

Extension:PdfHandler


From MediaWiki.org

This is the README file for the **PdfHandler extension for MediaWiki** software. The extension is only useful if you've got a MediaWiki installation; it can only be installed by the administrator of the site.

The extension shows uploaded pdf files in a multipage preview layout. With enabled WebStore the extension generates automatically images from the specified page.

This is the first version of the extension and it's almost sure to have bugs. See the BUGS section below for info on how to report problems.

MediaWiki extensions manual - list

 <h3>PdfHandler</h3> <p>Release status: stable</p>	
Implementation	Media
Description	Handles PDF files like multipage DJVU
Author(s)	xarax <jodeldi at gmx dot de>
MediaWiki	1.11+
License	GPL
Download	<p>Download snapshot Subversion (http://svn.wikimedia.org/svnroot/mediawiki/trunk/extensions/PdfHandler/) ^[Help]</p> <p>Browse source code (http://svn.wikimedia.org/viewvc/mediawiki/trunk/extensions/PdfHandler/) View code changes (http://www.mediawiki.org/w/index.php?title=Special:Code/MediaWiki&path=%2Ftrunk%2Fextensions%2FPdfHandler%2F)</p>
Example	<p>usability.wikimedia.org example</p> <hr/> <p>j-crew.de example (http://spiele.j-crew.de/wiki/Datei:13q_gro.pdf)</p>
<p>Check usage (http://s23.org/wikistats/mediawiki_extensions.php?name=Extension:PdfHandler) (<i>experimental</i>)</p>	

Contents

- 1 License
- 2 Pre-requisites
- 3 Installation
- 4 Usage
- 5 Bugs and enhancements
 - 5.1 Enhancements
- 6 See also

License

Copyright 2007 xarax (s:de:Benutzer:Xarax)

This program is free software; you can redistribute it and/or modify it under the terms of the GNU General Public License as published by the Free Software Foundation; either version 2 of the License, or (at your option) any later version.

This program is distributed in the hope that it will be useful, but WITHOUT ANY WARRANTY; without even the implied warranty of MERCHANTABILITY or FITNESS FOR A PARTICULAR PURPOSE. See the GNU General Public License for more details.

You should have received a copy of the GNU General Public License along with this program; if not, write to the Free Software Foundation, Inc., 59 Temple Place, Suite 330, Boston, MA 02111-1307 USA

Pre-requisites

This software was tested with 1.11alpha (r25574) – 1.12alpha (r28001) – 1.13.0rc1, and also 1.16wmf forward. It may or may not work with earlier versions, but please test it.

It requires the following packages:

- `gs-gpl` for `gs` to render the page images (<http://www.ghostscript.com/>)
- `imagemagick` for convert to resize the extracted images (<http://www.imagemagick.org/script/index.php>)
- `xpdf-utils` for `pdfto` to extract metadata from pdf (<http://www.foolabs.com/xpdf/download.html>). The `poppler-utils` package may be substituted for `xpdf-utils` on Ubuntu and Debian systems.

It needs at least PHP 5.1.3 to work (dependency on `SimpleXMLElement::addChild`).

Installation

To install, copy all the files in the archive you downloaded to the PdfHandler subdirectory of the extensions subdirectory of your MediaWiki installation.

In your MediaWiki `LocalSettings.php`, add the following line some place towards the bottom of the file:

```
require_once( "$IP/extensions/PdfHandler/PdfHandler.php" );
```

Then you can set some variables in your `LocalSettings.php` (if you wish to override one of the defaults):

- `$wgPdfProcessor` – path to your ghostscript implementation (default to `gs`)
- `$wgPdfPostProcessor` – path to your imagemagick convert (default to `convert`)
- `$wgPdfInfo` – path to your pdfto (default to `pdfto`)
- `$wgPdfOutputExtension` – defaults to `jpg`
- `$wgPdfHandlerDpi` – defaults to 150 dpi
 - The extension extracts a bitmap image for each page of the PDF, using this resolution (dpi = dots per inch). For example, a PDF page with the European size A4 is 210 mm wide, corresponding to 595 points (at 72 dpi). This yields an image 1240 pixels wide (at 150 dpi). If instead this parameter is set to 300 dpi, the width will be 2480 pixels.
- `$wgPdfCreateThumbnailsInJobQueue` - defaults to `false`
 - Put creating pages' thumbnails into a job queue, so they do not have to be created while browsing a file page, but during normal wikibrowsing. **Be advised** that setting this to `true` may significantly increase CPU load of your webserver on a high-traffic website. Job queue was designed to perform quick tasks on page views, and creating thumbnails can be included into that queue, according to this definition. Nevertheless, as quick as it is, it also requires certain CPU load to convert PDF pages to another format. The solution to that would be either setting `$wgJobRunRate` to rather small value (ie. 0.05) or disabling job queue on the high-traffic wiki and setting up another server to do just job queue (like Wikimedia did).

Variables below are not specific to this extension:

- Enable PDF uploads, if you haven't already: `$wgFileExtensions[] = 'pdf';`
- `$wgMaxShellMemory` - memory limit for `gs`, `convert` and `pdfto`. The default value of 102400 kB (100 MB) might be too low.

Example for a WindowsXP installation:

```
$wgPdfProcessor = 'C:\Programme\gs\gs8.60\bin\gswin32.exe';
$wgPdfPostProcessor = $wgImageMagickConvertCommand;
$wgPdfInfo = 'C:\Programme\xpdf-3.02p11-win32\pdfto.exe';
$wgPdfToText = 'C:\Programme\xpdf-3.02p11-win32\pdftotext.exe';
```

If `$wgImageMagickConvertCommand` is not already defined, use the definition alike

```
$wgPdfPostProcessor = 'C:\Programme\ImageMagick-6.6.2-Q16\convert.exe';
```

Example for a Ubuntu installation:

```
$wgPdfProcessor = 'gs';
$wgPdfPostProcessor = $wgImageMagickConvertCommand;
```

```
$wgPdfInfo = 'pdfinfo';
```

Ubuntu Note: You must have already installed the following 3 packages; imagemagick ghostscript xpdf-utils

Usage

- The main usage of the PdfHandler extension is without user interaction. If you upload a new pdf file, the metadata will be stored in the database, and then this file can be shown in a multipage preview layout like the djvu handler does. ~~The extension does not render PDF files that were uploaded prior to installation; they will be displayed in the same way as before.~~

The latter is apparently no longer true; existing files also get a preview. I tested it on my wiki, and also recall this stated in bugzilla/maillinglist/wherever, although I can't seem to locate the reference --Dror Snir 00:40, 25 August 2010 (UTC)

- Another option, introduced quite long ago (r25575 (http://svn.wikimedia.org/viewvc/mediawiki/trunk/extensions/PdfHandler/PdfHandler_body.php?revision=25575&view=markup)), is to use it to display PDF files as an image, showing a single page at a time, like so: [[File:myPdfFile.pdf|page=1|600px]]. The page and size parameters are optional; the default page is page #1. Instead of a size-parameter, you can also use the thumb-parameter, with or without captions: [[File:myPdfFile.pdf|page=1|thumb|My PDF]].

If you would like to present a 2-page pdf, for example, do the following: [[File:myPdfFile.pdf|page=1]]
[[File:myPdfFile.pdf|page=2]]

Bugs and enhancements

Bugs can be reported at Bugzilla (<http://bugzilla.wikimedia.org>)

Bug list:


- PdfHandler bug reports (https://bugzilla.wikimedia.org/buglist.cgi?query_format=advanced&bug_status=UNCONFIRMED&bug_status=NEW&bug_status=ASSIGNED&bug_status=REOPENED&product=MediaWiki%20extensions), all of them, old and new

Enhancements

- There is a patch to run a background job for thumbnails to be created. Thanks to that, the thumbnails are created during normal browsing through wiki pages, instead of during browsing the PDF file page. The patch saves time by creating one thumbnail at the time, not causing any load deficiencies, unlike rendering them on PDF file pages.

See also

- Extension:Proofread Page (could be use with PdfHandler)
- Manual:How to use DjVu with MediaWiki

	<p>This extension is being used on one or more of Wikimedia's wikis. This means that the extension is stable and works well enough to be used by such high traffic websites. A full list of the extensions installed on a particular wiki can be seen on the wiki's Special:Version page.</p>
---	---

Retrieved from "<http://www.mediawiki.org/w/index.php?title=Extension:PdfHandler&oldid=478125>"

Categories: Stable extensions | Media handling extensions | Extensions in Wikimedia version control | All extensions | Extensions used on Wikimedia | UploadVerifyFile extensions

- This page was last modified on 2 January 2012, at 10:23.
- Text is available under the Creative Commons Attribution/Share-Alike License; additional terms may apply. See Terms of use for details.