



Advanced Adobe Acrobat (PDF)

2001 FDI Track D-Intermediate web
Course Development

Presented by Troy Herring

E-Mail troyfish@vt.edu Phone 540.231.4826



Searching Across Multiple PDF Documents

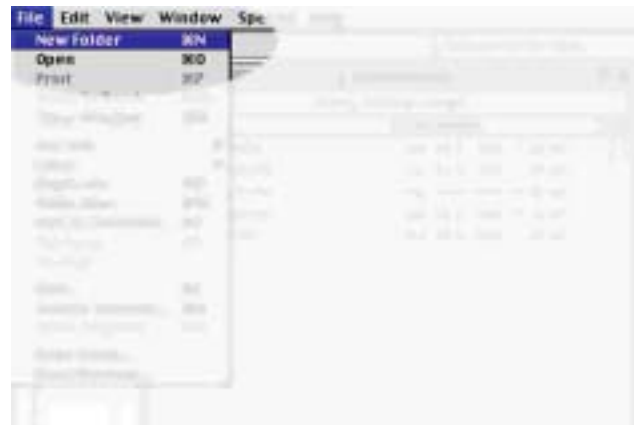
The Adobe Acrobat Search command allows you to perform full text searches of PDF documents and collections that have been indexed with the Acrobat Catalog feature. The Search command is more powerful and flexible than the Find command. It lets you search multiple documents, and define advanced query criteria. Search is faster than Find because it reads the index rather than the entire document.

To search an index created using Adobe Acrobat Catalog, you first **select the indexes** to search, **define a search query**, and **view the occurrences** of the search term within the documents you selected to review. You can also select the documents to review from those returned by the search. A search query is an expression made up of text and other items to define the information you want to find.

How to create a Searchable Index.

You can use the Adobe Acrobat Catalog command to create a full-text index of your PDF documents and document collections. A full-text index is a searchable database of all the text in a document or set of documents. The following sections guide you through the steps to create indexes.

- 1) Create and name a new folder to contain the document collection and the index files.
- 2) Move or copy all the PDF documents you want to index into the new folder. (Drag and Drop works just fine)
- 3) Open Acrobat 5.0
- 4) Choose **Tools > Catalog** to open the Adobe Acrobat Catalog dialog box.





Advanced Adobe Acrobat (PDF)

2001 FDI Track D-Intermediate web
Course Development

Presented by Troy Herring

E-Mail troyfish@vt.edu Phone 540.231.4826



Searching Across Multiple PDF Documents

5) Click the **New Index** button.



6) In the **New Index Definition** dialog box, enter a name for the index.

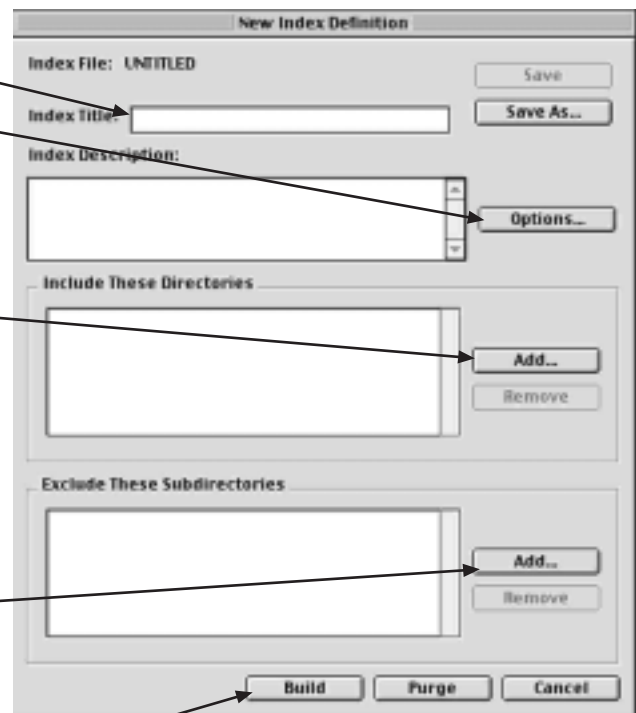
7) To change index options, select **Options**, make the necessary selections, You can exclude specific words (stop-words) from the index, **exclude numbers**, and disable some of the user's search options (**Case Sensitive, Sounds Like, Word Stemming**).

8) In the **Include These Directories** section, click **Add** and navigate to the folder you want to include in the index. Select the folder and click **OK** (Windows) or **Choose** (Mac OS).

NOTE: On Mac OS, if you do not plan on moving the index and document collection, you can add folders from multiple servers or disk drives.

9) To exclude folders from the index, click **Add** under **Exclude These Subdirectories**, navigate to the folders, and select them. You cannot exclude individual files inside a folder; you have to exclude the entire folder.

10) Click **Build**.





Advanced Adobe Acrobat (PDF)

2001 FDI Track D-Intermediate web
Course Development

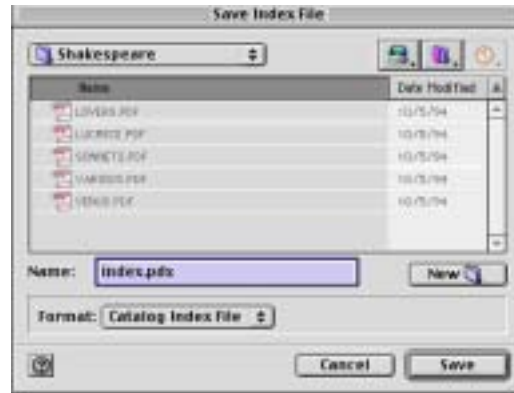
Presented by Troy Herring

E-Mail troyfish@vt.edu Phone 540.231.4826



Searching Across Multiple PDF Documents

- 11) Click **Build**. (Retain the **.pdx** extension provided for the filename.)
- 12) Determine the location where the index will be saved.
- 13) Click **SAVE**.



How to add or remove a Searchable Index from your choices.

You can search any or all indexes displayed in the Index Selection dialog box. Dimmed indexes are not available for searching.

- 1 Choose **Edit > Search > Select Indexes** to list the currently available indexes and to **add** or **delete** indexes, and then do one of the following in the **Index Selection dialog box**:

- A) To **add an index** to the available indexes list, click **Add**, navigate to the index, and double-click on the index file.
- B) To **remove an index**, select the index name, click **Remove**, and then click **OK**.
- C) To select or deselect an index, select the box for the index, and then click **OK**.

NOTE: Dimmed indexes are currently unavailable for searching.

- D) To view information about an available index, highlight the index name, click **Info**, and then click **OK**. The information displayed includes the build date, creation date, number of documents in the index, location of the index, status, and information provided by the builder of the index.





Advanced Adobe Acrobat (PDF)

2001 FDI Track D-Intermediate web

Course Development

Presented by Troy Herring

E-Mail troyfish@vt.edu Phone 540.231.4826



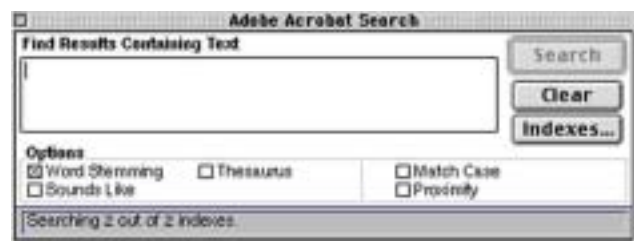
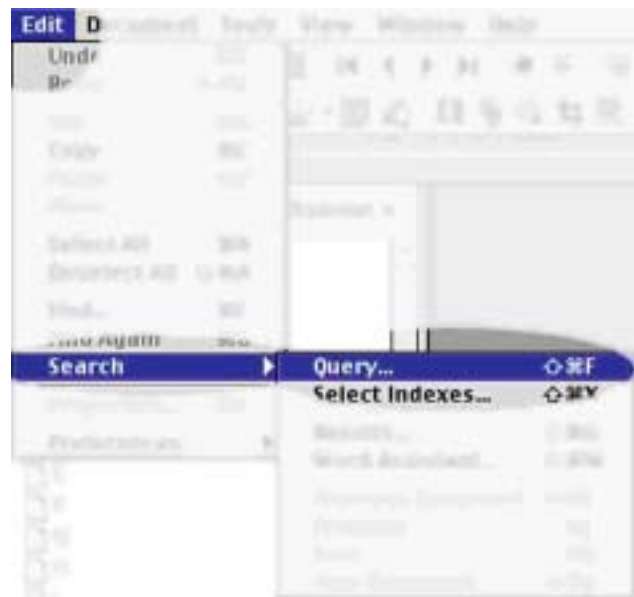
Searching Across Multiple PDF Documents

How Search many PDF Documents.

The Search command allows you to perform a search on indexed PDF documents. You can search for a simple word or phrase, or you can expand your search query by using wild-card characters and Boolean operators. You can use the search options to further refine your search. And if document and date information were provided for the documents you are searching, you can use that information to further narrow your search.

The text that you type in can be a single word, a number, a term, or a phrase. It can be a word, with or without wild-card characters (*, ?), or any combination of letters, numbers, and symbols. Because you can use Boolean operators in the text box, you must enclose any search term that includes and, or, or not in quotation marks. You can also use the operators =, ~, and != with text, but only to perform exact matches, contains, and does not contain searches, respectively. You can use comparison operators (<, <=, >, >=) with values of the same type.

- 1) Launch Acrobat 5.0
- 2) Choose **Edit > Search > Query**.
- 3) Enter the text you want to search for in the **Find Results Containing Text** box. To clear the Search dialog box and redefine the search, click **Clear**.
- 4) Select any combination of the search options:
 - A) **Word Stemming**. Finds words that contain part of (a word stem) the specified search word. It applies to single words, not phrases; does not apply to words that contain wild-card characters (*, ?); finds words that end in ing, ed, s, ion, and so on, but not er; and cannot be used with the **Match Case** option. **Word Stemming** works only for indexes built with this option.
 - B) **Sounds Like**. Finds different spellings for proper names. It applies to single words, not phrases; does not apply to words that contain wild-card characters; and cannot be used with the **Match Case** option. **Sounds Like** works only for indexes built with this option.
 - C) **Thesaurus**. Finds similar words that appear in the documents you are searching, not necessarily all the similar words you might find in a complete thesaurus. It applies to single words, not phrases; does not apply to words that contain wild-card characters; and cannot be used with the **Match Case** option.





Advanced Adobe Acrobat (PDF)

2001 FDI Track D-Intermediate web
Course Development

Presented by Troy Herring

E-Mail troyfish@vt.edu Phone 540.231.4826



Searching Across Multiple PDF Documents

D) **Match Case.** Limits the results of the search by finding only those documents that contain words with the same capitalization. It can be used with a **Boolean** expression and with terms that use wild-card characters. Characters matched by wild-card characters can be either uppercase or lowercase.

E) **Proximity.** Limits the results of simple **AND** searches to one pair of matches per document—the pair closest together. The two matches must be within three pages or fewer of each other. This option is useful for locating a document that concentrates on some topic of interest. Proximity affects relevancy ranking in searches. The closer the matches are within a document, the higher the ranking. **Proximity** does work with complex **AND** searches—such as, Hawaii AND (cruise OR fly).

5) Select **Search.** The Search dialog box is hidden, and documents that match your search query are listed in the **Search Results** window in order of relevancy. The results are displayed in the Search Results window. Documents more likely to contain relevant information are listed first on the list. The relevancy ranking of each document is indicated by an icon. The degree of fill in the circle in the icon indicates the probability that the document contains the search information.

- i) A **solid fill** indicates a high probability that the document contains your search term
- ii) an **empty circle** indicates a low probability that the document contains your search term.

6) View the document by doing one of the following:

A) **Select** the document, and click the **View** button.

or

B) **Double-click** the document.

