

dtSearch PDF Search Highlighter

dtSearch Corp.

<https://www.dtsearch.com/pdfhl>

(301) 263-0731

info@dtsearch.com

Copyright © 2011-2022 dtSearch Corp. All rights reserved.

Contents

- Overview..... 4
- Options 7
- Technical Information..... 9
 - How the plug-in works 9
 - Security Options 12
 - Information for Developers..... 14
 - Troubleshooting steps..... 17
 - Troubleshooting steps - Servers 19

Overview

The dtSearch PDF Search Highlighter is a free plug-in for Adobe Reader that makes it possible for dtSearch and other search programs to highlight hits in retrieved documents inside Adobe Reader DC, XI, and X. Earlier versions of Adobe Reader had this feature built in, so there was no need for a plug-in.

Subject to the included license agreement, you can download the plug-in at no charge from the [dtSearch web site](#) and install the plug-in on any number of computers within your organization. The plug-in may be used with any compatible product or web site -- not just dtSearch products or products that use the dtSearch Engine.

For additional information, please see the [dtSearch support page](#).

Compatibility and usage

dtSearch Desktop/Network Users

To use the plug-in with dtSearch, you will need dtSearch version 7.67 or later. dtSearch will automatically detect the plug-in if it is installed.

Web sites

The plug-in works with web sites to highlight hits in PDF files after a search. If a web site that you use to search PDF files has been updated to work with this plug-in, then you can just install the plug-in and leave the default settings in place. Web sites that use dtSearch Web 7.67 or later will automatically work with the plug-in.

If a web site has not been updated to work with this plug-in, then you will need to uncheck the box in [Options](#) to "Verify that highlighting requests are only sent to compatible servers."

Keyboard shortcuts

When a PDF file opens after a search, press Ctrl+Shift+Spacebar to go to the next hit and Ctrl+Shift+Backspace to go to the previous hit.

Toolbar buttons - dtSearch Desktop

In dtSearch Desktop, check the box in Options > Preferences > PDF View Options to "Use dtSearch toolbar to navigate hits in PDF files". With this option enabled, the dtSearch Desktop "Next Hit" and "Previous Hit" buttons will locate the next and previous pages with hits in a PDF file.

Toolbar buttons - Adobe Reader DC

Unfortunately Adobe Reader DC does not provide a way to make the buttons visible at all.

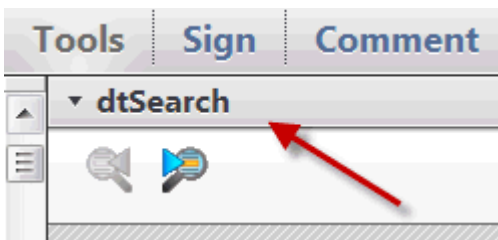
Toolbar buttons - Adobe Reader 11

To make the Next Hit and Previous Hit buttons visible in Adobe Reader XI:

1. If the Adobe Reader toolbar is not visible, press F8 to make it visible. (To make the toolbar visible by default, open Adobe Reader and click Edit > Preferences > Internet and un-check the option to "Display in Read Mode by default".)
2. Click the "Tools" button in the Adobe Reader toolbar to open the Tools panel



3. Click "dtSearch" to show the Next Hit and Previous Hit buttons.



Note: Adobe Reader does not allow plug-ins to add buttons to the main toolbar, so the plug-in can only add buttons to this Tools panel in Adobe Reader.

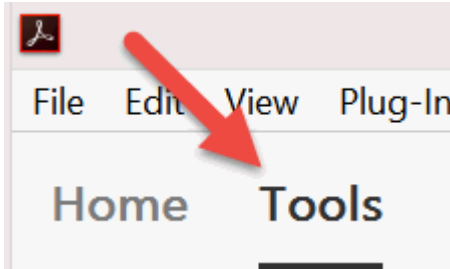
Toolbar buttons - Adobe Acrobat 11

1. In Adobe Acrobat, follow the steps above to make the toolbar buttons visible.

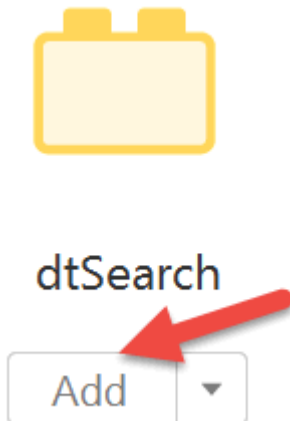
2. Right-click each button and select "Add to Quick Tools". This will make the buttons visible in Adobe Acrobat, but only when it is open outside a browser window.

Toolbar buttons - Adobe Acrobat DC

1. Start Adobe Acrobat and click Tools



2. Locate the dtSearch item in the list of tools and click "Add"



This will add a dtSearch button to the tools panel on the right. When viewing a file with highlighting, click this button and Next Hit and Previous Hit buttons will appear in a bar at the top of the document.

If the tools panel is hidden, look for a small triangle near the middle of the Adobe Acrobat toolbar and click the triangle to show the tools panel.

Options

After installing the plug-in, click Start > All Programs > dtSearch > dtSearch PDF Search Highlighter > dtSearch PDF Search Highlighter Options to access security and other options for the plug-in.

These options only affect highlighting of hits in Adobe Reader and will have no effect on internet search engines such as Google and Bing that do not do this.

Which web sites should be allowed to highlight hits on PDF files?

Only allow programs on my computer to highlight hits in PDF files.
Choose this option if you only use dtSearch Desktop/Network or other desktop search products and do not want to use web sites to highlight search hits PDF files.

Allow web sites to highlight hits, but only on PDF files from the same web site.
Choose this option if you use a web site to search PDF files and the PDF files are located on that same web site.

Allow web sites to highlight hits on PDF files from the same domain (recommended).
Choose this option if you use a web site to search PDF files and the PDF files are located on a different web site but within the same domain. For example, you might use a search engine at <https://search.example.com> to search PDF files at <https://www.example.com>.

Also allow these trusted web sites to highlight hits on PDF files on any web site.
This option lets you list web sites that you trust to provide highlight information on other sites' PDF files. For example, you might go to <https://search.example.com> to search legal opinions that are on unrelated web sites.

Verify that highlighting requests are only sent to compatible servers

Check this box to enable a security enhancement that verifies that highlight requests are only sent to servers that are compatible with this plug-in. For more information on the purpose of this, please see "[Security Options](#)."

You may need to un-check this box to enable hit highlighting to work with servers that use older versions of dtSearch Web (before version 7.67).

Diagnostics...

Click the Diagnostics... button to access options to temporarily disable the plug-in and to enable diagnostic logging for troubleshooting purposes.

Check for Updates...

Click Check for Updates to check for a newer version of the PDF Search Highlighter. You can also check the [dtSearch PDF Search Highlighter download page](#) for information about new versions.

Technical Information

Overview

The dtSearch PDF Search Highlighter is a plug-in for Adobe Reader that makes it possible for search programs to highlight hits in retrieved documents inside Adobe Reader X. Earlier versions of Adobe Reader had this feature built in, so there was no need for a plug-in.

When a user clicks on a link to a PDF file on a web page, the browser loads Adobe Reader as a plug-in and uses it to display the page.

Adobe Reader versions prior to Adobe Reader X provided a way for search programs to send hit highlighting information so search hits could be highlighted in Adobe Reader, and users could click "Next Hit" and "Previous Hit" buttons in Adobe Reader to find hits quickly.

Adobe Reader no longer supports this URL syntax. dtSearch Corp. has created the plug-in to restore this very useful Adobe Reader feature. Subject to the included license agreement, a user can download the plug-in at no charge from the [dtSearch web site](#) and install the plug-in on any number of computers within the user's organization. The plug-in may be used with any compatible product or web site -- not just dtSearch products or web sites that use the dtSearch Engine.

Installing the plug-in

To install the plug-in, download and run [dtSearchPdfHighlighter.exe](#). dtSearchPdfHighlighter.exe is a self-extracting ZIP archive containing an MSI installation package, dtSearchPdfHighlighter.msi.

The installer adds the following to the user's system:

- A folder in C:\Program Files named "dtSearch\dtSearch PDF Search Highlighter" with dtswebhits.api (the plug-in), dtspdfcfg.exe (a configuration program for the plug-in), and some HTML help files.
- A Start Menu shortcut for the configuration program.
- Copies of dtswebhits.api in the Adobe Reader and Adobe Acrobat plugins folders.

- A registry entry that adds "dtsPdfWh" to the Internet Explorer User Agent string. (For the reason this is necessary, please see "[Internet Explorer Compatibility](#).")
- A Browser Helper Object for Internet Explorer that enables hit highlighting to work in Internet Explorer. (In a "Custom" installation, this can be disabled.)
- A file named ProtectedModeWhitelistConfig.txt in the Adobe Reader program folder to allow dtswebhits.api to read highlighting data. (If this file already exists, the dtSearch-related entries will be added to it.)
- A registry value bUseWhitelistConfigFile under HKEY_LOCAL_MACHINE\SOFTWARE\Policies\Adobe\Acrobat Reader\<version>\FeatureLockDown set to 1 to enable use of the ProtectedModeWhitelistConfig.txt file.

Uninstalling the plug-in removes all of the above items from the user's system.

The plug-in uses settings stored under this registry key:

HKEY_CURRENT_USER\Software\dtSearch Corp.\dtSearch\PlugIns\dtswebhits.api

Each time a PDF file opens in Adobe Reader, the plug-in checks the URL used to open the file for hit highlighting data. If the URL includes a reference to highlighting data, the plug-in verifies that the reference to the highlighting data complies with the user's [security settings](#), downloads the highlighting data, and uses it to add hit highlighting to the PDF file in Adobe Reader.

How hit highlighting data is included in URLs

The format for the URLs that provide hit highlighting information is:

<https://www.example.com/sample.pdf#xml=http://www.example.com/hits.xml>

The #xml= portion of the link points to a URL that returns an XML stream describing the location of the hits in the PDF file. The format of the XML file is described Adobe Technical Note 5172 -- Highlight File Format (no public link available).

The #xml= portion of the URL can request the XML from a script or program, like this:

<https://www.example.com/sample.pdf#xml=http://www.example.com/getPdfHits.asp?idoc=5>

The dtSearch plug-in supports this URL syntax and also supports a variant, using ?xml= instead of #xml=. The variant is needed because the #xml= URLs are not visible to the plug-in when Adobe Reader is running inside a WebBrowser control.

Security Options

The plug-in has two categories of security settings: (1) settings that limit which servers can generate hit highlighting information based on the location of the PDF file, and (2) a verification protocol to ensure that the highlighting server is able to generate hit highlighting data.

These are intended to guard against use of the plug-in to exploit security vulnerabilities in web sites that do not sufficiently guard against cross-site scripting and cross-site request forgery attacks.

Servers that can generate hit highlighting information

The default settings only allow a site to hit-highlight its own PDF files. In cases where a site hit-highlights content located on other web sites, the Security Options dialog box allows the user to enter the web site as a "trusted" site that is allowed to do this.

Validation of highlighting script

To prevent use of the plug-in to send forged requests to web sites, the plug-in will send a standard validation request to make sure the target URL is really a PDF search highlighter. The validation request replaces the query in the original URL with "IsPdfHighlighter", and expects a response that contains "YesPdfHighlighter".

For example, suppose a user clicks this link:

```
https://www.example.com/harmless.pdf#xml=https://www.example.com/shopping.asp?buy=Product&ShipTo=Attacker
```

Adobe Reader will open the PDF file, and the dtSearch plug-in will see the #xml= in the URL. To verify the target URL, before requesting the highlighting data, the dtSearch plug-in will first send this request:

```
https://www.example.com/shopping.asp?IsPdfHighlighter
```

A shopping script that is not aware of the PDF search highlighting protocol will not respond with the specific "YesPdfHighlighter" code. If the server does not respond with this text, the dtSearch plug-in will decide the server does not support PDF highlighting and the request will not be sent.

To preserve backward compatibility with servers that have not been updated, users will have the option to disable this verification.

Developer Information

Internet Explorer Compatibility

When PDF files are opened inside an Internet Explorer window, the #xml= syntax is not visible to the plug-in, so it cannot access the hit highlighting information. To work around this, the plug-in installs a browser helper object (BHO) in Internet Explorer.

For situations where the browser helper object is not used, such as applications that embed the WebBrowser control, the plug-in supports an alternative syntax, ?xml= instead of #xml=, that works without the browser helper object.

dtSearch Web

If your web site uses dtSearch Web, upgrade to dtSearch Web 7.67 or later to make the site compatible with the plug-in.

Your users who have Adobe Reader DC will need to download and install the plug-in, using this download link:

<https://download.dtsearch.com/pdfhl/dtSearchPdfHighlighter.exe>

dtSearch Publish applications

(1) The lbview.exe viewer that comes with dtSearch 7.67 or later is needed to highlight hits in PDF files. If you have an older version, install dtSearch Publish, run the dtSearch CD Wizard and click "Change CD Settings", and check the box to "Update program files on the CD".

(2) When the CD starts, users who do not have the plug-in installed will see the page in /builtin/getplugin.html which has a prompt and link to download and install the plug-in.

Server-based applications

(1) Server-based applications should support the [IsPdfHighlighter verification protocol](#). To do this, your script that responds to PDF hit highlighting requests should respond to an IsPdfHighlighter request with a response that contains "YesPdfHighlighter". To make troubleshooting easier, you can embed this in a web page so you can see the response in a web browser. Example:

```
<html><head><title>Pdf Highlighter Confirmation</title></head><body><!--
YesPdfHighlighter --><P>This server supports PDF highlighting</P>
```

(2) Adobe Reader runs in a separate process from the web browser, so highlighting requests will not use the same web browser session that was used to search. As a result, session state and authentication done in the browser will not be available when the plug-in requests highlighting data from the server.

Desktop applications

Desktop applications can highlight hits in PDF files by embedding a WebBrowser control and navigating to a URL that provides the location of the PDF file and the highlighting data.

(1) The URL your application uses has to use ?xml= instead of #xml=. This is necessary because the #xml= syntax does not work in an embedded WebBrowser control.

(2) Your application should generate file-based URLs to send highlighting data to the plug-in:

- Call SearchResults.MakePdfWebHighlightFile as before to generate the highlight data.
- Write the highlight data to a temp file.
- Generate a URL with the location of the PDF file and the temp file with a unique name starting with "dtSearchPdfHits", like this:
c:\documents\example.pdf?xml=c:\temp\dtSearchPdfHits12345.xml
- Call the WebBrowser's Navigate method with the URL to open the PDF file.

The name of the XML file with the highlight data must be in the format dtSearchPdfHits*.xml, because the dtSearch plug-in adds a whitelist entry to an Adobe Reader configuration file enabling access to files with this name pattern. If the name of the highlight file is not whitelisted, the plug-in may receive an "Access Denied" error when it tries to read the highlight data.

Diagnostic tools

To help with troubleshooting, the dtspdfcfg.exe utility can enable a diagnostic logging mode in the plug-in. To run the dtspdfcfg.exe utility, click Start > All Programs > dtSearch PDF Search Highlighter > dtSearch PDF Search Highlighter Options.

Click the Diagnostics... button to access diagnostic options.

To create a diagnostic log, check the "Enable diagnostic logging" box. The plug-in will create a log file in the Windows TEMP folder, in a TEMP folder under the AppData\LocalLow folder, or in a "Low" folder under the Windows TEMP folder. Click "Zip logs for email" to collect all of the log files into a ZIP file.

You can also monitor the log in real time using the dbgview utility from the Microsoft [Sysinternals](#) web site. Because of Adobe Reader Protected Mode, you may need to run dbgview in low integrity mode to see the log. To do this, use the psexec utility (also available from the Microsoft [Sysinternals](#) web site) to launch dbgview.exe with the -l command-line switch, like this:

```
psexec -l dbgview.exe
```

For troubleshooting steps when highlighting does not work, please see:

Client-side troubleshooting: [Troubleshooting steps](#)

Server-side troubleshooting: [Troubleshooting steps - Servers](#)

Distribution of the plug-in

Subject to the included license agreement, a user can download the plug-in from the dtSearch web site and install the plug-in on any number of computers within the user's organization. The plug-in may be used with any compatible product or web site -- not just dtSearch products or web sites that use the dtSearch Engine.

To comply with Adobe licensing requirements, the plug-in must be deployed using the dtSearchPdfHighlighter.msi installer and must be distributed only from the dtSearch download URL (<https://download.dtsearch.com/pdfhl/dtSearchPdfHighlighter.exe>).

Adobe Reader plug-ins can only be created using a digital signature issued by Adobe, which dtSearch Corp. must renew each year with permission from Adobe. Each time the license is renewed, the plug-in must be updated with a new digital signature, and renewal is subject to Adobe's discretion.

Troubleshooting steps

Problem: PDF files open in web browser but without hit highlighting

(1) Check that PDF files are opening in Adobe Reader or Adobe Acrobat versions DC, XI, or X. The dtSearch PDF Search Highlighter plug-in is designed only to work with these versions of Adobe Reader and Adobe Acrobat. If you are using a non-Adobe viewer, or an older version of Adobe Reader or Adobe Acrobat, then please see this article:

PDF viewers that support highlighting hits
<https://support.dtsearch.com/faq/dts0229.htm>

(2) Click this link to see if hit highlighting works:

<https://support.dtsearch.com/pdftest/constitution.pdf#xml=https://support.dtsearch.com/pdftest/states.xml>

If hit highlighting does appear, then your computer is set up to highlight hits in PDF files, and so the highlighting problem must be with the web site. For troubleshooting steps, please see [Troubleshooting steps - Servers](#).

If hit highlighting does not appear, either the plug-in is not installed or it is disabled in your browser.

(3) Click Start > Programs > dtSearch PDF Search Highlighter > dtSearch PDF Search Highlighter Options

If the program does not start, or if the menu option does not exist, the plug-in is not installed.

(4) Check that the web browser is set up to use Adobe Reader for PDF files. Currently only the Internet Explorer browser supports display of PDF files in Adobe Reader. If you are using Internet Explorer, in Tools > Manage Add-ons, check that the "dtSearch PDF Search Highlighter BHO" is enabled.

Problem: PDF files do not open inside the web browser

If you click a PDF file and either you get a download prompt or the PDF file opens in Adobe Reader outside of the web browser, then Adobe Reader is not set up to run inside web browsers. To correct this, run Adobe Reader and click Edit > Preferences > Internet and check the box to "Display PDF in browser".

Problem: The Adobe Reader toolbar is missing

Adobe Reader X is set up to hide the toolbar by default. To change this, open Adobe Reader and click Edit > Preferences > Internet and un-check the option to "Display in Read Mode by default".

To make the toolbar visible in a browser window when it is hidden, press F8.

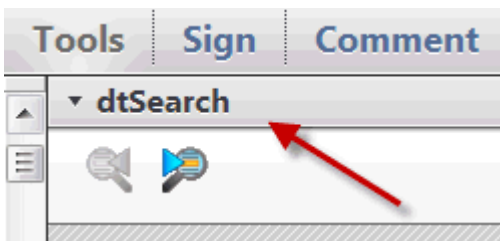
Problem: The toolbar is visible but "Next hit" and "Previous hit" buttons are missing

To make the Next Hit and Previous Hit buttons visible in Adobe Reader:

1. Click the "Tools" button in the Adobe Reader toolbar to open the Tools panel



2. Click "dtSearch" to show the Next Hit and Previous Hit buttons.



Note: Adobe Reader no longer allows plug-ins to add buttons to the main toolbar, so the plug-in can only add buttons to this Tools panel in Adobe Reader.

Troubleshooting steps - Servers

First, check that the client machine is set up correctly to support PDF highlighting. To check, click this link both in Internet Explorer and in Firefox:

<https://support.dtsearch.com/pdf/test/constitution.pdf#xml=https://support.dtsearch.com/pdf/test/states.xml>

If hit highlighting does not work, please see [Troubleshooting steps](#).

If PDF highlighting works when you click the link above but not in your application,

(A) Check the URLs that your application generates

Check that the URLs your application generates have the right format. The format for the URLs that provide hit highlighting information is:

`https://www.example.com/sample.pdf#xml=https://www.example.com/hits.xml`

The `#xml=` portion of the link points to a URL that returns an XML stream describing the location of the hits in the PDF file.

(B) Check the `IsPdfHighlighter` protocol

Check that the script that handles highlighting requests handles the `IsPdfHighlighter` protocol. To prevent use of the plug-in to send forged requests to web sites, the plug-in will send a standard validation request to make sure the target URL is really a PDF search highlighter. The validation request replaces the query in the original URL with `"IsPdfHighlighter"`, and expects a response that contains `"YesPdfHighlighter"`.

For example, suppose a user clicks this link:

`https://www.example.com/harmless.pdf#xml=https://www.example.com/gethits.aspx?search=123456`

Adobe Reader will open the PDF file, and the `dtSearch` plug-in will see the `#xml=` in the URL. To verify the target URL, before requesting the highlighting data, the `dtSearch` plug-in will first send this request:

`https://www.example.com/gethits.aspx?IsPdfHighlighter`

To test your script, you can enter the URL for your highlighting script in a browser window with the `IsPdfHighlighter` request added and check that the response includes `YesPdfHighlighter`.

(C) Generate a diagnostic log on a client machine where highlighting does not work

To generate a diagnostic log:

(1) Click Start > Programs > dtSearch Pdf Search Highlighter > dtSearch PDF Search Highlighter Options to open the configuration program for the highlighter.

(2) Click Diagnostics... and check the box to Enable diagnostic logging to create a text log file of the highlighting process.

(3) Try to open a PDF file that should have highlighting.

(4) Close all browser windows and all Adobe Reader windows

(5) In the dtSearch PDF Search Highlighter Options program, click Zip logs for email to find the diagnostic logs.

The plug-in creates a log named `dtSearch_WebHitsPlugin*.log`. If the log is not present in the ZIP file, make sure you closed all Adobe Reader and browser windows before clicking the Zip logs for email button (an open Adobe Reader instance will keep the log file locked).

In the diagnostic log:

- Check that the URL detected in the log includes the `#xml=` syntax. If you do not see the `#xml=` syntax in the log, either your application may not be generating the `#xml=` links, or, if you are using Internet Explorer, then the dtSearch PDF Search Highlighter BHO may be disabled in Internet Explorer (see [Troubleshooting steps](#)).

- Check for any error messages recorded in the log identify any [security issues](#) that prevented highlighting from working.

- Check that the `IsPdfHighlighter` query was processed correctly with a response that includes "YesPdfHighlighter"

- Check that the XML your application returns is correctly formatted and does not include any extra content such as HTML headers or messages.