PDF to Html Converter

User Manual of PDF to Html Converter-command line

Brief Introduction

PDF to HTML Converter Command Line can be used to convert PDF to HTML in batches. It can be used to process unattended conversion from large volumes PDF to HTML in batch mode under MS-DOS. With this software application, you can convert password protected PDF to HTML if you have the permission to do so. And it will output html file with SEO necessary elements. If you need, it can help you remove image or text parts and output html with the elements you need.

System requirement

Microsoft Windows 98/ME/NT/2000/XP/2003/Vista/7/2008 of both 32-bit and 64-bit.

How to Use

- **Download PDF to HTML Converter Command Line** and unzip it.
- Read readme.txt to know the usage and parameters.
- Open the MS-DOS interface to run the command line application.

Parameter List

Usage: pdf2html [options] <PDF-file> <HTML-file>

```
 -f <int>          : first page to convert
 -l <int>           : last page to convert
 -r <int>           : resolution, in DPI (default is 100)
 -quality <int> : set quality for JPEG file, default is 95
 -imgformat <int> : 0 is JPEG, 1 is PNG, 2 is GIF
 -opw <string>     : owner password (for encrypted PDF file)
 -upw <string>      : user password (for encrypted PDF file)
 -onehtm           : convert PDF to one continuous html page
 -oneword          : take single word as coordinate to create accurate HTML files
 -noimg            : remove images from HTML files
 -notxtidx         : remove text index file from HTML files
 -notextinbody      : remove text from HTML's body
 -notextinmeta      : remove text from HTML's meta
 -noseo             : don't optimize HTML files for search engines
 -homeurl <string>: add a home URL into the left index page
 -yoffset <int>     : set Y offset for HTML page contents
 -noutf8            : remove UTF8 header to compatible with Firefox
 -$ <string>        : input your License Key
```

Examples

```
pdf2html.exe C:\in.pdf C:\out.htm
```

```
pdf2html.exe -r 150 C:\in.pdf C:\out.htm
pdf2html.exe -imgformat 1 C:\in.pdf C:\out.htm
pdf2html.exe -noimg C:\in.pdf C:\out.htm
pdf2html.exe -onehtm C:\in.pdf C:\out.htm
pdf2html.exe -oneword C:\in.pdf C:\out.htm
pdf2html.exe -homeurl "http://www.verypdf.com" C:\in.pdf C:\out.htm
pdf2html.exe -yoffset 20 C:\in.pdf C:\out.htm
pdf2html.exe -notextinbody -notextinmeta C:\in.pdf C:\out.htm
```

Important Parameter Explanation

-oneword

Output HTML files will remain the original formatting and interrelation of text and graphical elements on each page by default. But sometimes if you need to output HTML files with accurate positions, please use the parameter "-oneword" because it will use single word as coordinate to generate HTML files for easy searching and editing.

-noseo

Parameter "-noseo" is used to remove the SEO necessary elements such as keywords and title, which are contained in each output HTML by default. Though such elements can help your webiste to be found by search engines, sometimes, they might be redundant according to your basic needs.

-imgformat

When converting PDF files with images, by this software you can choose output image in which kind of format, like jpg, png or gif. That depends on your need, as ".png" is small in size, gif is good when making video.
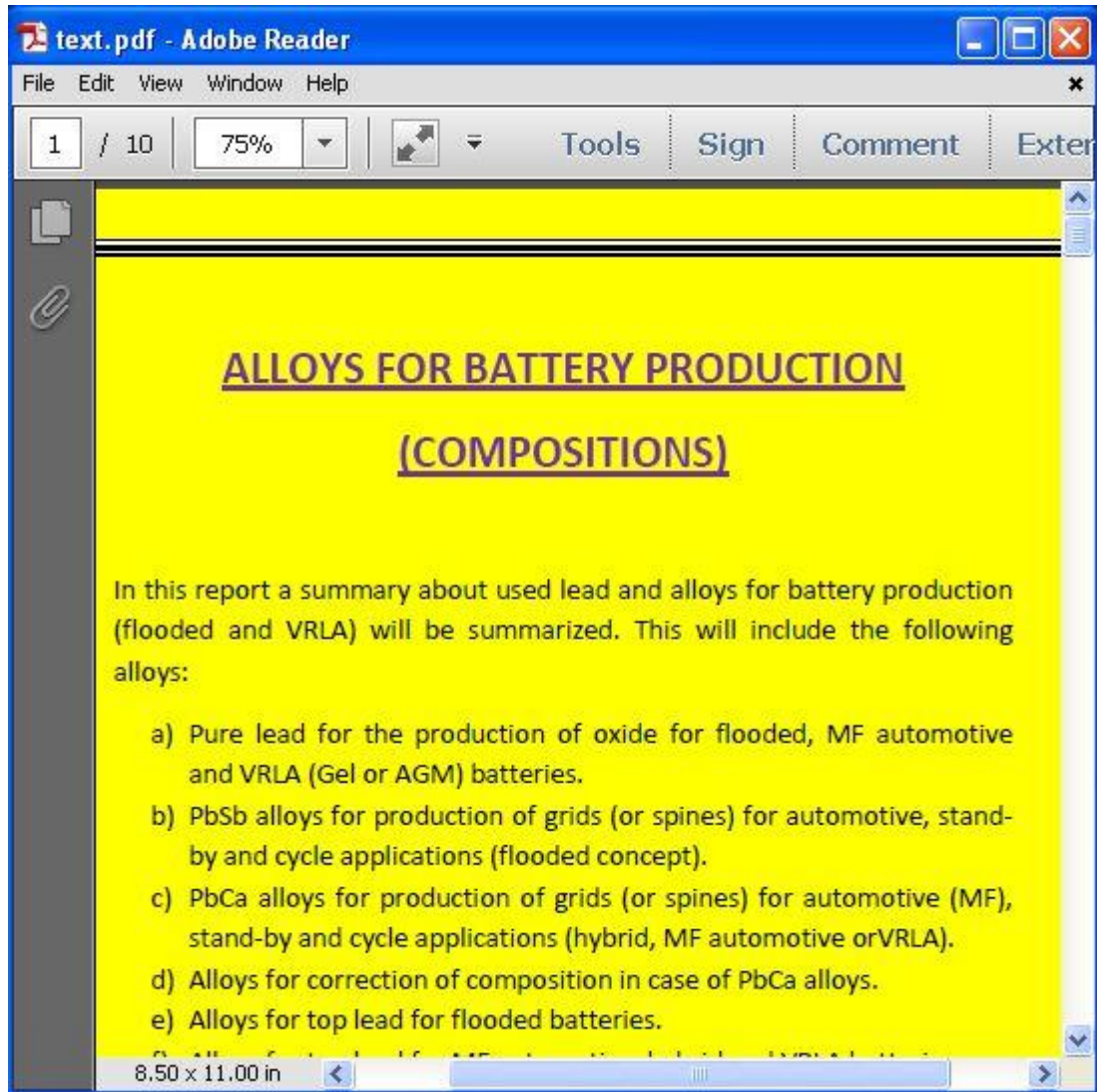
-homeurl

When you need to add your home URL at the left index, please use this parameter. Like: *pdf2html.exe -homeurl "http://www.verypdf.com" C:\in.pdf C:\out.htm,* by this command line, you can add VeryPDF website URL at the left index page.

Demo

Input PDF file

Output Html File
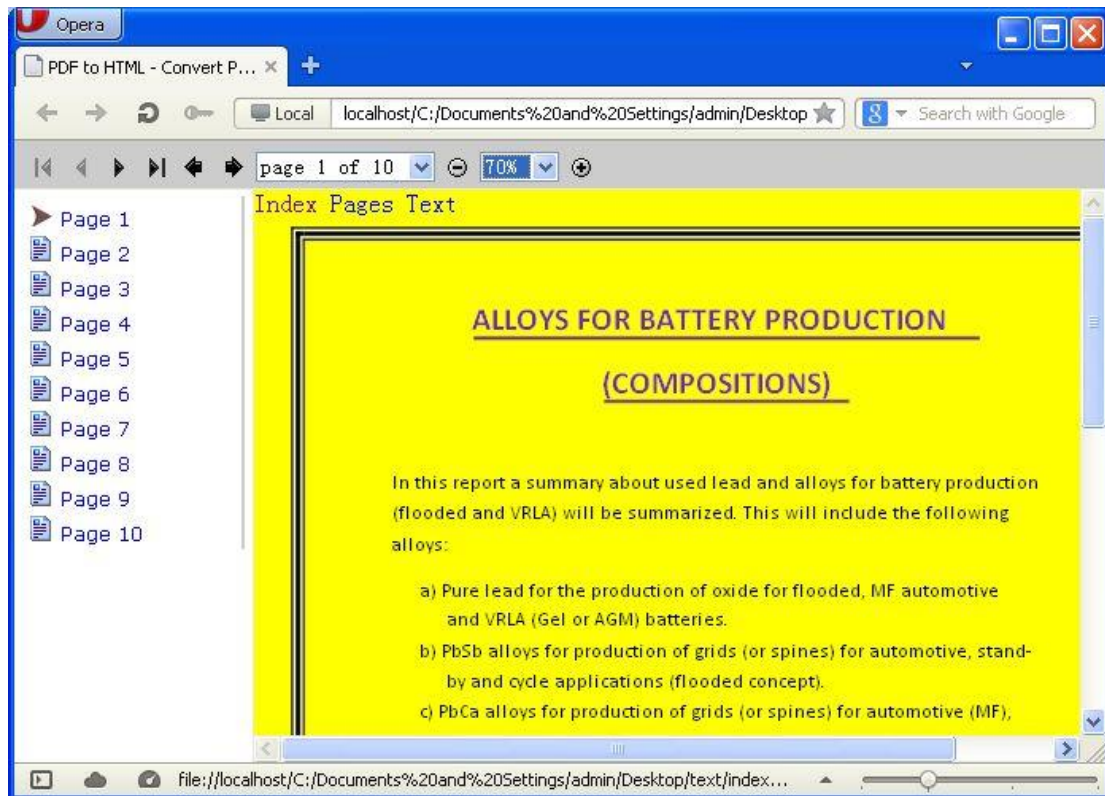
Contact US

During the using, if you have any question, please send us a ticket to support@verydoc.com. We will solve the problem at the first time.