

NAME

pdftohtml – Portable Document Format (PDF) to HTML converter (version 3.04)

SYNOPSIS

pdftohtml [options] *PDF-file HTML-dir*

DESCRIPTION

Pdftohtml converts Portable Document Format (PDF) files to HTML.

Pdftohtml reads the PDF file, *PDF-file*, and places an HTML file for each page, along with auxiliary images in the directory, *HTML-dir*. The HTML directory will be created; if it already exists, pdftohtml will report an error.

CONFIGURATION FILE

Pdftohtml reads a configuration file at startup. It first tries to find the user's private config file, *~/xpdfrc*. If that doesn't exist, it looks for a system-wide config file, typically */usr/local/etc/xpdfrc* (but this location can be changed when pdftohtml is built). See the **xpdfrc(5)** man page for details.

OPTIONS

Many of the following options can be set with configuration file commands. These are listed in square brackets with the description of the corresponding command line option.

-f *number*

Specifies the first page to convert.

-l *number*

Specifies the last page to convert.

-r Specifies the resolution, in DPI, for background images. The default is 150 DPI.

-opw *password*

Specify the owner password for the PDF file. Providing this will bypass all security restrictions.

-upw *password*

Specify the user password for the PDF file.

-q Don't print any messages or errors. [config file: **errQuiet**]

-cfg *config-file*

Read *config-file* in place of *~/xpdfrc* or the system-wide config file.

-v Print copyright and version information.

-h Print usage information. (**-help** and **--help** are equivalent.)

BUGS

Some PDF files contain fonts whose encodings have been mangled beyond recognition. There is no way (short of OCR) to extract text from these files.

EXIT CODES

The Xpdf tools use the following exit codes:

- | | |
|----|-----------------------------------|
| 0 | No error. |
| 1 | Error opening a PDF file. |
| 2 | Error opening an output file. |
| 3 | Error related to PDF permissions. |
| 99 | Other error. |

AUTHOR

The pdftohtml software and documentation are copyright 1996-2014 Glyph & Cog, LLC.

SEE ALSO

xpdf(1), **pdftops(1)**, **pdftotext(1)**, **pdfinfo(1)**, **pdffonts(1)**, **pdfdetach(1)**, **pdftoppm(1)**, **pdftopng(1)**, **pdfimages(1)**, **xpdfrc(5)**

