

check if OCR is required for PDF in C# using ByteScout PDF Extractor SDK

How To: tutorial on check if OCR is required for PDF in C#

Today you are going to learn how to check if OCR is required for PDF in C#. Check if OCR is required for PDF in C# can be implemented with ByteScout PDF Extractor SDK. 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.

The SDK samples like this one below explain how to quickly make your application do check if OCR is required for PDF in C# with the help of ByteScout PDF Extractor SDK. C# sample code is all you need: copy and paste the code to your C# application's code editor, add a reference to ByteScout PDF Extractor SDK (if you haven't added yet) and you are ready to go! Enjoy writing a code with ready-to-use sample C# codes to add check if OCR is required for PDF functions using ByteScout PDF Extractor SDK in C#.

Trial version can be downloaded from our website. Source code samples for C# and documentation are included.

C# - Program.cs

```
using Bytescout.PDFExtractor;
using System;

namespace CheckIfOCRIIsRequired
{
    class Program
    {
        static void Main(string[] args)
        {
            try
            {
                // Loop through all files in directory and check whether OCR
                operation is required
                foreach (string filePath in
                    System.IO.Directory.GetFiles("InputFiles"))
                {
                    _CheckOCRRequired(filePath);
                }
            }
        }
    }
}
```

```

        catch (Exception ex)
        {
            Console.WriteLine("Error: " + ex.Message);
        }

        Console.WriteLine("Press enter key to exit...");
        Console.ReadLine();
    }

    ///
    /// Check whether OCR Operation is required
    ///
    ///
    private static void _CheckOCRRequired(string filePath)
    {
        //Read all file content...
        using (TextExtractor extractor = new TextExtractor())
        {
            extractor.RegistrationKey = "demo";
            extractor.RegistrationName = "demo";

            // Load document
            extractor.LoadDocumentFromFile(filePath);
            Console.WriteLine("\n*****\n\nFilePath: {0}",
filePath);

            int pageIndex = 0;

            // Identify OCR operation is recommended for page
            if (extractor.IsOCRRecommendedForPage(pageIndex))
            {
                Console.WriteLine("\nOCR Recommended: True");

                // Enable Optical Character Recognition (OCR)
                // in .Auto mode (SDK automatically checks if needs to use OCR or
not)
                extractor.OCRMode = OCRMode.Auto;

                // Set the location of 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;
            }
            else
            {
                Console.WriteLine("\nOCR Recommended: False");
            }
        }
    }
}

```

```
//Read all text
var allExtractedText = extractor.GetText();
Console.WriteLine("\nExtracted Text:\n{0}\n\n", allExtractedText);
}
}
}
}
```

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](#)