you need to quickly find keyword and extract page in PDF in your C# application. In order to implement the functionality, you should copy and paste this code for C# below into your code editor with your app, compile and run your application. Implementing C# 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.

Trial version of ByteScout PDF Extractor SDK can be downloaded for free from our website. It also includes source code samples for C# and other programming languages.

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:

Program.cs

```csharp
// This example page extraction by found keyword.

using System;
using Bytescout.PDFExtractor;

namespace SplittingExample
{
    class Program
    {
        static void Main(string[] args)
        {
            string inputFile = @".\sample2.pdf";

            // Create Bytescout.PDFExtractor.TextExtractor instance
            TextExtractor extractor = new TextExtractor();
            extractor.RegistrationName = "demo";
            extractor.RegistrationKey = "demo";

            // Load sample PDF document
            extractor.LoadDocumentFromFile(inputFile);

            int pageCount = extractor.GetPageCount();

            // Search each page for a keyword
            for (int i = 0; i < pageCount; i++)
            {
                if (extractor.Find(i, "bombardment", false))
                {
                    // Extract page
                    using (DocumentSplitter splitter = new DocumentSplitter("demo", "de
                    {
                        splitter.OptimizeSplittedDocuments = true;

                        int pageNumber = i + 1;  // (!) page number in ExtractPage() i
                        string outputFile = @".\page" + pageNumber + ".pdf";
                        splitter.ExtractPage(inputFile, outputFile, pageNumber);

                        Console.WriteLine("Extracted page " + pageNumber + " to file \
                    }
                }
            }

            // Cleanup
            extractor.Dispose();

            Console.WriteLine();
            Console.WriteLine("Press any key...");
            Console.ReadKey();
        }
    }
}
```

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