

How to find website addresses in PDF with regex in C# with ByteScout PDF Extractor SDK

What is ByteScout PDF Extractor SDK? It 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.

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

Source Code Files:

FindWebsiteAddressWithRegex.csproj

```
<?xml version="1.0" encoding="utf-8"?>
<Project ToolsVersion="15.0" xmlns="http://schemas.microsoft.com/developer/msbuild/2003">
  <Import Project="$(MSBuildExtensionsPath)\$(MSBuildToolsVersion)\Microsoft.Common.props" Condition="Exists('$(MSBuildExtensionsPath)\$(MSBuildToolsVersion)\Microsoft.Common.props')">
  </Import>
  <PropertyGroup>
    <Configuration Condition="'$(Configuration)' == ''">Debug</Configuration>
    <Platform Condition="'$(Platform)' == ''">AnyCPU</Platform>
    <ProjectGuid>{99735776-2956-463D-9795-EBCE16928C30}</ProjectGuid>
    <OutputType>Exe</OutputType>
    <RootNamespace>FindWebsiteAddressWithRegex</RootNamespace>
    <AssemblyName>FindWebsiteAddressWithRegex</AssemblyName>
    <TargetFrameworkVersion>v2.0</TargetFrameworkVersion>
  </PropertyGroup>
</Project>
```

```

<FileAlignment>512</FileAlignment>
</PropertyGroup>
<PropertyGroup Condition="'$(Configuration)|$(Platform)' == 'Debug|AnyCPU' ">
  <PlatformTarget>AnyCPU</PlatformTarget>
  <DebugSymbols>>true</DebugSymbols>
  <DebugType>full</DebugType>
  <Optimize>>false</Optimize>
  <OutputPath>bin\Debug\</OutputPath>
  <DefineConstants>DEBUG;TRACE</DefineConstants>
  <ErrorReport>prompt</ErrorReport>
  <WarningLevel>4</WarningLevel>
</PropertyGroup>
<PropertyGroup Condition="'$(Configuration)|$(Platform)' == 'Release|AnyCPU' ">
  <PlatformTarget>AnyCPU</PlatformTarget>
  <DebugType>pdbonly</DebugType>
  <Optimize>>true</Optimize>
  <OutputPath>bin\Release\</OutputPath>
  <DefineConstants>TRACE</DefineConstants>
  <ErrorReport>prompt</ErrorReport>
  <WarningLevel>4</WarningLevel>
</PropertyGroup>
<ItemGroup>
  <Reference Include="Bytescout.PDFExtractor, Version=9.1.0.3170, Culture=neutral, PublicKeyToken=f7dd1bd9d4"
    <SpecificVersion>False</SpecificVersion>
    <HintPath>c:\Program Files\Bytescout PDF Extractor SDK\net2.00\Bytescout.PDFExtractor.dll</HintPath>
  </Reference>
  <Reference Include="System" />
  <Reference Include="System.Data" />
  <Reference Include="System.Xml" />
</ItemGroup>
<ItemGroup>
  <Compile Include="Program.cs" />
</ItemGroup>
<ItemGroup>
  <None Include="SampleInvoice.pdf">
    <CopyToOutputDirectory>Always</CopyToOutputDirectory>
  </None>
</ItemGroup>
<Import Project="$(MSBuildToolsPath)\Microsoft.CSharp.targets" />
</Project>

```

Program.cs

```

using Bytescout.PDFExtractor;
using System;

namespace FindWebsiteAddressWithRegexp
{
    // Note: if you are looking for a more higher level API to extract data from invoices, reports, statements
    // then please check Document Parser SDK and Web API at https://bytescout.com/products/developer/documentpa
    // and https://pdf.co for secure and scalable web api

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

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

```


[Explore Web API Samples](#)

[visit www.ByteScout.com](#)

[visit www.PDF.co](#)

www.bytescout.com