

Foxit PDF IFilter 1.0

For Microsoft Office SharePoint Server

User Manual

© 2009 Foxit Corporation.

All rights reserved.

Copyright © 2009 Foxit Corporation. All Rights Reserved.

No part of this document can be reproduced, transferred, distributed or stored in any format without the prior written permission of Foxit.

Anti-Grain Geometry - Version 2.3

Copyright (C) 2002-2005 Maxim Shemanarev (<http://www.antigrain.com>)

Permission to copy, use, modify, sell and distribute this software is granted provided this copyright notice appears in all copies. This software is provided "as is" without express or im-plied warranty, and with no claim as to its suitability for any purpose.

Contents

License Agreement for Foxit PDF IFilter	4
Overview	5
Why PDF IFilter?.....	5
Why Foxit PDF IFilter?	5
How Foxit PDF IFilter works?	6
Key benefits	6
Getting Started	7
System requirements.....	7
Installing Foxit PDF IFilter.....	7
Updating Foxit PDF IFilter	9
Uninstalling Foxit PDF IFilter	9
Working with Foxit PDF IFilter	10
Indexing PDFs.....	10
Searching PDFs	14
More Help.....	15
Additional Information	16
Microsoft Developer Support.....	16
Language Support	16
FAQ	17
Contact Us	20

License Agreement for Foxit PDF IFilter

By using software of Foxit Corporation or its subsidiaries ("Foxit"); you agree to the following terms and conditions. If you do not agree with such terms and conditions; do not use the software. The terms of an end user license agreement accompanying a particular software file upon installation or download of the software shall supersede the terms presented below.

If the software is designed for use with an application software product (the "Host Application") published by Foxit; Foxit grants you a non-exclusive license to use such software with the Host Application only; provided you possess a valid license from Foxit for the Host Application. Except as set forth below; such software is licensed to you subject to the terms and conditions of the End User License Agreement from Foxit governing your use of the Host Application.

Foxit PDF IFilter can be used on desktop environment for free. However, users need to acquire additional license when they use Foxit PDF IFilter on server environment, e.g. Microsoft SQL Server.

DISCLAIMER OF WARRANTIES: YOU AGREE THAT FOXIT HAS MADE NO EXPRESS WARRANTIES TO YOU REGARDING THE SOFTWARE AND THAT THE SOFTWARE IS BEING PROVIDED TO YOU "AS IS" WITHOUT WARRANTY OF ANY KIND. FOXIT DISCLAIMS ALL WARRANTIES WITH REGARD TO THE SOFTWARE; EXPRESS OR IMPLIED; INCLUDING; WITHOUT LIMITATION; ANY IMPLIED WARRANTIES OF FITNESS FOR A PARTICULAR PURPOSE; MERCHANTABILITY; MERCHANTABLE QUALITY OR NONINFRINGEMENT OF THIRD PARTY RIGHTS. Some states or jurisdictions do not allow the exclusion of implied warranties; so the above limitations may not apply to you.

LIMIT OF LIABILITY: IN NO EVENT WILL FOXIT BE LIABLE TO YOU FOR ANY LOSS OF USE; INTERRUPTION OF BUSINESS; OR ANY DIRECT; INDIRECT; SPECIAL; INCIDENTAL; OR CONSEQUENTIAL DAMAGES OF ANY KIND (INCLUDING LOST PROFITS) REGARDLESS OF THE FORM OF ACTION WHETHER IN CONTRACT; TORT (INCLUDING NEGLIGENCE); STRICT PRODUCT LIABILITY OR OTHERWISE; EVEN IF FOXIT HAS BEEN ADVISED OF THE POSSIBILITY OF SUCH DAMAGES. Some states or jurisdictions do not allow the exclusion or limitation of incidental or consequential damages; so the above limitation or exclusion may not apply to you. Foxit's liability is limited to replacement or refund of only the virus-infected software downloaded from Foxit website.

Overview

Welcome to Foxit PDF IFilter 1.0 -- a professional tool that helps you to index PDF files.

Foxit PDF IFilter is an application designed to help users index a large amount of PDF documents and then quickly find text within these documents. The PDF documents can not only be files, but also email attachments or database records.

Foxit PDF IFilter supports PDF 1.7 specification and can be used in many Microsoft search engines, such as Windows Indexing Service, MSN Desktop Search, Internet Information Server, SharePoint Portal Server, Windows SharePoint Services (WSS), Site Server, Exchange Server, SQL Server and all other products based on Microsoft Search technology. With Foxit PDF IFilter, all these search engines can find PDF files with ease.

Why PDF IFilter?

PDF IFilter is necessary for several reasons listed below:

- Microsoft creates and maintains office file formats, such as .doc, .xls, .ppt, etc. However, many other file formats are created and maintained by other companies. For example, PDF was created by Adobe.
- With the advantages of security, reliability and accessibility, PDF is now being used by more and more people all over the world.
- PDF IFilter is a plug-in that allows Microsoft search engines to index PDF files so that they become searchable. Without an appropriate PDF IFilter, contents of a PDF file cannot be indexed.

Why Foxit PDF IFilter?

Foxit PDF IFilter is favored by a large number of government organizations and well-known companies because of the following reasons:

- Foxit PDF IFilter helps to do a two-stage search – Index and Search, which makes finding information much easier.
- Foxit PDF IFilter supports multi-threading and makes full use of multi-CPU/core in a server.
- Foxit PDF IFilter supports native 64-bit code.
- Smaller and faster than other PDF IFilters.

- Supports Chinese/Japanese/Korean PDF documents.

How Foxit PDF IFilter works?

Foxit PDF IFilter acts as a plug-in for full-text search engines. A search engine usually works in two steps:

Step 1, the search engine goes through a designated place, e.g. a file folder or a database, and indexes all documents or newly modified documents, including PDF documents, in the background and creates internal data to store indexing result.

Step 2, a user specifies some keywords he/she would like to search and the search engine answers the query immediately by looking up the indexing result and responds to the user with all the documents that contains the keywords.

During Step 1, the search engine itself doesn't understand format of a PDF document. Therefore, it looks in windows registry for an appropriate PDF IFilter and finds the Foxit PDF IFilter. Foxit PDF IFilter understands PDF format. It filters out embedded formatting and extracts text from the document and return text back to the search engine. Further information about IFilter specification can be found on the [Microsoft Web site](#).

Key benefits

Here's a list of benefits that Foxit PDF IFilter brings to you:

- Integrates with existing operating systems and tools within your company.
- Provides an easy solution to search within PDF documents located on local computer, local network and intranet.
- Greatly increases your ability to accurately locate information.
- Much smaller and faster than IFilter offered by other vendors.
- Supports Chinese/Japanese/Korean PDF documents.
- Supports native 64-bit code.
- Supports multi-threading and takes full use of multiple-CPU in a server.

Getting Started

This section provides everything you'll need to know about getting started with Foxit PDF IFilter, including its system requirement, installation and uninstallation.

System requirements

Foxit PDF IFilter 1.0 requires the following environment:

- Microsoft Windows 2003 Server

In addition, the environment must also contain Microsoft Office SharePoint Server.

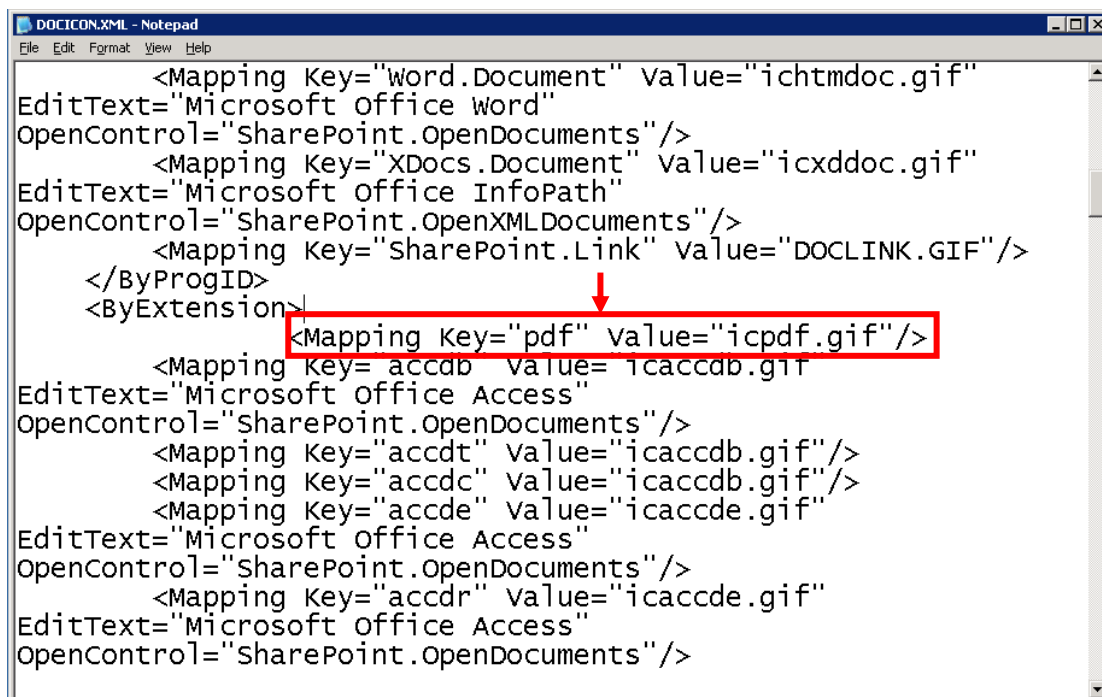
Installing Foxit PDF IFilter

Foxit PDF IFilter has two versions respectively for 32-bit Windows and 64-bit Windows. You can choose one of them to install into your machine according to your system requirements. Before installing Foxit PDF IFilter, please note that Microsoft Office SharePoint Server has been installed successfully.

To install Foxit PDF IFilter 1.0, please do the following:

- Download Foxit PDF IFilter from Foxit Website (<http://www.foxitsoftware.com>) onto the machine where you wish to install it.
- Stop the IIS Admin service: Start > Settings > Control Panel > Administrator Tools > Services > IIS Admin Service > Stop. Close window.
- Run the Foxit PDF IFilter 1.0 Setup program to install the filter on the server.
- Download PDF icon from <http://www.foxitsoftware.com/pdf/ifilter/pdficon.gif>.
- Copy the downloaded .GIF file to "Drive:\Program Files\Common Files\Microsoft Shared\Web Server Extensions\12\Template\Images".
- Edit the file "Drive:\Program Files\Common Files\Microsoft Shared\Web server extensions\12\Template\Xml\DOCICON.XML":
 1. Right-click DOCICON.XML file > click Open With > select Notepad.
 2. Add an entry for the .pdf extension. For example, where ICPDF is the name of the .gif file:

```
<Mapping Key="pdf" Value="icpdf.gif"/>
```



```

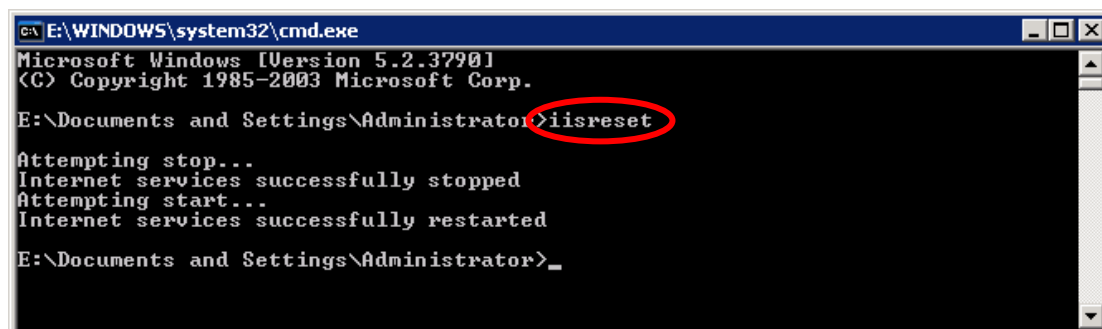
DOCICON.XML - Notepad
File Edit Format View Help
    <Mapping Key="word.Document" value="ichtmdoc.gif"
EditText="Microsoft Office word"
OpenControl="SharePoint.OpenDocuments"/>
    <Mapping Key="XDocs.Document" value="icxddoc.gif"
EditText="Microsoft Office InfoPath"
OpenControl="SharePoint.OpenXMLDocuments"/>
    <Mapping Key="SharePoint.Link" value="DOCLINK.GIF"/>
  </ByProgID>
  <ByExtension>
    <Mapping Key="pdf" value="icpdf.gif"/>
    <Mapping Key="accdb" value="icaccdb.gif"
EditText="Microsoft Office Access"
OpenControl="SharePoint.OpenDocuments"/>
    <Mapping Key="accdt" value="icaccdb.gif"/>
    <Mapping Key="accdc" value="icaccdb.gif"/>
    <Mapping Key="accde" value="icaccede.gif"
EditText="Microsoft Office Access"
OpenControl="SharePoint.OpenDocuments"/>
    <Mapping Key="accdr" value="icaccede.gif"
EditText="Microsoft Office Access"
OpenControl="SharePoint.OpenDocuments"/>

```

3. On the File menu, click Save, and then quit Notepad.

- Perform an iisreset:

1. Click Start > Run > type "cmd" in the Open line> click OK > type "iisreset" at the command prompt > hit Enter.



```

C:\ E:\WINDOWS\system32\cmd.exe
Microsoft Windows [Version 5.2.3790]
(C) Copyright 1985-2003 Microsoft Corp.

E:\Documents and Settings\Administrator>iisreset

Attempting stop...
Internet services successfully stopped
Attempting start...
Internet services successfully restarted

E:\Documents and Settings\Administrator>_

```

2. Close the window.

- Stop and start the Sharepoint Search Service by running the following commands:
 1. net stop osearch
 2. net start osearch
 3. net stop spsearch
 4. net start spsearch

Notes:

1. Foxit PDF IFilter appears in the list of programs on the Programs Start Menu after installation.

2. *Foxit PDF IFilter has no user interface.*
3. *The installation package will unzip a language file called `fpdfcjk.bin` into the installation folder, which will help you to index PDF documents containing Chinese/Japanese/Korean characters.*

Updating Foxit PDF IFilter

When updating Foxit PDF IFilter, you need to uninstall the old version of Foxit PDF IFilter to install the new one. You do not have to reboot your machine after update.

Generally speaking, Foxit PDF IFilter consists of a core layer and an installation layer. When you upgrade the installation layer, you do not need to re-do a full crawl. But when you upgrade the core layer, you can decide if you need to rebuild index on your computer according to the importance of the updates by checking what's new details on Foxit Official Website.

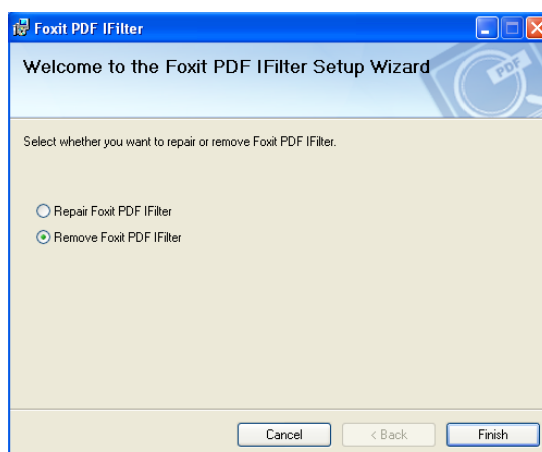
Uninstalling Foxit PDF IFilter

To uninstall Foxit PDF IFilter, please perform the following:

- Open the Start Menu > Settings > select "Control Panel".
- Click on the "Add or Remove Programs" tab.
- Click on the "Foxit PDF IFilter" record and then on the "Remove" button to uninstall it.

Or you can double-click the Foxit PDF IFilter.msi setup file that you have used to install Foxit PDF IFilter, and do the following:

- In the Foxit PDF IFilter dialog box, select Remove Foxit PDF IFilter.



- Click Finish button to uninstall it completely.

Working with Foxit PDF IFilter

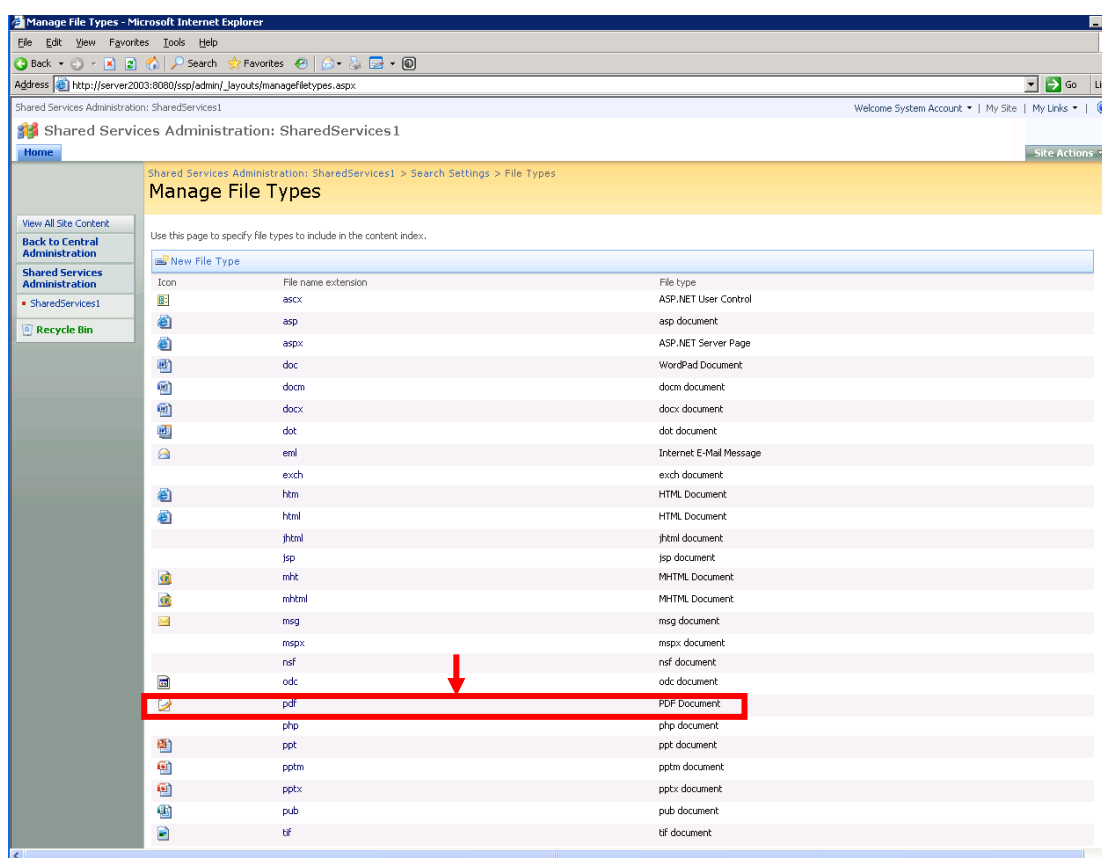
Foxit PDF IFilter is able to search text in PDF files. After you install Foxit PDF IFilter, you should upload your PDF documents and perform a full crawl with Microsoft Office SharePoint Server (MOSS) before you start a searching.

You don't have to do full crawls every time you add some new PDF files. You may schedule the incremental crawls for all the new ones.

Indexing PDFs

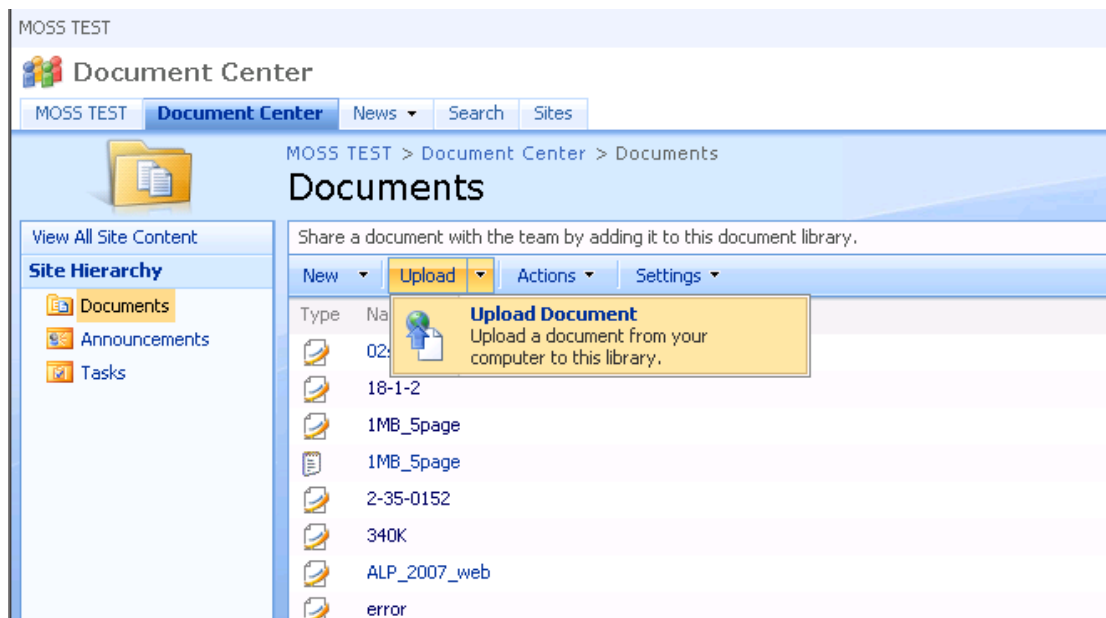
To index PDF files on your computer, please do the following:

- Add the .pdf file type to the index list:
 1. Go to the **Shared Service Administration: SharedServices 1** home page, click **Search Settings** link > under **Crawl Settings** click **File Types**.
 2. Click **New File Type** > type "pdf" in the **File extension** field > click **OK**.

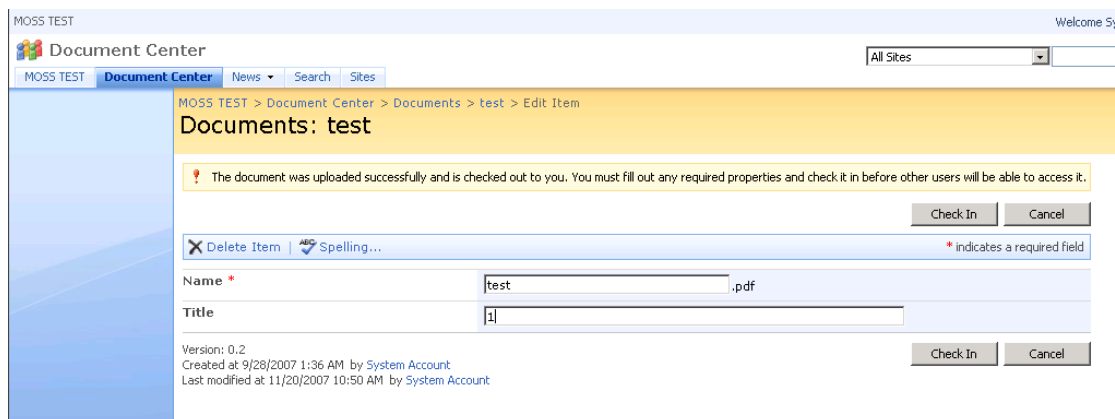


The PDF file type has been added

- Upload a PDF file, and perform a Full Crawl on the Search content indexes:
 1. Go to the **Document Center**, and click **Documents**.
 2. Click **Upload**, and then **Upload Document**.

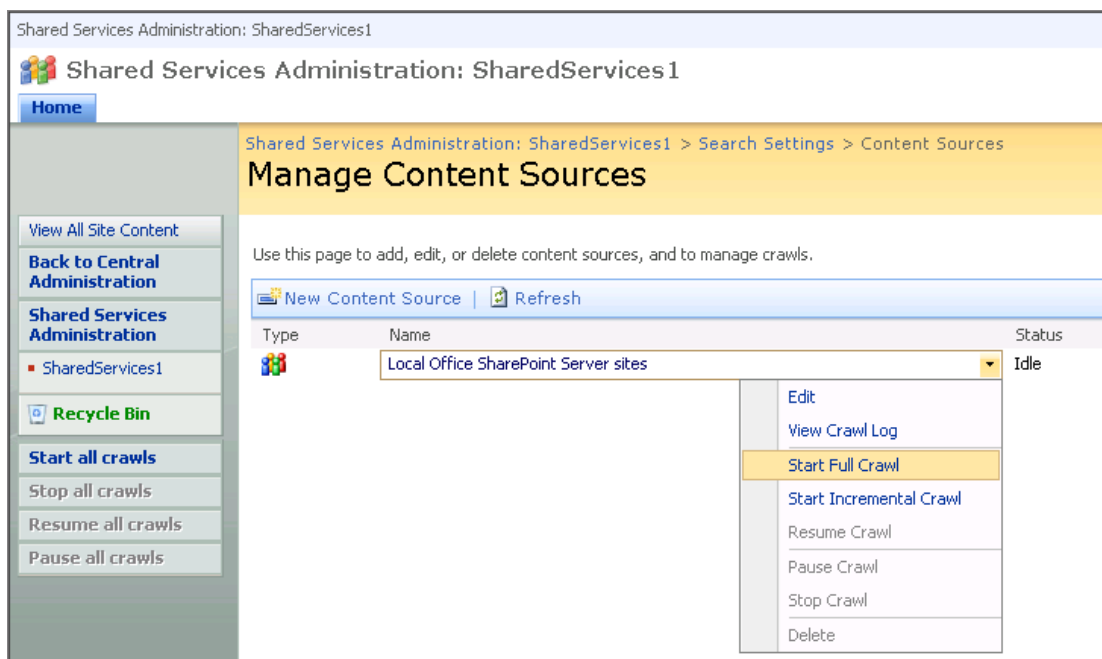


3. Click the **Browse** button to upload a PDF document and click **OK**. When it has uploaded, the **Documents Edit Item** page is displayed. You can give a title, if you like, to your uploaded PDF file. Click the **Check In** button.



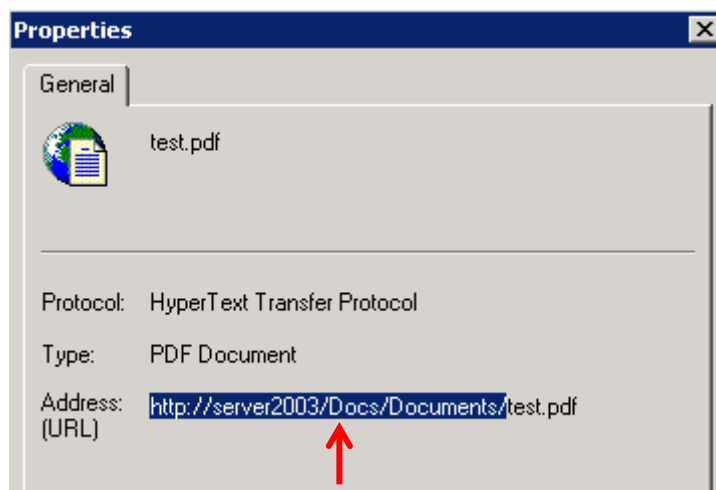
4. When you are done, you will be returned to the Documents page where you will see that your PDF file has been uploaded.
5. Return to the **Shared Service Administration: SharedServices 1** home page, click **Search Settings** link> under **Crawl Settings** click **Content sources and crawl schedules**.

6. On the **Manage Content Sources** page, open the content source's context menu by left-clicking the down arrow and select **Start Full Crawl** to index all files. This will start the indexing process.

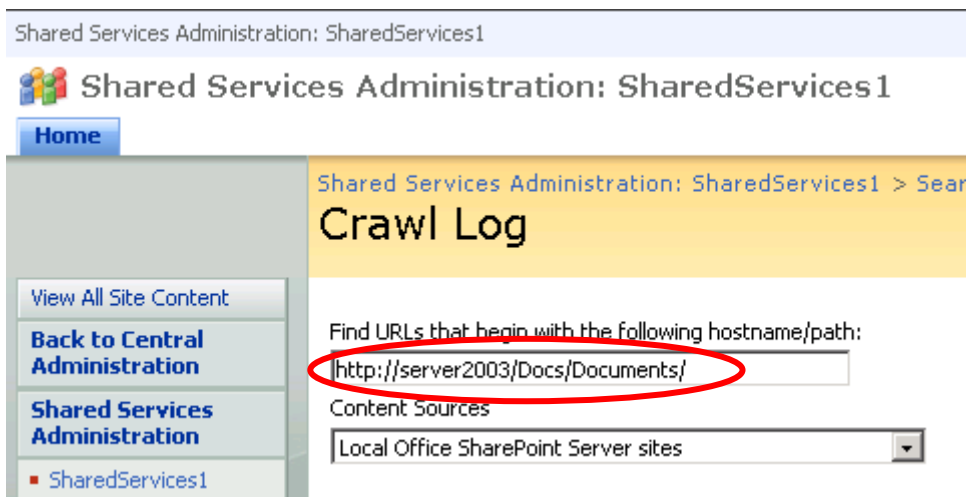


7. (Optional) To start an incremental crawl, please open the content source's context menu by left-clicking the down arrow and select **Start Incremental Crawl**.

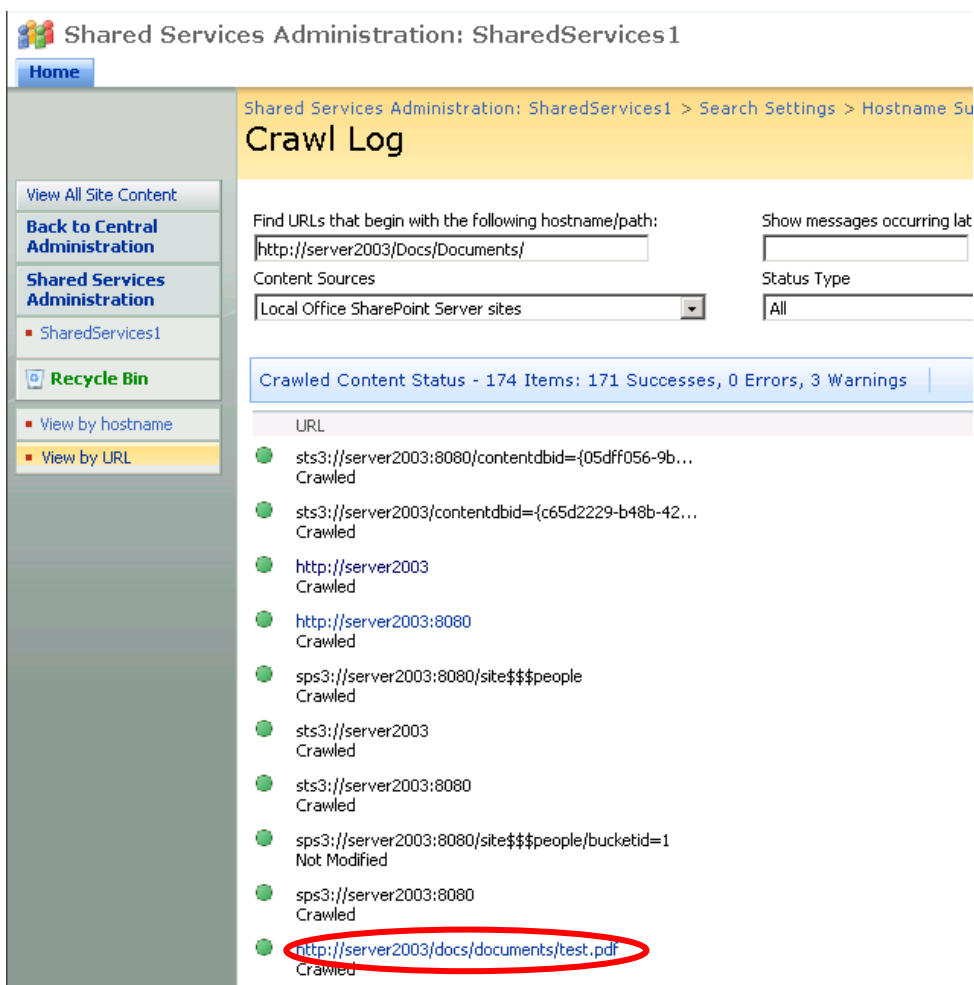
- View Crawl Logs.
 1. Return to the **Document Center**, and click **Documents**.
 2. Right-click the PDF document you have uploaded > choose **Properties**.
 3. Select the **address (URL)** of the PDF file except [filename.pdf] > right-click the selected area > choose **Copy**.



- Go to the **Manage Content Sources** page, open the content source's context menu by left-clicking the down arrow and select **View Crawl Log** > paste the URL to **Find URLs that begin with the following hostname/path** field > click the **Filter** button.



- When you're done, you can see the whole URL of the uploaded PDF file appear in the **Crawl Log** list.

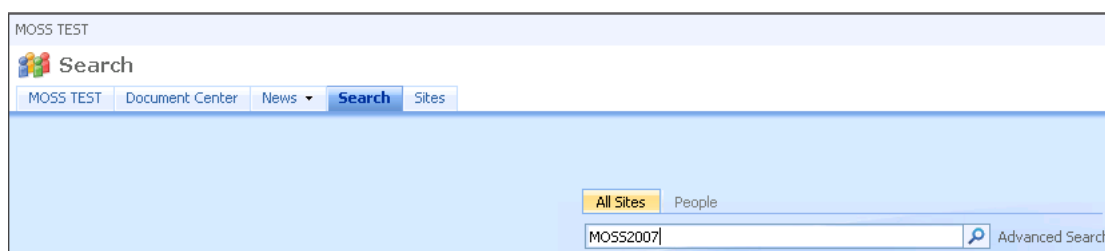


Searching PDFs

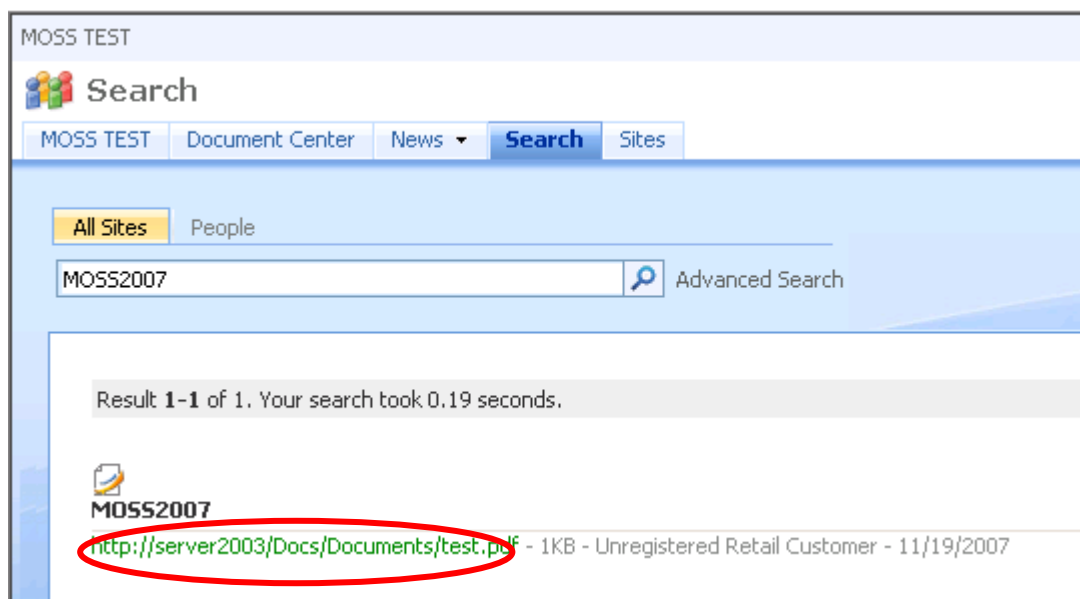
After you finish indexing PDF files, you can start a search to find text within PDF files. Also, you can search PDF document properties, such as author, title, etc.

To start a search, please perform the following:

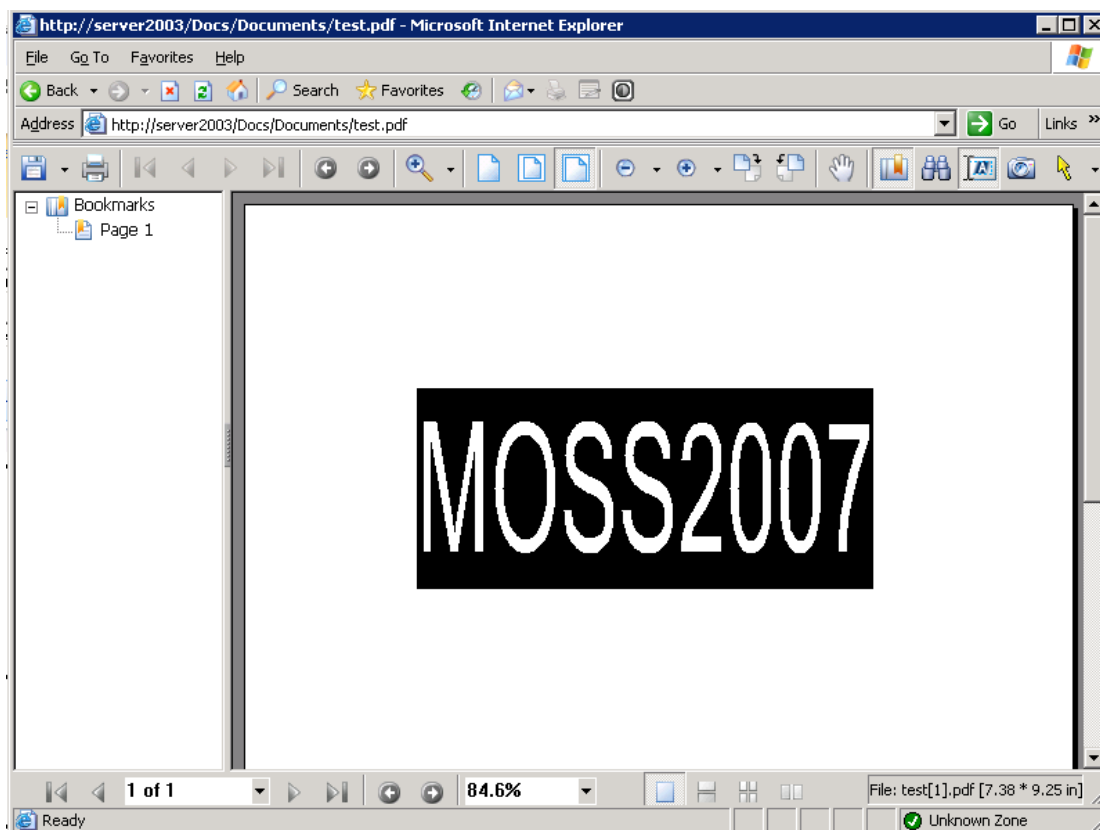
- Go to the **Document Center**, and click **Search** to visit the Search Center > type the text you want to search > hit Enter.



- You will see your search results. Click to open one of the PDF files.



Search results



Open one of the search results

More Help

For more information on how to work with Microsoft Office SharePoint Server (MOSS), please refer to Microsoft Office SharePoint Server (MOSS) Help included in the product.

Additional Information

Microsoft Developer Support

Foxit PDF IFilter is a Dynamic Link Library (DLL) file that provides a bridge between a Microsoft indexing client and a library of PDF files. It consists of code that understands the PDF Format as well as code that can interface with the indexing client. When an indexing client needs to index content from PDF files, it will look in its registry for an appropriate DLL and will find the Foxit PDF IFilter. Foxit PDF IFilter will return text to the indexing client. The indexing client will then index the results and return the appropriate results to the user.

Language Support

This section lists languages supported by Foxit PDF IFilter 1.0.

- German
- French
- Italian
- English
- Spanish
- Portuguese
- Russian
- Chinese (traditional and simplified)
- Korean
- Japanese
- Arab
- Hebrew

FAQ

1. What is Foxit PDF IFilter?

Foxit PDF IFilter is an application designed to help users index a large amount of PDF documents and then quickly find text within these documents. It acts as a plug-in for full-text search engines that scans documents for text and properties (also called attributes), extracts text from documents, filters out formatting and retaining information about text position, and extracts properties of an entire document.

2. Is Foxit PDF IFilter free?

Foxit PDF IFilter can be used for free on desktop environment. However, users need to purchase licenses when they use Foxit PDF IFilter on server environments.

3. I have a computer with one hyper-threading processor inside. Do I need to buy a license for 2 CPUs for the machine?

We will only count the “real” numbers of processors reported by your system. Please go to Start > Programs > Accessories > System Tools > System Information to check your CPU numbers.

4. What is the Foxit PDF IFilter's role in the PDF indexing process and PDF text search?

Foxit PDF IFilter acts as a plug-in for full-text search engines. A search engine usually works in two steps:

Step 1, the search engine goes through a designated place, e.g. a file folder or a database, and indexes all documents or newly modified documents, including PDF documents, in the background and creates internal data to store the indexing result.

Step 2, a user specifies some keywords he would like to search and the search engine answers the query immediately by looking up the indexing result and responds to the user with all the documents that contain the keywords.

During Step 1, the search engine itself doesn't understand the format of a PDF document. Therefore, it looks in the windows registry for an appropriate PDF IFilter and finds the Foxit PDF IFilter. Foxit PDF IFilter understands the PDF format. It filters out embedded formatting and extracts text from the document and returns text back to the search engine. Further information about IFilter specification can be found on the Microsoft Web site.

5. Do I have to remove Adobe PDF IFilter to install Foxit PDF IFilter?

Yes.

6. Where do I need to install the Foxit PDF IFilter in MOSS 2007?

In MOSS 2007, the Foxit PDF IFilter only needs to be installed on the index server.

7. How to verify the correct installation of Foxit PDF IFilter?

Firstly, please check if Foxit PDF IFilter appears in the list of programs on the Program Start Menu after installation.

Secondly, you can download IFilter detection tools to inspect IFilter installation information in the registry, e.g. the filtreg.exe program.

Thirdly, install Windows Desktop Search to verify if the text in PDF files can be found.

8. How to upgrade Foxit PDF IFilter? Do I need to reboot the machine?

You need to uninstall the old version of Foxit PDF IFilter to install the new one. You do not have to reboot your machine after update.

9. Do I have to rebuild index after updating?

That depends. Foxit PDF IFilter consists of a core layer and an installation layer. When you upgrade the installation layer, you do not need to re-do a full index. But when you upgrade the core layer, you may need to rebuild index on your computer.

10. Why can't I find some PDF files?

There might be three reasons:

First, Foxit PDF IFilter cannot index password-protected PDF files.

Second, the PDF file doesn't contain searchable text. You can check if there is searchable text in the PDF file by opening it in Foxit Reader and selecting the text with the Select Text tool. Or you can enter into the Text Viewer mode to select text. If neither method highlights any text, it is likely that the PDF contains a source image of text which Foxit PDF IFilter can not index.

Third, the PDF file doesn't contain text that is properly encoded. If the file contains searchable text, yet you cannot find the text with Find Tool in Foxit Reader, then the text may not be properly encoded.

11. Why can't Foxit PDF IFilter work after Microsoft updates? Do I have to re-install it?

If you installed the Foxit PDF IFilter on your search server, the IFilter does not work after you install a Microsoft update. It is because that the upgrade program automatically overwrites the registry value of Foxit PDF IFilter with that of Adobe PDF IFilter.

To make it work again, you must reinstall Foxit PDF IFilter. In this case, you do not have to uninstall IFilter. You can just run the Foxit IFilter MSI program again and

select the repair function. In addition, we will contact Microsoft to address this issue shortly.

Contact Us

Feel free to contact us should you need any information or have any problems with our products. We are always here, ready to serve you better.

- **Office Address:**

Foxit Corporation
39819 Paseo Padre Parkway
Fremont CA 94538
USA

- **Mailing Address:**

Foxit Corporation
39819 Paseo Padre Parkway
Fremont CA 94538
USA

- **Sales:**

1-866-MYFOXIT or 1-866-693-6948 (8AM-5PM PST Monday - Friday)
510-438-9090 (8AM-5PM PST Monday - Friday)
408-307-9358 (8AM-5PM PST Monday - Friday)

- **Support:**

1-866-MYFOXIT or 1-866-693-6948 (8AM-5PM PST Monday - Friday)
979-446-0280 (6AM-5PM PST Monday - Friday)

- **Fax:**

510-405-9288

- **Website:**

www.foxitsoftware.com

- **E-mail:**

Sales and Information - sales@foxitsoftware.com

Technical Support - support@foxitsoftware.com

Marketing Service - marketing@foxitsoftware.com