PDF to XML with images in C# with ByteScout PDF Extractor SDK

How to code PDF to XML with images in C#: How-To tutorial

Source code documentation samples provide quick and easy way to add a required functionality into your application. ByteScout PDF Extractor SDK was made to help with PDF to XML with images in C#. ByteScout PDF Extractor SDK is the SDK is designed to help developers with pdf tables and pdf 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 and other utilities.

This rich sample source code in C# for ByteScout PDF Extractor SDK includes the number of functions and options you should do calling the API to implement PDF to XML with images. This C# sample code should be copied and pasted into your application's code editor. Then just compile and run it to see how it works. Use of ByteScout PDF Extractor SDK in C# is also explained in the documentation included along with the product.

Free trial version of ByteScout PDF Extractor SDK is available on our website. Get it to try other samples for C#.

FOR MORE INFORMATION AND FREE TRIAL:

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

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
using System;
using Bytescout.PDFExtractor;

namespace PDF2XML
{
    class Program
    {
        static void Main(string[] args)
        {
            // Create Bytescout.PDFExtractor.XMLExtractor instance
            XMLExtractor extractor = new XMLExtractor();
            extractor.RegistrationName = "demo";
            extractor.RegistrationKey = "demo";

            // Load sample PDF document
            extractor.LoadDocumentFromFile("sample1.pdf");

            // Uncomment this line to get rid of empty nodes in XML
            //extractor.PreserveFormattingOnTextExtraction = false;

            // Set output image format
            extractor.ImageFormat = OutputImageFormat.PNG;

            // Save images to external files
            extractor.SaveImages = ImageHandling.OuterFile;
            extractor.ImageFolder = "images"; // Folder for external images
            extractor.SaveXMLToFile("result_with_external_images.xml");

            // Embed images into XML as Base64 encoded string
            extractor.SaveImages = ImageHandling.Embed;
            extractor.SaveXMLToFile("result_with_embedded_images.xml");

            // Cleanup
                    extractor.Dispose();
        }
    }
}
```

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](#)

[www.bytescout.com](#)