









Installation of the Attachment Indexing plugins

Article Number: 47 | Rating: 3.8/5 from 13 votes | Last Updated: Thu, Aug 22, 2013 at 7:33 AM

PHPKB knowledge base software is able to index text content of the files attached to knowledge base articles in order to make them searchable. Attachment indexing is supported in all editions of PHPKB knowledge base software. Some documents types can be searched without any additional tools, others need PHP modules enabled or third party tools (plugins) installed. All these modules (plugins) and 3rd party tools are free.

List of Supported File Types

File Type	Supported Format	Required Tool (Plug-in)
 MS Office 2003 Word Document	.DOC	AntiWord (free) is required
 MS Office 2003 Excel Workbook	.XLS	xlhtml (free) is required
 MS Office 2003 PowerPoint Presentation	.PPT	pphtml (free) is required
 MS Office 2007 Word Document	.DOCX	PHP ZIP library is required
 MS Office 2007 Excel Workbook	.XLSX	PHP ZIP library is required
 MS Office 2007 PowerPoint Presentation	.PPTX	PHP ZIP library is required
 Adobe PDF Documents	.PDF	pdftotext (free) is required
 Plain-text Documents	.TXT, .HTM, .HTML, .CSV, .XML	No Plugin Required

Installation of Plugins to Search Attached Files on Windows Server

We strongly recommend you to use the latest version of PHP. These plugins work under PHP 5.3+ correctly. Earlier versions of PHP have [bugs](#) and may freeze when launch external programs (e.g. attachment indexation plugins) using Windows command line.

[Download the latest PHP package for Windows](#) (VC9 x86 Non Thread Safe is recommended).

PHP 5.3 doesn't support ISAPI anymore. So you need to use FastCGI instead.

1. Enabling Required PHP Extensions (Modules) on Windows Server

You need to enable certain PHP modules in order to index content of of MS Office 2007 documents.

1. Find the "ext" subdirectory of your PHP installation (it is C:\PHP\ext\ by default).
2. Check if the following files exist in that folder: php_mbstring.dll, php_zip.dll.
3. If any of these files do not exist, you should run PHP installation and install appropriate modules (Mbstring or PHP ZIP respectively).
Note: If you have PHP 5.3 or higher, you need not to enable PHP ZIP extension as it is already built-in to the PHP engine.
4. Open the php.ini configuration file of your PHP engine in any text editor such as notepad.
5. Search for the "extension=" (without quotes).
6. You'll find the section with the list of PHP extensions. Some of them are commented with # symbol.
7. You should enable these modules by removing comment symbol (#).
8. Save php.ini file.
9. Restart the web server for changes to take effect.

2. Attachment Indexing Plugins Installation on Windows Server

1. The "phpkb/admin/" folder contains the following plugins:
 - o xlhtml (required for Excel 2003 files)
 - o pphtml (required for PowerPoint 2003 files)
 - o pdftotext (required for PDF files)
2. If your server runs under IIS, you should grant the Internet Guest Account with the read and execute permissions to "admin" folder. If it is Apache server, you should check which user runs the Apache process and grant read and execute permissions for that user.
3. There is a folder called "antiword" under "admin" folder of your PHPKB installation. You need to copy that folder to "C:\" of your server so that its path becomes "C:\antiword\".
4. Add read and execute permissions for the web-server to this folder and its contents as well.

Installation of Plugins to Search Attached Files on Linux/Unix Server

Please follow the instructions below to install/enable the attachment indexation plugins and required PHP extensions.

1. Enabling Required PHP Extensions (Modules) on Linux Server

1. Open the php.ini configuration file.
2. Find the "extension_dir" parameter. It indicates the path to PHP extensions directory. Go to that directory and check that zip.so file exists there.
3. Add reference for the PHP ZIP extension to the php.ini:

```
extension=zip.so
```

4. Restart the web server for changes made in php.ini to take effect.

2. Attachment Indexing Plugins Installation on Linux Server

If your system (RedHat / Fedora / CentOS) supports [Yum Package Manager](#) you can run this command instead to install necessary modules:

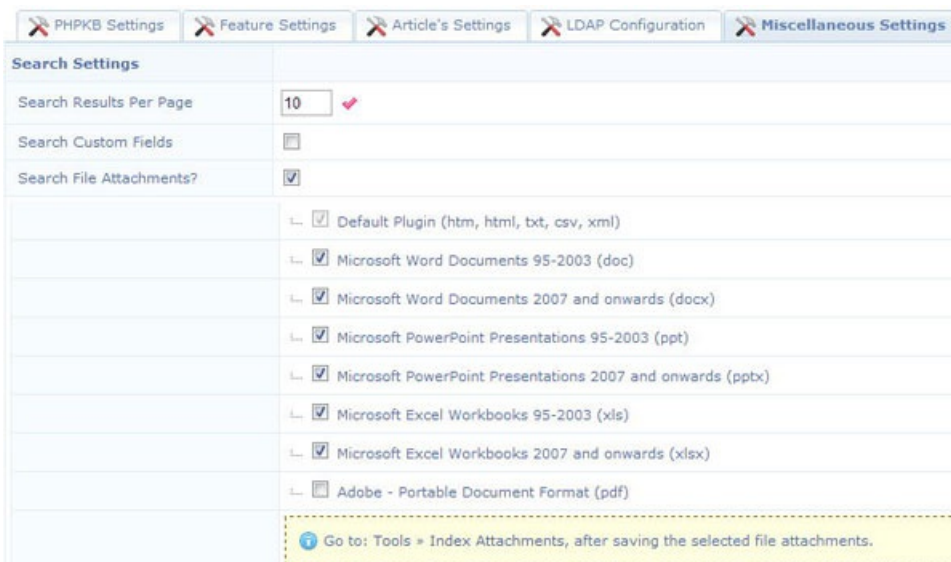
```
# yum install poppler-utils
```

OR run following command to install necessary modules on a system that has APT library (e.g. Ubuntu, Debian):

```
$ sudo apt-get install poppler-utils
```

What to do after installation of indexing plug-ins?

1. Go to the PHPKB admin control panel.
2. Go to the Tools » Manage Settings » Miscellaneous Settings.
3. Set check box on the "Search File Attachments" and on each document type as shown in the image below.



4. Click on "Save Changes" button to save the settings.

How to enable automatic indexing of file attachments when they are uploaded?

If you would like to auto-index the attached files as soon as they are uploaded, then there is another setting under "File Upload Settings" section of "Miscellaneous Settings". You can set the check box for "Index Attachments" as shown in image below.



Now you can upload attachments and they will be automatically indexed for search.

How to index attached files manually?

You can also manually run indexation for existing file attachments whenever required from "Tools" » "Index Attachments" section of the admin control panel as show below.



How does PHPKB searches the content of attached PDF Files?

PHPKB knowledge base software is able to index the text content of PDF documents and make them searchable. It converts a PDF file to text file format to search its contents. It uses "pdftotext utility" to convert Portable Document Format (PDF) files to plain text. It reads the PDF file, and writes a text file thus making itself able to search within the contents of PDF documents.

How to Disable Indexing for Specific File Types?

It may be necessary to disable attachment indexing for specific file types if conversion is unable to process the attachments due to size or other error. When disabled, conversion will be bypassed for only specified file types. Attachment indexing can also be disabled for all types if desired.

1. Go to the PHPKB admin control panel.
2. Go to the Tools » Manage Settings » Miscellaneous Settings.
3. Remove check box on the "Search File Attachments" if you would like to disable indexing for all file types.
4. Remove check box for specific file types if you would to disable indexing for those file types.

Posted by: Knowledge Base Administrator - Tue, Feb 12, 2013 at 2:35 AM. This article has been viewed 83268 times.

Online URL: <https://www.knowledgebase-script.com/kb/article/installation-of-the-attachment-indexing-plugins-47.html>