

repair text in PDF in C# using ByteScout PDF Extractor SDK

How to use ByteScout PDF Extractor SDK for repair text in PDF in C#

Sample source codes below will show you how to cope with a difficult task, for example, repair text in PDF in C#. ByteScout PDF Extractor SDK was made to help with repair text in PDF in C#. 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.

Fast application programming interfaces of ByteScout PDF Extractor SDK for C# plus the instruction and the C# code below will help you quickly learn repair text in PDF. To do repair text in PDF in your C# project or application you may simply copy & paste the code and then run your app! Further enhancement of the code will make it more vigorous.

Our website provides free trial version of ByteScout PDF Extractor SDK. It comes along with all these source code samples with the goal to help you with your C# application implementation.

C# - Program.cs

```
using Bytescout.PDFExtractor;
using System;

namespace RepairText
{
    class Program
    {
        static void Main(string[] args)
        {
            try
            {
                //Read all text from pdf file
                using (TextExtractor extractor = new TextExtractor())
                {
                    // Load PDF document
                    extractor.LoadDocumentFromFile("sample.pdf");

                    // Set the font repairing OCR mode
                    extractor.OCRMode =
                    OCRMode.TextFromImagesAndVectorsAndRepairedFonts;

                    // Set the location of OCR language data files
                    extractor.OCRLanguageDataFolder = @"c:\Program Files\Bytescout
```

```
PDF Extractor SDK\ocrdata\";

        // Set OCR language
        extractor.OCRLanguage = "eng"; // "eng" for english, "deu" for
German, "fra" for French, "spa" for Spanish etc - according to files in "ocrdata"
folder

        // Find more language files at
https://github.com/bytescout/ocrdata

        // Set PDF document rendering resolution
        extractor.OCRResolution = 300;

        //Read all text
        string allText = extractor.GetText();

        Console.WriteLine("Extracted Text: \n\n" + allText);
    }
}
catch (Exception ex)
{
    Console.WriteLine(ex.Message);
}
Console.ReadLine();
}
}
```

FOR MORE INFORMATION AND FREE TRIAL:

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

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

[Explore documentation](#)

[Visit www.ByteScout.com](#)

or

[Get Your Free API Key for www.PDF.co Web API](#)