

HOWTO: Extract Text from a Pdf Document

The PdfTextDocument class is the base class for extracting searchable text from an existing pdf file. Below is a method showing how to extract an entire page of text from the pdf file:

C#

```
using (PdfTextDocument doc = new PdfTextDocument(docpath)) { PdfTextPage page =  
doc.GetPage(0); string s = page.GetText(0, page.CharCount); MessageBox.Show(s); }
```

VB.NET

```
using doc As New PdfTextDocument(docpath) Dim page As PdfTextPage = doc.GetPage(0) Dim s As  
String = page.GetText(0, page.CharCount) MessageBox.Show(s) End Using
```

Original Article:

Q10291 - HOWTO: Extract Text from a Pdf Document

Atalasoftware Knowledge Base

<https://www.atalasoftware.com/kb2/KB/50184/HOWTO-Extract-Text-from-a-Pdf-Docume...>