Adobe™ Acrobat™ Search Online Guide

This online guide contains all the information you need to use the Search feature with Acrobat Reader. You can use the Search feature to quickly locate information in collections of PDF documents that have been indexed with the Acrobat Catalog™ program. Click one of the following topics to go to an explanation of that topic:

How to use this online guide

Use these procedures to navigate through this guide:

Topic Click underlined text to go to the topic indicated. Underlined text indicates text that is linked to another part of this guide.
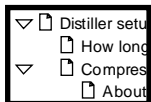
Click the Go Back button in the toolbar to return to your previous location.

Click the Next Page button in the toolbar to go to the next page of the guide.

Click the First Page button in the toolbar to return to the opening screen of this guide.

Click the bookmark name to go to the topic marked by that bookmark. Click the triangle to the left of a bookmark to show and hide subordinate bookmarks. The bookmarks for this guide provide a complete list of topics.

Click the arrow button to go to the next screen of any continued topic.

About the Search feature

With the Acrobat Reader Search command, you can perform <u>full-text</u> searches of collections of PDF documents that have been indexed with the Adobe Acrobat Catalog program.

Unlike the Find command of a word-processing program, which searches for words by reading every word on every page, the Search command searches full-text indexes created by Acrobat Catalog. The full-text indexes created by Acrobat Catalog can represent hundreds or thousands of documents, and—compared with searching a document word by word one page at time—searching a full-text index is very fast.

A full-text index is an alphabetized list of all the words and terms used in a collection of documents. The Search command uses full-text indexes to find words and terms quickly in the documents without having to open the documents.

→

Indexed document collections can be extremely large. Electronic document publishers can, for example, build an index for all the PDF documents that fit on a CD-ROM. Corporate index publishers can build an online index for all the PDF documents that fit on a network file server. The Search command lets you search thousands of documents in seconds.

Using the Search command, you can quickly search indexed PDF documents for single words or terms, phrases, or arbitrary character patterns specified with wild-card characters. You can also search for documents that contain combinations of words and phrases. You can, for example, search for documents that contain the phrase *status report* and that contain some form of the word *succeed*, such as *succeed*, *success*, or *successful*.

Unlike most other full-text search programs, the Acrobat Search command can find words and terms in illustrations, graphs, and formatted tables. Because it searches PDF documents and because PDF is based on the PostScript page description language, which represents illustrations and artwork as well as text, the Search command can find all the words on a page, no matter where or how they are used. The Search command can even find words that are rotated or attached to a curved line in an illustration.

In addition to finding indexed words and terms, the Search command can find documents that contain specific field values. The Search command can find, for example, all the documents that have Author fields that contain the name *von Neumann*, and that have Keyword fields that contain *binary* or *program*.

The Search command can also find documents that were created or modified during a specific period. You can limit your searches, for example, to documents that were created before or after a particular date.

Search buttons

The Acrobat Reader toolbar contains tools for selecting, viewing, annotating, and linking documents. The Acrobat Search feature adds the following buttons to the toolbar:

The **Search** button searches indexed PDF documents for single words, phrases, or combinations of words.

The **Search Results** button displays the results of your cross-document search.

The **Search Next** button displays the next occurrence of a search term found by the Search command.

The **Search Previous** button displays the previous occurrence of a search term found by the Search command.

Starting a search

1 Choose Search > Query from the Tools menu, or click the Search button 🔍 on the toolbar.

2 Select one or more indexes. The indexes represent the document collections you want to search.

3 Choose one or more of the following options:

▪ Enter a search term, phrase, or Boolean expression. You don't have to enter search text; you can perform a search with just Document Info field values or a document creation or modification date range.

▪ Enter one or more Document Info field values, either to limit your search or to find documents with field values you specify.

▪ Enter a creation or modification date range, either to limit your search or to find documents created or modified in the periods you specify.

▪ Choose one or more search options.

4 Click Search to perform the search. Acrobat Reader searches the index or indexes and displays a list of all the found documents in the Search Results window.

5 View the search results. By default, the documents returned from a search are listed according to a score (or relevance ranking), which places the documents most likely to contain relevant information at the top of the list. Double-click a document title to display the first page that contains a search term in that document. With one of the documents returned from the search displayed, you can use the Search Next and Search Previous buttons to see all the pages containing search terms in all the documents returned from the search.

See Viewing search results for more information.

After performing a search, you can use the Refine feature to perform another search that is limited to just those documents returned by the previous search.

Selecting indexes to search

The Search command searches for words and character sequences in documents that have been indexed by the Acrobat Catalog program. By selecting a single index, you can find words and terms in any document in the collection of indexed documents. By selecting two or more indexes, you can search for words in two or more collections at the same time.
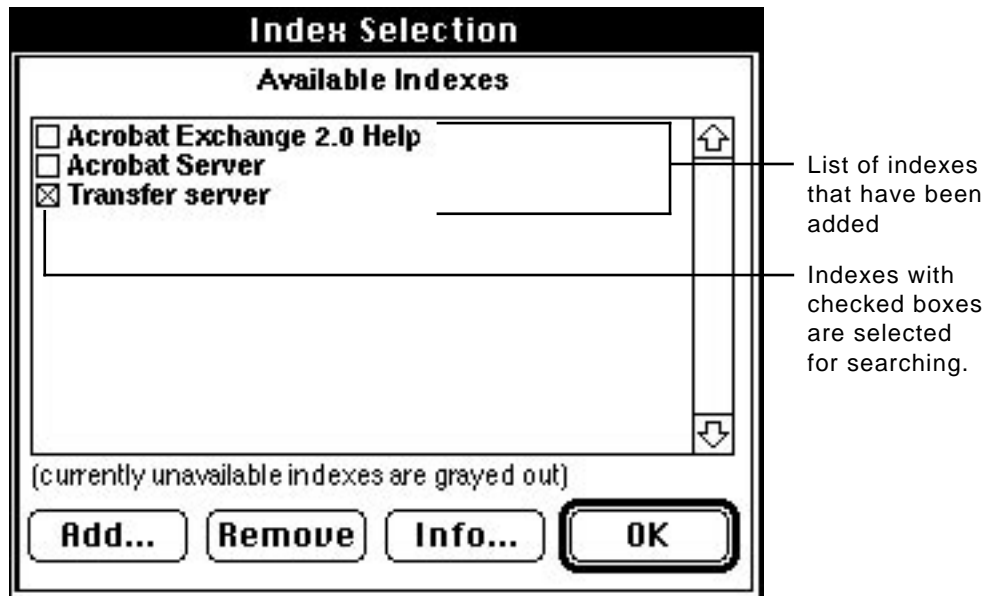
The publisher of a document collection can automatically add an index to the list of available indexes. To search additional indexes, you must add them to the list of available indexes. As you search for information, you choose which indexes to search by selecting and deselecting indexes in the list of available indexes.

To help decide which index to use for your search, you might want to view a description of an index. See <u>Viewing information about indexes</u> for instructions.

To select one or more document indexes to search:

1 Click the Indexes button in the Acrobat Search window. Alternatively, choose Search > Indexes from the Tools menu. The Index Selection dialog box appears.

Index Selection

Available Indexes

☐ **Acrobat Exchange 2.0 Help**
☐ **Acrobat Server**
☒ **Transfer server**

(currently unavailable indexes are grayed out)

[Add...] [Remove] [Info...]   **OK**

List of indexes that have been added

Indexes with checked boxes are selected for searching.

→

2 Click the check box to the left of the index you want to use.

3 Click OK to close the Index Selection dialog box and return to the Search window.

Note: *Dimmed indexes are currently unavailable for searching. Usually, an index is unavailable because the network connection between your computer and the server where the index resides has been lost. See* The index is unavailable *for more information.*

To add an index to the search list:

1 Select Search > Indexes from the Tools menu, or if the Acrobat Search window is already open, click the Indexes button. The Index Selection dialog box appears.

2 Click Add. The Add Index dialog box appears.

3 Locate and select the index you want to use. Acrobat index definition filenames end with .pdx.

4 Click OK (with Windows) or Open (on the Macintosh) to select the index and close the Add Index dialog box. Alternatively, double-click the name of the index you want to use. The new index is added to the search list.

→

To remove an index from the search list:

1 Click the Indexes button in the Acrobat Search window. Alternatively, Choose Search > Indexes from the Tools menu. The Index Selection dialog box appears.

2 Click to highlight the name of the index you want to remove.

3 Click Remove. The index is removed from the list.

4 Click OK to return to the Acrobat Search window.

Viewing information about indexes

After you have <u>added an index to the search list</u>, you can display the following information about the index:

- **Title** is the index name that appears in the Index Selection dialog box.

- **Description** is the index description written by the publisher who created it.

- **Path** is the location and name of the PDX index file as seen from your computer.

- **Last Built** is the date the index was last updated to include new and changed documents.

- **Created** is the date the index was created.

- **Documents** is the number of PDF documents in the collection of indexed documents.

→

▪ **Status** indicates the index's current status. *Available* means that you can select the index for searching. *Unavailable* means that the Search command cannot open the index files, usually because the network connection to the server containing the index files has been lost. You cannot search an unavailable index. *Not Loaded* means the index has not been selected; it may or may not be available for searching.

To display information about an index:

1 Click the Indexes button in the Acrobat Search window. Alternatively, choose Search > Indexes from the Tools menu. The Index Selection dialog box appears.

2 Click to highlight the name of the index in which you are interested.

3 Click Info. The Index Information dialog box appears.

Searching for a single term

To find occurrences of a single term, enter the term in the Find text box and click Search. A search term can be a word, a word with wild-card characters, or any combination of letters, numbers, and symbols.

You cannot find stopwords, which are words excluded from an index, and you might not be able to find numbers. Contact the person who created the index to learn which words, if any, are excluded from the index and whether numbers are included in the index.

You can use the Match Case, Word Stemming, Sounds Like, and Thesaurus options when searching for a single term. The Match Case, Word Stemming, and Sounds Like options might not be available with all indexes, because index publishers can choose not to add support for these features.

You can also use Document Info and Document creation and modification date field values to limit the number of documents returned by a search for a single term.

→

See [About Acrobat indexes](#) for a description of how the Acrobat Catalog program recognizes words and other terms in PDF documents and how index publishers can prevent stopwords and numbers from being included in an index.

Searching for a phrase

To search for a phrase, type the phrase in the Find text box and click Search. For example,

    status report

finds all occurrences of *status report*, *Status report*, and *Status Report*, as well as all other combinations of uppercase and lowercase letters for the phrase *status report*.

You can use a phrase as part of a [Boolean expression](#). For example,

    status report AND anaconda project

finds all documents that contain both the phrase *status report* and the phrase *anaconda project*.

If the phrase you want to find contains a [Boolean operator](#), you must enclose the phrase in quotation marks. For example,

    "cats and dogs" OR "rain or shine"

finds all documents that contain either the phrase *cats and dogs* or the phrase *rain or shine.*

Use the Match Case option to find phrases that are capitalized exactly like the phrase you enter in the Find box.

The Thesaurus, Sounds Like, and Word Stemming options have no effect on phrase searches.

You can use Document Info and Document creation and modification date field values to limit the number of documents returned by a search for a phrase.

Stopwords and phrase searches

Phrases that include stopwords cannot be found as phrases. For example, if Lincoln's Gettysburg address is indexed with the typical stopword list—which includes *of*, *the*, *by*, and *for*—users cannot find the address by searching for its most famous phrase,

Of the people, by the people, for the people

or with

people people people

which is how—with the typical stopword list—the phrase is indexed. In this case, the only way to find the words in Lincoln's famous phrase is to search for *people*. The limitation that stopwords place on phrase searches is a good reason not to exclude stopwords from an index. However, stopwords reduce the size of an index, which is why they are sometimes excluded.

Punctuation and other special characters in phrase searches

The Search command ignores any punctuation characters you enter as part of a phrase.

For example, the search phrase *Go home!* finds *Go home!* but it also finds strings such as *I want to go home* and *Go home? You must be joking*.

Similarly, the Search command ignores other special characters such as the at sign [@] and the asterisk [*]. See How terms are indexed for details.

Searching with wild-card characters

You can use wild-card characters to find all the words that contain a word fragment or all the words and terms that match an arbitrary character pattern. The wild-card characters are the following:

*   the asterisk, which matches zero, one, or more characters

?   the question mark, which matches any one character

Wild-card examples

geo*   matches words such as *geode*, *geodesic*, *Geoffrey*, *geography*, *geometry*, *George*, and *geothermal*

*nym   matches words such as *antonym*, *homonym*, and *synonym*

?ight   matches words such as *fight*, *light*, *might*, *right*, and *sight*

555-????   matches all seven-digit phone numbers with the 555 prefix

pr?m*   matches words such as *premature*, *premeditate*, *prim*, *primate*, *promise*, and *promontory*

You can use wild-card characters in a term that is part of a [Boolean expression](#). You can also use wild-cards to specify [Document Info](#) field values.

Note: *You cannot use wild-card characters to represent separator characters such as the hyphen [-] and the slash [/]. When the Catalog program builds an index, it uses separator characters to recognize where words begin and end and does not place them in the index. To find both* begin/end *and* begin-end*, for example, you cannot search for* begin?end*. Searching for* begin?end *finds neither* begin/end *nor* begin-end*. To find both, you must search for the phrase* begin end*, or for both* begin/end *and* begin-end*. See* [How terms are indexed](#) *for a description of separator characters.*

Search tips

When you begin using the Search command, you might encounter the following two problems:

- A search finds too many documents. A search for a single word can, for example, return hundreds of documents—far too many to open and view.

- A search finds too few, or no documents.

This topic gives advice for solving both of these problems.

Handling a search that finds too many documents

When a search returns too many documents, you can use the following techniques to restrict or refine the results of the search.

- Ignore documents with a medium or low relevance ranking. By default, found documents are listed in order of a relevance ranking. The information you seek might be in the documents with the highest relevance ranking.

- Use AND to limit searches to documents that contain two or more words or phrases. You can search, for example, for documents that contain *North* and *Nicaragua* and *senate hearings*. You can also use NOT to find documents that contain one, two, or more words but not another word. You can search, for example, for documents that contain *human* and *fossil* but not *piltdown*. See [Searching with AND, OR, and NOT](#) for more information about AND and NOT.

- When you use AND to search for documents containing two or more words, you can use the Proximity search option to change the way found documents are assigned a relevance ranking. With the Proximity option selected, the closer together the words are in a document, the higher the relevance ranking. The Proximity option also requires that terms in an AND search be within three or four pages of one another to be found.

- Use the [Refine Search feature](#) to limit a search to the documents returned by a previous search. You can, for example, first find all the documents that contain the phrase *Anaconda project*. Then, by using the Refine option, you can enter any combination of search terms and options, and the search will be limited to only those documents that also contain *Anaconda project*.

- Use [field values](#) to restrict searches to only those documents with field values you specify. When searching an index of English literature, for example, you can limit a search to just those documents that have an Author field that contains *Shakespeare* and a Title field that contains *king*.

- Use [date ranges](#) to restrict searches to only those documents that were, for example, created before a specific date or modified after a specific date.

- Use the [Match Case](#) search option to limit searches to documents that contain the words or phrases that are capitalized exactly like the search word or phrase you enter.

Locating hard-to-find information

If a search returns too few documents, first make sure that the search is set up correctly. Make sure that you are using the index or indexes built for the documents that contain the information you seek. Also, make sure that you don't have search options selected or field values entered from a previous search.

If a search is set up correctly and too few documents are returned from the search, use the following techniques to increase the number of found documents:

- Use the <u>Word Stemming</u> search option to find words that share a stem with a search word. Searching for *programming*, for example, also finds *program* and *programs,* but not *programmer*. (See <u>Word Stemming search option</u> for details.)

- Use OR to search for alternative spellings of the same word or to search for two or more words that are used to describe the information for which you are searching. When looking for information about political revolutions, for example, you might search for *revolution* or *coup* or *uprising*. See <u>Searching with AND, OR, and NOT (Boolean expressions)</u> for more information about OR searches.

- Use the <u>Thesaurus</u> search option to find words that have the same meaning as the search word. Searching for crypt, for example, also finds *grave*, *mausoleum*, *sepulcher*, and *tomb*.

→

- Use the [Sounds Like](#) search option to find different spellings of last names. Searching for *Smith*, for example, also finds *Smythe*. (Searching for *Smith*, however, also finds *send* and *somewhat*. See [Using the Word Assistant](#) for a way to use the Sounds Like feature without finding irrelevant terms.)

Limiting search results

One way to limit the number of documents returned by a search is to refine your search query. Another way to limit search results is to limit the number of documents that are searched. For example, you can limit a search to documents that were written after a certain date.

See the following topics for more information:

Using Document Info fields

Using document creation and modification dates

Using Document Info fields

You can use Document Info field values to limit the results of your searches to just those documents that contain specific field values. To add the Document Info fields to the Search window, choose Preferences > Search from the Edit menu and select Show Fields in the Preferences dialog box.

```
┌ With Document Info ─────────────────────────────────┐
:        Title │                                      │ :
:      Subject │                                      │ :
:       Author │                                      │ :
:     Keywords │                                      │ :
└─────────────────────────────────────────────────────┘
```

When you enter a Document Info field value or expression, only documents with fields that contain the value and that meet the search criteria are returned.

→

Document Info fields may be empty, because PDF document creators are not required to enter field values. You may also see additional fields not described in this guide, because PDF document creators can add custom fields.

When you enter one or more Document Info field values or expressions in the Find text box, you do not have to enter a search term. You can use fields to find all the documents that contain a specified field value.

Note: *The Title field contains a document title that has been entered by the author. For Macintosh documents, it does not contain the Macintosh filename unless the author entered the filename as a title.*

Using two or more fields

Entering values for two or more fields specifies an AND condition. Documents must contain all the field values you specify. For example, if you enter *Samuelson* in the Author field, *cost analysis* in the Subject field, and *spreadsheet* in the Keywords field, only documents that contain these values in these fields are returned from subsequent searches.

Wild-card characters

You can use wild-cards in Document Information text boxes. For example, if you enter

Sm?th*

in the Author text box, documents with Author field values that include *Smith*, *Smyth*, and *Smythe* can be returned with subsequent searches. See Searching with wild-card characters for instructions on how to use wild-card characters.

→

Boolean expressions

You can use Boolean expressions in the Document Info text boxes. For example, if you enter

    Shakespeare or Milton

in the Author text box, only documents that include either *Shakespeare* or *Milton* in the Author field are found. See [Searching with AND, OR, and NOT (Boolean expressions)](#) for instructions on how to build Boolean expressions.

Boolean expressions that use more than one field

By combining the field expression with the search expression in the Find text box, you can build a Boolean expression that uses more than one field. For example, if you enter

("Sixteen to one project" OR "16 to 1 project") AND
(Author ~ Raskin) AND (Keywords ~ "slide show" OR
keywords ~ presentation OR keywords ~ spreadsheet)

in the Find text box, the search returns only documents that

- contain either the phrase *Sixteen to one project* or
16 to 1 project and

- Have an Author field that contains *Raskin*, and

- Have a Keywords field that contains either the phrase *slide show*
or the word *presentation* or *spreadsheet.*

The tilde [~] operator is described in the following section.

→

Operators that can be used only with Document Info fields

Certain operators can be used only in Document Info field values.

Operator	Meaning
=	matches exactly (for text, numeric, and date values)
~	contains (for text values)
!=	does not contain (for text, numeric, and date values)
<	is less than (for date or numeric values)
<=	is less than or equal to (for date or numeric values)
>	is greater than (for date or numeric values)
>=	is greater than or equal to (for date and numeric values)

The comparison operators (<, <=, >, and >=) can be used only with values of the same type. For example, a date value can be compared only with another date value, and a text value can match exactly only another text value.

The standard Document Info fields—Title, Subject, Author, and Keywords—are all text values.

Using document creation and modification dates

You can use the Date Info fields to limit search results to just those documents that were created or modified in a specific time period. To add the Date Info fields to the Acrobat Search window, choose Preferences > Search from the Edit menu and select the Show Date option.

With Date Info	mo dy yr		mo dy yr
Creation after	/ /	before	/ /
Modification after	/ /	before	/ /

To specify a date:

1 Click the month field box for the date you want to specify, and enter the number of the month. The current day and year appear in the day and year fields.

2 To change the day, click the day field and either enter the day number or use the up and down arrows to select the day number.

3 To change the year, click the year field and either enter the year or use the up and down arrows to select the year.

→

To limit a search to documents created or modified after a specific date:

Specify the After date, and leave the Before date blank.

To limit a search to documents that were created or modified before a specific date:

Specify the Before date, and leave the After date blank.

To clear the date fields:

Click Clear.

To use both the creation and modification dates:

Entering time periods for both creation and modification dates specifies an AND condition. Only documents created in the specified period and modified in the specified period are returned from a search.

Using search options

The search options expand and restrict the results of searches with single terms, phrases, and Boolean expressions. The search options do not apply to Document Info field values. To add the search options to the Acrobat Search window, choose Preferences > Search from the Edit menu and select Show Options.

```
Options
☒ Word Stemming    ☐ Thesaurus         ☐ Match Case
☐ Sounds Like                          ☐ Proximity
```

- Word Stemming finds words that share a stem with the search term.

- Thesaurus finds words that have the same meaning as the search term.

- Sounds Like finds words that share a phonetic resemblance to a search term.

- Match Case finds only words that are capitalized exactly like the search term.

- Proximity changes AND searches so that words must be within three or four pages of one another.

→

See also [Using the Refine Search feature](#).

You can use the [Word Assistant](#) to see the effects of using the Word Stemming, Thesaurus, and Sounds Like search options before performing a search. Choose Search > Word Assist from the Tools menu to display the Word Assistant dialog box.

Word Stemming search option

Word stemming expands the search to find words that share word stems with the search terms you enter in the Find box. For example, with Word Stemming selected, searching for *building* finds *build* and *builds* as well as *building*.

Note: *The Word Stemming option does not find words that end with* er*. For example, searching for* build *does not find* builder*, searching for* bake *does not find* baker*, and searching for* publish *does not find* publisher*.*

Use the [Word Assistant dialog box](#) to see all the words in the current index or indexes that share a stem with a specified search term. Choose Search > Word Assist from the Tools menu to display the Word Assistant dialog box.

The Word Stemming feature does not apply to terms that contain [wild-card characters](#).

→

You cannot use the Word Stemming option with the Match Case option. When Word Stemming is selected, the Match Case option is ignored.

Not all indexes support the Word Stemming option; index publishers must build support for word stemming into indexes. When an index does not support word stemming, selecting the option has no effect.

Thesaurus search option

Thesaurus expands the search to find words that bear some semantic resemblance to the search terms you enter in the Find box. For example, with Thesaurus selected, searching for *begin* finds *start*, as well as *attack*, *produce*, and many other terms.

Sometimes the number of words that the Search command considers to have the same meaning as a given search term is surprising. Before searching with the Thesaurus option, it is usually a good idea to use the Word Assistant to see all the words that the Search command considers to have the same meaning as a specific search term. Choose Search > Word Assist from the Tools menu to display the Word Assistant dialog box.

The Thesaurus option does not apply to terms in a search phrase or to terms that contain wild-card characters.

You cannot use the Thesaurus option with the Match Case option. When the Thesaurus option is selected, the Match Case option is ignored.

Sounds Like search option

Sounds Like expands the search to find words that begin with the same letter as a search term and that share some phonetic relation to a search term. For example, with Sounds Like selected, searching for *Smith* also finds *Smyth* and *Smythe*. Searching for *Smith*, however, also finds *send* and *somewhat*. Before searching with the Sounds Like option, it is usually a good idea to use the [Word Assistant](#) to see all the words that the Sounds Like option will find. To display the Word Assistant dialog box, choose Search > Word Assist from the Tools menu.

The Sounds Like feature does not apply to terms in a search phrase or terms that contain wild-card characters.

You cannot use the Sounds Like option with the Match Case option. When Sounds Like is selected, the Match Case option is ignored.

Not all indexes support the Sounds Like option; index publishers must build support for the Sounds Like option into indexes. When an index does not support Sounds Like, selecting the option has no effect.

Match Case search option

Match Case limits the results of the search to find only those documents that contain words with the same capitalization as the search term or terms you enter in the Find box. With Match Case selected, for example, searching for *He* finds all occurrences of *He*, the symbol for Helium, but not all occurrences of *he*, the common pronoun.

You can use the Match Case option with phrase searches, and the Match Case option applies to any terms you use with a <u>Boolean expression</u>.

You can also use the Match Case option with terms that contain <u>wild-card characters</u>. Characters matched by wild-card characters, however, can be either uppercase or lowercase.

You cannot use the Match Case option with the Word Stemming, Thesaurus, or Sounds Like options. When any of these options are selected, the Match Case option is ignored.

Not all indexes support the Match Case option; index publishers must build support for the Match Case option into indexes. When an index does not support Match Case, selecting the option has no effect.

Proximity search option

Proximity changes the way AND searches work. With Proximity selected, words in an AND search must be within three or four pages of each other to be found. Without the Proximity option, two words may appear anywhere in a document and still be found by an AND search.

Also, the closer the words are in the document, the higher the relevance ranking of the document in which they appear.

Using the Word Assistant

The Word Assistant dialog box shows how the [Sounds Like](#), [Word Stemming](#), and [Thesaurus](#) search options affect a search query. Use the Word Assistant to see which terms, in addition to the search term, will be found with a search option selected.
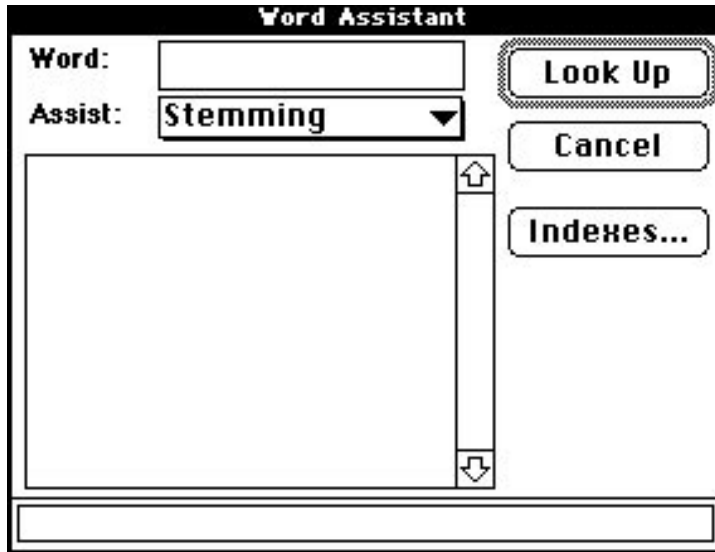
You can also cut terms from the Word Assistant dialog box and paste them into the Search window to construct quickly a list of terms to find. Cutting and pasting terms is especially useful for the long lists of words that are often suggested by the Thesaurus or Sounds Like option.

To see the effects of a search option on a search term:

1 Choose Search > Word Assist from the Tools menu. The Word Assistant dialog box appears.

2 If you do not want to use your current indexes, click Indexes to display the Index Selection dialog box and select the indexes you want to use.

3 Select the search option from the Assist menu.

4 Enter the search term in the Word text box.

5 Click Lookup. All the words that will be found are listed in the scrolling list.

→

To search with a term listed in the Word Assistant dialog box:

1 Click the [Search button](#) to display the Search window.

2 Choose Search > Word Assistant from the Tools menu to display the Word Assistant dialog box.

3 Enter the search term in the Word box, select the search option from the Assist list, and click Lookup. The related words are listed in the list box.

4 Double-click a word that you want to use. The word appears in the Word text box.

5 Double-click the word in the Word text box to select it, and press Ctrl/Command+C to copy it.

6 Click the Search window title to select it; then click the Find text box and press Ctrl/Command+V to paste the selected word into the text box.

7 Repeat steps 4 through 6 for each word you want to use; separate each pair of words in the Find text box with [OR](#).

Using the Refine Search feature

Refine search restricts the next search to only those documents found by the previous search. Only documents currently listed in the <u>Search Results window</u> are searched.

To use the Refine Search feature:

1 Use any combination of search terms, <u>Document Info field</u> values, <u>Date Info field</u> values, and <u>search options</u> to perform a search. The found documents are listed in the Search Results window.

2 With one or more documents in the Search Results window, click the <u>Search button</u> in the toolbar to open the Acrobat Search window.

3 Enter any new combination of search terms, Document Info field values, Date Info field values, and search options.

4 Press Ctrl (for Windows) or Option (on the Macintosh) and click the Refine button. (Pressing Ctrl or Option with documents in the Search Results list changes the Search button to the Refine button.) The Search Results window lists documents that were previously in the Search Results list and that meet the new search criteria.

Searching with AND, OR, and NOT (Boolean expressions)

You can use AND, OR, and NOT operators to build a Boolean expression that specifies combinations of words that a document must have to be returned by a search.

Suppose, for example, you were searching a collection of U.S. history articles for information about Noah Webster, who wrote the first U.S. American English dictionary. If you searched for *Webster*, you would find references to Noah Webster, but you would also find references to Daniel Webster, a nineteenth century U.S. senator from Massachusetts. To find references to Noah Webster and avoid references to Daniel Webster, you could search for articles that contain *Webster* and *dictionary* but not *congress* or *senate*.

You can also use Boolean expressions in [Document Info](#) field text boxes to specify a combination of field values that documents must contain to be returned from a search.

See the following topics for more information about Boolean expressions:

[AND searches](#)

[OR searches](#)

[NOT searches](#)

[Combining Boolean operators](#)

[Using parentheses to change the order of evaluation](#)

[Using operator names and parentheses in a Boolean expression](#)

For information about using Boolean expressions to specify Document Info field values, see also [Searching with AND, OR, and NOT (Boolean expressions)](#) or [Using Document Info fields](#).

AND searches

Use AND to specify two or more search terms that a document must contain to be returned by a search. For example,

installation AND instructions

finds only those documents that contain both *installation* and *instructions.*

You can also use the ampersand [&] to specify an AND search. For example,

installation & instructions

also finds documents that contain both *installation* and *instructions.*

→

Choosing the [Proximity search option](#) changes the way AND searches work. Without the Proximity option, terms in an AND search can be anywhere in a document. Searching for *red* and *blue*, for example, can find a document in which *red* is the first word on the first page and *blue* is the last word on the last page. With the Proximity option, however, terms in an AND search must be within two or three pages of each other to be found. Also, with the Proximity option selected, the closer the terms appear together, the higher the [relevance ranking](#) of the document that contains them.

OR searches

Use OR to specify two or more search terms, any of which a document can contain and can be returned by a search. For example,

    television OR tv

finds all documents that contain either *television* or *tv*.

You can also use a comma [,] or a vertical bar [|] to separate terms in an OR search. For example,

    television, tv

and

    television | tv

also find all documents with either *television* or *tv* or both.

Although any document that contains any term used with an OR operator is returned from a search, documents that contain both terms rank higher in the Search Results list. See <u>Relevance ranking</u> for more information.

NOT searches

Use NOT to exclude documents that contain a search term. For example,

> not quarterly

finds all documents that do not contain the word *quarterly*.

You can also use an exclamation point [!] to specify a NOT search. For example,

> ! quarterly

also finds all documents that do not contain the word *quarterly*.

Note: *You must place a space between the exclamation point and the search term.*

Combining Boolean operators

This topic explains how to combine Boolean operators in a search query.

Caution: Be careful when combining operators to build Boolean expressions. Unless you work with Boolean expressions frequently, it is easy to build an expression that does not mean what you think it means.

Combining AND and OR operators

When you combine AND and OR in the same expression, AND is evaluated before OR. For example,

    darwin OR origin AND species

finds all documents that contain *darwin* or that contain both *origin* and *species*.

Combining AND, OR, and NOT operators

When NOT is used with either or both of the AND and OR operators, it is evaluated before either the AND or the OR. For example,

    evolution AND not darwin

finds all documents that contain the word *evolution* but not *darwin*.

Using parentheses to change the order of evaluation

You can use parentheses to change the default order of evaluation for Boolean operators. For example,

    (darwin or origin) AND species

finds all documents that contain either *darwin* and *species* or *origin* and *species*. Parentheses can be nested to any depth.

Using operator names and parentheses in a Boolean expression

When you use a phrase that contains operator names or parentheses, the phrase must be enclosed in quotes. For example,

"cats and dogs" OR "cats & dogs"

finds all documents that contain either the phrase *cats and dogs* or the phrase *cats & dogs*.

Similarly, when you use a search term that contains an operator character or parentheses, the term must be enclosed in quotes. For example,

"(cats|dogs)"

finds all documents that contain the term (*cats|dogs*) as well as all documents that contain the phrase *cats dogs*. The phrase *cats dogs* is found because parentheses and the vertical bar are discarded before terms are entered into the index. They are also discarded when a search is performed. See How terms are indexed for details.

As another example,

"origin of the species" OR "step 36(a)" AND "cats and dogs"

finds all documents that contain the phrase *origin of the species*, or that contain both the phrase *step 36(a)* and the phrase *cats and dogs*.

When the following words and characters are used within a phrase, the phrase must be enclosed in quotation marks:

AND	&[AND operators]
OR	|,[OR operators]
NOT	![NOT operators]

Viewing search results

When you enter a search query, Acrobat Reader lists in the Search
Results window the documents it finds.

Search Results

Score Title   Found 18 out of 42 documents.

◑ ROMEO AND JULIET
◑ MACBETH
◑ THE WINTER'S TALE
◑ A MIDSUMMER NIGHT'S DREAM
◕ ANTONY AND CLEOPATRA
◔ CORIOLANUS
◔ KING HENRY VIII
◔ HAMLET
○ KING LEAR
○ THE TAMING OF THE SHREW
○ THE FIRST PART OF KING HENRY VI

[**View**]   [**Info...**]

SHAKES:TRAGEDY.PDF:ROMEOSP.PDF

To view a document returned from the search:

1 Select the document name, and click View (or double-click the document's name).

2 Use the Search buttons on the toolbar to view all the pages containing search terms:

▪ To highlight the next occurrence of a search term, click the Search Next button 📄. When the last occurrence of a search term in a document is displayed, Search Next displays the first occurrence in the next document. Search Next displays documents in the order that they are listed in the Search Results window. Pressing the Shift key and clicking Search Next displays the first occurrence of a search term in the next document.

▪ To highlight the previous occurrence of a search term, click the Search Previous button 📄. When there are no more occurrences of a search term in a document, the last occurrence of a search term in the previous document is displayed. Search Previous displays documents in the reverse of the order in which they are listed in the Search Results window. Pressing the Shift key and clicking Search Previous displays the first occurrence of a search term in the previous document.

Changing the order in which found documents are listed

By default, documents are listed in the order of their score or relevance ranking, which places the documents most likely to contain relevant information at the top of the list. With the Acrobat Search Preferences dialog box, you can change the order in which documents are listed. For example, you can sort documents by creation date, by title, or by author. Choose Preferences > Search from the Edit menu to display the Acrobat Search Preferences dialog box.

Changing the way search terms are displayed

By default, all the search terms on a page are highlighted and the Search Next and Search Previous buttons display the next and previous pages with found terms. You can, however, change the way found terms are highlighted by changing the Highlight option in the Acrobat Search Preferences dialog box:

▪ **By Word** highlights search terms one at a time; each click of the Search Next or Search Previous button moves the highlighting from one term to the next.

- **No Highlighting** displays pages with found terms without highlighting terms; each click of the Search Next or Search Previous button displays a new page.

- **By Page**, the default, displays pages with all found terms highlighted; each click of the Search Next or Search Previous button displays a new page.

Choose Preferences > Search from the Edit menu to display the Acrobat Search Preferences dialog box.

Limiting the number of documents returned by a search

By default, no more than 100 documents are listed in the Search Results list. You can, however, change this limit to a higher or lower number in the Acrobat Search Preferences dialog box.

The minimum is 1 document; the maximum is 1000 documents.

Choose Preferences > Search from the Edit menu to display the Acrobat Search Preferences dialog box.

→

Keeping the Search Results window open

By default, the Search Results window closes when you view a document returned from a search. You can, however, keep the Search Results window open while viewing search results by clearing the Hide on View option in the Acrobat Search Preferences dialog box. Choose Preferences > Search to display the Search Preferences dialog box.

When you keep the Search Results window visible while viewing search results, the Search Results window reports how many documents were found, how many pages in the current document contain a search term, and which page in the current document is being displayed.

Search Results

Score Title	Found 18 out of 42 documents.

◑ ROMEO AND JULIET
◑ MACBETH
◑ THE WINTER'S TALE
◑ A MIDSUMMER NIGHT'S DREAM
◔ ANTONY AND CLEOPATRA
◔ CORIOLANUS
◔ KING HENRY VIII
◔ HAMLET
○ KING LEAR
○ THE TAMING OF THE SHREW
○ THE FIRST PART OF KING HENRY VI

[**View**] [Info...]

Document 8 of 18: 1 of 1 pages of hits.
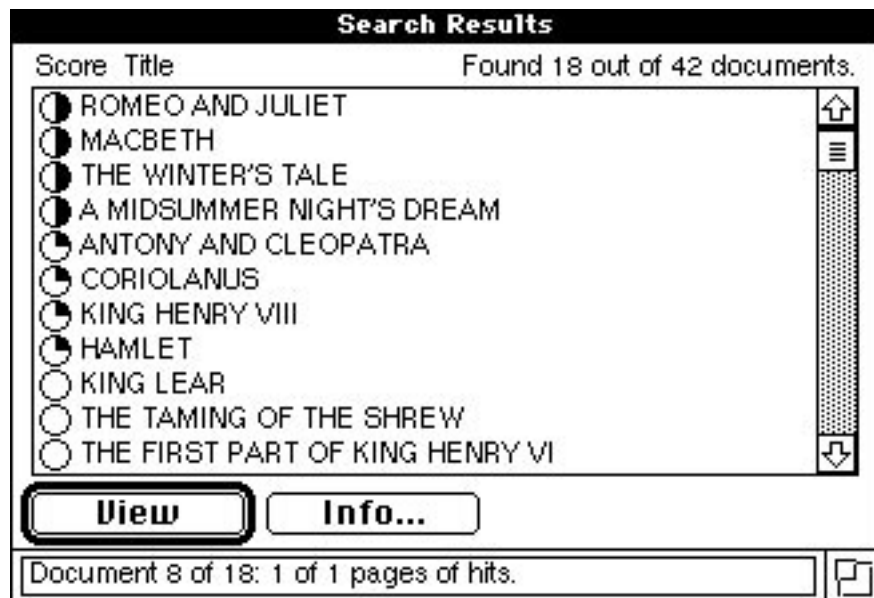
Relevance ranking (score)

Each document returned from a search is assigned a relevance ranking. A document's relevance ranking indicates how likely it is that the document contains the information for which you are searching.

The Search Results window uses five icons to indicate a document's relevance ranking:

A full circle indicates a very high relevance ranking; the document is very likely to contain relevant information.

A three-quarter circle indicates a high relevance ranking; the document is likely to contain relevant information.

A half circle indicates a medium relevance ranking; the document probably contains relevant information.

A one-quarter circle indicates a low relevance ranking; the document might contain relevant information.

An empty circle indicates a very low relevance ranking; the document is unlikely to contain relevant information.

The method used to determine a document's relevance ranking depends on the kind of search performed. With a single search term, for example, the relevance ranking indicates how frequently the term appears in the document. When two search terms are separated by an [OR operator](#), however, documents that contain both terms have a relevance ranking higher than documents that contain just one of the terms. Similarly, when you use the [Proximity search option](#), the closer search terms are within a document, the higher the relevance ranking for the document.

Another factor that contributes to a document's relevance ranking is the density of search terms relative to other terms in the document. The higher the proportion of search terms in a document, the higher the document is ranked.

Setting search preferences

You can use the Acrobat Search Preferences dialog box to

- Control whether the Search window closes when you perform a search

- Control the size and contents of the Search window. (You can show or hide the [Document Info fields](#), the [document date fields](#), and the [search options](#).)

- Control whether the Search Results window closes when you view a document returned from a search

- Control the number of found documents listed in the Search Results window, and the order in which they are listed

- Control whether all search terms on a page are highlighted, whether terms on a page are highlighted one at a time, or whether no terms are highlighted

- Control whether Macintosh file server volumes that contain indexes you are using are automatically mounted when you start a search

See [Query preferences](#) for details about the preference settings that control the Acrobat Search window. See [Results preferences](#) for details about the preference settings that control the Search Results window. See [Indexes preference (Macintosh only)](#) for details about the automount setting for file servers that contain indexes.

Query preferences

The Query preferences control the size, contents, and behavior of the Acrobat Search window.

```
┌─Query─────────────────────────────────────────┐
│ ☒ Show Fields          ☒ Show Date            │
│ ☒ Show Options                                 │
├ ─ ─ ─ ─ ─ ─ ─ ─ ─ ─ ─ ─ ─ ─ ─ ─ ─ ─ ─ ─ ─ ─ ┤
│ ☒ Hide on Search                               │
└────────────────────────────────────────────────┘
```

▪ **Show Fields** displays the <u>Document Info fields</u> with the Search window.

▪ **Show Options** displays the <u>search options</u> with the Search window.

▪ **Show Date** displays the <u>Date Info selection options</u> with the Search window. The Date Info options let you limit the results of a search to documents that were created or modified during a specific time period.

▪ **Hide on Search** closes the Search window when you perform a search; clear this option to keep the Search window displayed while you view search results.

Results preferences

The Search Results preferences control the behavior of the [Search Results window](#).

```
┌ Results ─────────────────────────────────────┐
│  Sort By:     │Score              ▼│          │
│                                                │
│  Show Top:    │100│ Documents                 │
│                                                │
│  ☐ Hide on view                               │
└────────────────────────────────────────────────┘
┌ Highlight ───────────────────────────────────┐
│ Display:   │By Page         ▼│                │
└────────────────────────────────────────────────┘
```

- **Sort By** controls the order in which documents are listed in the Search Results window. By default, documents are sorted by score, which lists them in the order of their [relevance ranking](#). You can also sort found documents by creation date, modification date, or by one of the Document Info field values, such as Title or Author. When you sort found documents by Document Info field values, they are listed alphabetically by field values.

→

When you sort by Created or Modified date values, documents are sorted by date in descending order. The most recent dates are listed first.

Documents with empty fields are listed before documents with field values.

When documents are sorted by score, documents with the same score are listed alphabetically by title (or filename if they have no title).

When documents are sorted by any value other than score, documents with the same sort value are sorted by score.

- **Show Top __ Documents** defines the maximum number of documents you want returned from a search. Enter any value between 1 and 1000.

- **Hide on View** closes the Search Results window when you view a document returned from a search; clear this option to keep the Search Results window visible while you view search results.

- **Highlight Display** controls how search terms are highlighted and how the Search Next and Search Previous buttons work.

- **By Page** highlights all the search terms on a page. Search Next and Search Previous display the next or previous page with found terms.

- **By Word** highlights found search terms one at a time. Search Next and Search Previous highlight the next or previous found term.

- **No Highlighting** does not highlight search terms. Search Next and Search Previous display the next or previous page that contain found terms.

Indexes preference (Macintosh only)

The Indexes preference controls whether network file servers that contain currently selected indexes are automatically mounted when you start a search.

Indexes
☒ **Automount servers**

Select Automount Servers to connect automatically to file servers that contain selected indexes. When Automount Servers is not selected, you must manually mount servers that contain indexes before you can use the indexes.

About Acrobat indexes

The Acrobat Catalog program creates [full-text indexes](#) for collections of PDF documents. The Search feature uses these indexes to locate words quickly in documents that make up indexed collections.

Words and terms included in an index

By default, an Acrobat index lists every word, number, and alphanumeric term that appears in every document in the indexed document collection. An *alphanumeric term* is a sequence of letters and numbers. See [How terms are indexed](#) for details.

Stopwords

To reduce the size of an index, an index publisher can exclude stopwords from an index. Stopwords are not indexed. Usually, stopwords are common words that do not serve to identify topics or subjects. Carefully chosen stopwords can reduce the size of an index without reducing its usefulness.

Typical stopwords include

- Articles such as *the* and *a*
- Prepositions such as *for*, *by*, and *of*
- Conjunctions such as *or* and *and*

Each index can be built with a different stopword list.

Because they are not placed into the index, you cannot find stop-words with the Search command. Also, you cannot find a phrase that contains a stopword.

See [Typical stopword list](#) for a list of common stopwords.

Numbers

Another way index publishers can reduce the size of an index is to exclude numbers from an index. See [How Acrobat Catalog excludes numbers](#) for a description of how the Catalog program identifies numbers.

Full-text indexes

A *full-text index* is an alphabetized list of all the words and alphanumeric terms that occur in a collection of documents. For each word and alphanumeric term in the document collection, a full-text index lists the documents and pages where the word or term appears. Given a full-text index, search software can quickly locate all the occurrences of a search term without opening and searching the documents. The following table shows part of a full-text index for Shakespeare's plays.

Term	Document (Play)	Pages
abed	*All's Well That Ends Well*	150
	As You Like It	36
	Cymbeline	86
	Macbeth	35
	Othello	82
	Romeo and Juliet	111, 142
	Twelfth Night	38

→

Term	Document (Play)	Pages
abel	First Part of *King Henry VI*	25
Abel's	*King Richard II*	9
Abergavenny	*Henry VIII*	2, 6, 9, 10, 11, 19, 28
abet	*King Richard II*	71
abetting	*Comedy of Errors*	33

Note: *Indexes distinguish between uppercase and lowercase letters only when publishers build them to support the Match Case option. Otherwise, all index terms are spelled with lowercase letters.*

See [How terms are indexed](#) for a description of how the Acrobat Catalog program recognizes words and terms.

How terms are indexed

When indexing a PDF document, the Acrobat Catalog program uses separator characters and <u>white space</u> to recognize where one word or term ends and the next word or term starts. For example, in the sentence "We have nothing to fear but fear itself," the Acrobat Catalog program uses the space characters (or white space) and the period to identify seven terms: *we*, *have*, *nothing*, *to*, *fear*, *but*, and *itself*. Having identified the terms, Acrobat Catalog eliminates any <u>stopwords</u> defined for the index. After removing the stopwords, the Catalog program indexes the remaining terms. (Using the <u>typical stopword list</u>, the Catalog program would eliminate *to* and *but* and index the terms *we*, *have*, *nothing*, *fear*, and *itself*.)

Separator characters

Most punctuation marks and all symbols are separator characters. The most common separator characters are the space and common punctuation characters such as periods, commas, colons, semicolons, exclamation points, question marks, parentheses, and quotation marks.

→

The Catalog program removes separator characters when it recognizes words and alphanumeric terms; only words and alphanumeric terms are placed into the index.

As an example, consider the following two sentences:

Tell Mary Jones-Smith we're having roast chicken tonight.
(It's not a secret anymore.)

Using no stopwords, the Acrobat Catalog program would index the following terms from these two sentences:

tell	*mary*	*jones*	*smith*	*we're*
having	*roast*	*chicken*	*tonight*	
it's	*not*	*a*	*secret*	*anymore*

Note*: The apostrophe is one of the punctuation characters that is not considered a separator.*

The following table shows more examples of how separator characters indicate where terms begin and end.

Character sequence	Indexed terms (and phrases)
(415)555-1212	415, 555, 1212
#33-223A-BB	33, 223A, BB
2.5*(14-4)=25	2.5, 14, 4, 25
c:\project6\docs\proposal.doc	c, project6, docs, proposal, doc
mother-in-law	mother, in, law
6/18/50, 11-10-1954	6, 18, 50, 11, 10, 1954

Because the Search command also removes separator characters from search terms before beginning a search and because phrases do not have to be enclosed in quotes, you can usually enter a search term with separator characters and find the term as a phrase. Searching for a phone number such as

    (415)555-1212

searches for and finds the phrase

    415 555 1212

Every character except the characters listed in the table below is a separator character.

Characters that form words and alphanumeric terms in an index

all letters	digits [0–9]	single quote [']
right apostrophe [']	grave [`]	diereses [¨]
acute [´]	cedilla [¸]	macron [¯]
breve [˘]	dotaccent [˙]	ring [°]
hungarumlaut [˝]	ogonek [˛]	caron [ˇ]

→

Characters removed from indexed terms

Some characters, such as ligatures, in Type 3 fonts are dropped from indexed terms. For example, you cannot search for the term *Æsop's fables* if the phrase is set in a Type 3 font because the *Æ* is removed before the words are placed in the index.

Ligatures are removed only from words in Type 3 fonts. Normally, you can search for a word by typing all the characters of the word. For example, you can search for the word *Æsop* by typing **Aesop**.

How Acrobat Catalog excludes numbers

Index publishers can exclude numbers from Acrobat Catalog indexes. The Acrobat Catalog program defines a number to be a sequence of one or more numeric digits [0 through 9] that is

- Optionally preceded by a minus sign [-]

- Optionally separated by one or more commas [,] or periods [.]

Here are some examples of the kinds of numbers that the Acrobat Catalog program can exclude from an index:

    23   23.54671   -37.1   3,012,222   44.223.835,5

When numbers are separated by other punctuation characters and symbols, the Catalog program removes these other characters before determining whether the remaining characters are a number.

($3,000.23) becomes 3,000.23, which is an excluded number.

(37A) becomes 37A, which is a search term and not a number.

White space

When applications print to PostScript printers (or create PostScript files for the Acrobat Distiller), they often control the spaces between words by placing the words onto the page as separate units. That is, words in a PostScript file may not be separated by space characters. Rather, words may be individually placed onto the page as separate text objects.

> How   now   brown   cow   ?

When a block of text, such as a paragraph, is built as a group of individual text objects, the Acrobat Catalog program must reconstruct the block of text by finding and assembling the text objects in the correct order. The Acrobat Catalog program interprets each text object as a word or term, and it indexes each text object as a separate word or term. In this case, words and terms are not separated by space characters; they are separated by the white spaces between the text objects.

Typical stopword list

Here is a list of typical [stopwords](#):

About	about	Above	above	After	after
Also	also	Although	although	An	an
And	and	As	as	At	at
Be	be	Because	because	Before	before
Between	between	But	but	By	by
Either	either	For	for	From	from
Further	further	Here	here	However	however
If	if	In	in	Into	into
Now	now	Of	of	On	on
Or	or	Other	other	Out	out
Over	over	Since	since	Such	such
Than	than	That	that	The	the
Then	then	There	there	Therefore	therefore
These	these	This	this	Those	those
Through	through	Thus	thus	To	to
Under	under	Until	until	Upon	upon
What	what	When	when	Where	where
Whether	whether	Which	which	While	while
With	with	Within	within	Without	without

Troubleshooting searches

This topic provides advice for solving problems you might encounter when using the Search command. Solutions are described under the error message you will see.

The index is unavailable

Indexes might be unavailable for one or more of the following reasons:

- The connection to the network server containing the index has been lost.

- The index is being purged and rebuilt; this maintenance operation takes time but can significantly reduce the size of an index.

- The index has been moved to a new location.

- The index has been deleted.

Contact your network administrator if you suspect you have lost your connection to a server.

→

Contact the person responsible for the index if you suspect the index is being purged and rebuilt, or if you suspect that it has been moved or deleted.

Highlights will not be displayed

Acrobat Reader displays this message when it tries to open a PDF document that has changed since the index was built. The change may not have altered the locations of the words on the page. Using Acrobat Exchange to add a note to a PDF document or to change the security settings, for example, causes this message to be displayed, even though the contents of the document have not changed.

If the change to the file did alter the locations of words, the search term highlights will be incorrect. For example, if a page was deleted, highlights will indicate the wrong words.

If you think the change to the PDF document did not alter the locations of words in the document, or if you think you will be able to tell whether the highlights are correct, click the Highlight Anyway button to close the message box and view the document.

Note: Index creators can suppress this message by selecting the Optimize for CD-ROM option in the Adobe Acrobat Catalog Options dialog box. See Optimize for CD-ROM option *in the* Adobe Acrobat Catalog Online Guide*.*

No documents were found that matched your query

Acrobat Reader displays this message when no documents in the currently selected indexes meet the criteria you specify with a search query. If you get this message, try these steps:

▪ Make sure you have selected the indexes that contain the documents with the information you seek.

▪ Make sure you are not using Document Info or Date Info field values—or search options from the previous query by mistake.

▪ Make sure you are not searching for a [stopword] in an index that has stopwords.

▪ Make sure you are not searching for a number in an index from which numbers have been excluded.

▪ If you are using an operator such as AND or NOT in a phrase, make sure the phrase is enclosed in quotation marks (" ").

Out of file handles

You are running too many programs or you have too many windows open. Close some windows, and try the query again. If you are running two or more programs, quit all programs except Acrobat Reader and retry the query.

Not enough memory (Macintosh)

If you get a message saying you do not have enough memory for Acrobat Reader to find the documents that meet your search criteria, you can try increasing the memory available to the Acrobat Reader program.

1 If Acrobat Reader is running, quit the application.

2 From the Finder, click to select the Reader program icon.

3 Choose Get Info from the File menu. The Get Info dialog box appears.

4 In the Memory Requirements box, increase the Minimum size value. For example, you might increase the Minimum size value to 2500 or 3000. (The maximum amount of memory available for the application is limited by the amount of memory installed in your Macintosh.)

$\rightarrow$

5 Click the close box in the upper left corner of the Get Info dialog box.

6 Start Acrobat Reader. If a message appears, telling you that there is not enough memory to run the program, decrease the Minimum size value in the Acrobat Reader Get Info box.

Keyboard shortcuts

Tool or key	Plus	Result
Ctrl+Shift	F	Displays the Search window
Ctrl+Shift	X	Displays the Index Selection dialog box
Ctrl+Shift	J	Displays the Search Results window
Ctrl+Shift	W	Displays the Word Assistant dialog box
Ctrl+Shift	Y	Displays the first occurrence of search term in previous document
Ctrl	Y	Displays the previous page with search term, or the previous term if Display By Word is selected

→

| Ctrl | U | Displays the next page with search term, or the next term if Display By Word is selected |
| Ctrl+Shift | U | Displays first occurrence of search term in next document |