

How to convert PDF to XML with images with PDF extractor SDK in C# using ByteScout PDF Suite

Continuous learning is a crucial part of computer science and this tutorial shows how to convert PDF to XML with images with PDF extractor SDK in C#

Quickly learn how to convert PDF to XML with images with PDF extractor SDK in C# with this sample source code. ByteScout PDF Suite is the set that includes 6 SDK products to work with PDF from generating rich PDF reports to extracting data from PDF documents and converting them to HTML. This bundle includes PDF (Generator) SDK, PDF Renderer SDK, PDF Extractor SDK, PDF to HTML SDK, PDF Viewer SDK and PDF Generator SDK for Javascript and you can use it to convert PDF to XML with images with PDF extractor SDK with C#.

The following code snippet for ByteScout PDF Suite works best when you need to quickly convert PDF to XML with images with PDF extractor SDK in your C# application. IF you want to implement the functionality, just copy and paste this code for C# below into your code editor with your app, compile and run your application. Want to see how it works with your data then code testing will allow the function to be tested and work properly.

The trial version of ByteScout PDF Suite 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 Suite](#)

[Explore API Documentation](#)

[Get Free Training for ByteScout PDF Suite](#)

[Get Free API key for Web API](#)

[visit www.Bytescout.com](http://www.Bytescout.com)

Source Code Files:

```
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();
        }
    }
}
```

<https://www.youtube.com/watch?v=NEwNs2b9YN8>

ON-PREMISE OFFLINE SDK

[60 Day Free Trial](#) or [Visit ByteScout PDF Suite Home Page](#)
[Explore ByteScout PDF Suite Documentation](#)
[Explore Samples](#)
[Sign Up for ByteScout PDF Suite Online Training](#)

ON-DEMAND REST WEB API

[Get Your API Key](#)
[Explore Web API Docs](#)
[Explore Web API Samples](#)

[visit www.ByteScout.com](http://www.ByteScout.com)

[visit www.PDF.co](http://www.PDF.co)

www.bytescout.com