

data masking in PDF in VB.NET and ByteScout PDF Extractor SDK

How to code data masking in PDF in VB.NET: How-To tutorial

Writing of the code to data masking in PDF in VB.NET can be done by developers of any level using ByteScout PDF Extractor SDK. ByteScout PDF Extractor SDK was made to help with data masking in PDF in VB.NET. 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.

You will save a lot of time on writing and testing code as you may just take the code below and use it in your application. This VB.NET sample code should be copied and pasted into your application's code editor. Then just compile and run it to see how it works. Enhanced documentation and tutorials are available along with installed ByteScout PDF Extractor SDK if you'd like to dive deeper into the topic and the details of the API.

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

VB.NET - Program.vb

```
Imports System.IO
Imports System.Text.RegularExpressions
Imports Bytescout.PDFExtractor

Module Program

    Sub Main()

        Try

            ' Generate CSVExtractor instance
            Using extractor As New CSVExtractor("demo", "demo")

                ' Load PDF document
                extractor.LoadDocumentFromFile("sample.pdf")

                ' Get all data
                Dim allData = extractor.GetCSV()

                ' Regular expressions and replacements
                Dim ssnRegex = "\d{3}[-]?\d{2}[-]?\d{4}"
                Dim ssnReplace = "***_**_****"
```

```
Dim phoneRegex = "\d{3}[-]?\d{3}[-]?\d{4}"
Dim phoneReplace = "***_***_****"

' Find and mask SSN and phone numbers
allData = Regex.Replace(allData, ssnRegex, ssnReplace)
allData = Regex.Replace(allData, phoneRegex, phoneReplace)

' Write as CSV
File.WriteAllText("output.csv", allData)

' Open file
Process.Start("output.csv")

End Using

Catch ex As Exception
    Console.WriteLine(ex.Message)
End Try

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

End Sub

End Module
```

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

or

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