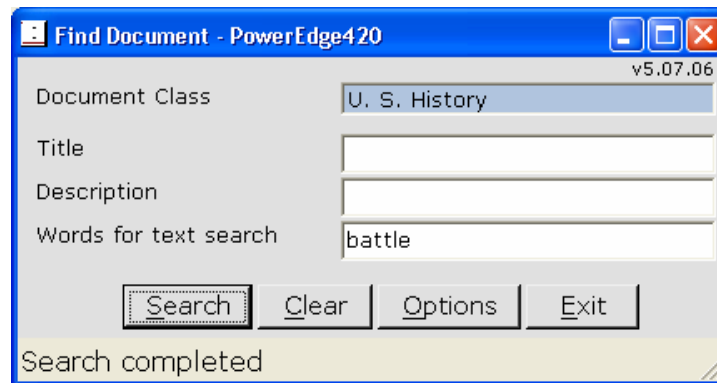


Searching for words and phrases with the Find Document application

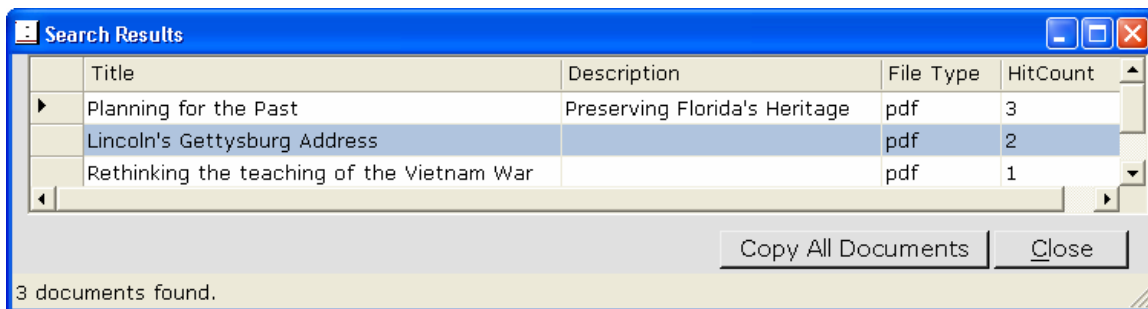
The Find Document application offers the capability to search for words or phrases within documents that have been saved in searchable formats, such as Microsoft Word, Microsoft Excel and the searchable versions of Adobe PDF. The Find Document application uses the Indexing Service provided by the Windows 2000, Windows Server 2003 and Windows XP operating systems to locate documents quickly using text searches.

One customer frequently searches *OurLibrary* document classes containing more than 400,000 pages for specific words and phrases. A list of the results usually appears within ten seconds.

First consider the simple case of searching for documents containing a single word. The word is entered into to “Words for text search” textbox as shown below:

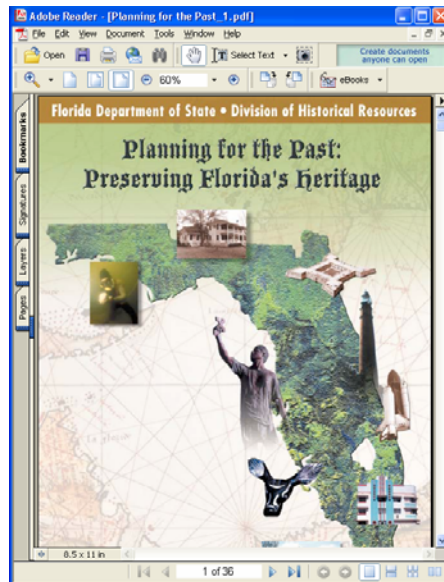


When the Search button is clicked, a grid will be displayed that shows all documents that were found that contain the word “battle” or any similar words that start with “battle”. For this example, the following grid was shown:

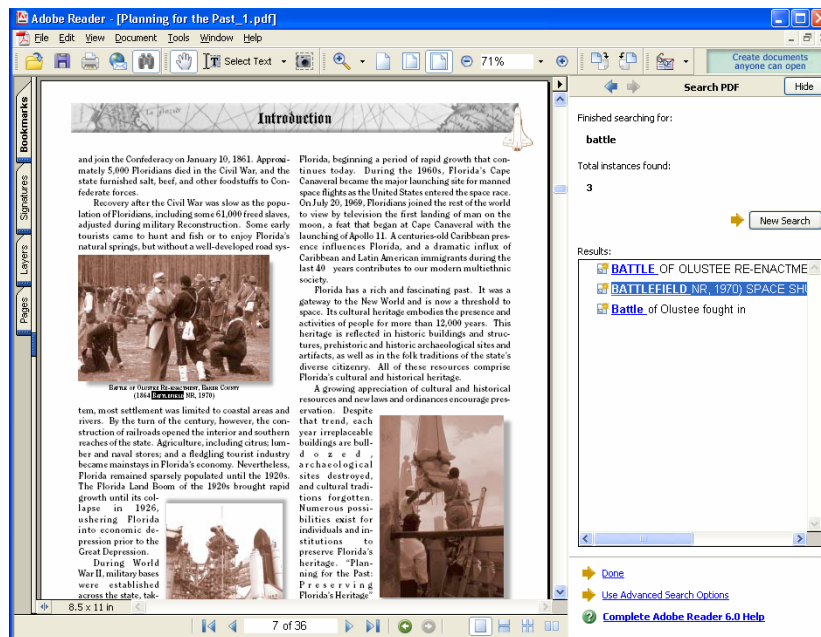


Note that the “Hit Count” Shows the number of times that words starting with “battle” were found. Double-clicking the mouse on the leftmost column in a document row will open and display the documents.

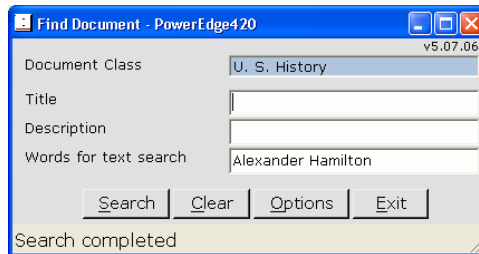
In this example, double-clicking the mouse on the first row opens Adobe Reader as shown below:



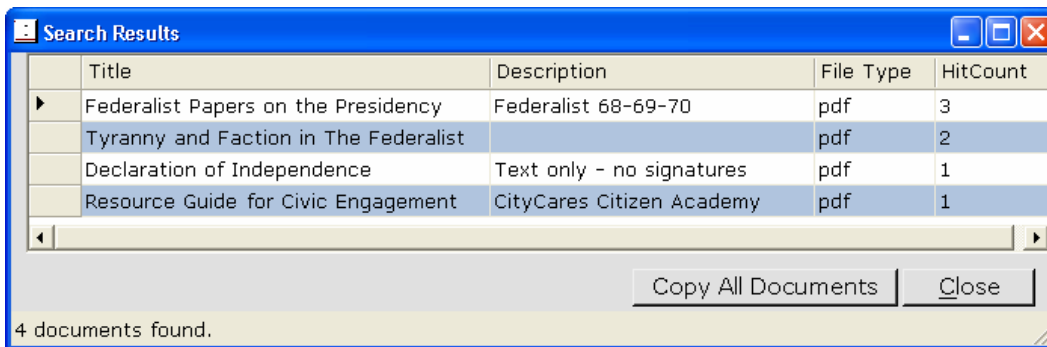
The user can click on the binoculars icon to perform a search within Adobe Reader (or Adobe Acrobat). Next, the user can enter the word to search for and click the Search button. The results will be shown on the right side of the windows, as demonstrated below:



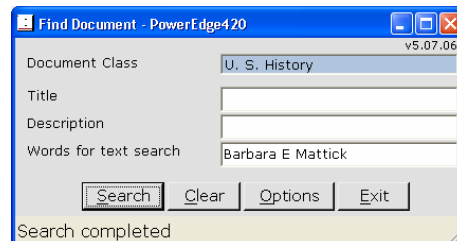
Next, we can search for a phrase or name simply by entering the text into the to “Words for text search” textbox as shown below:



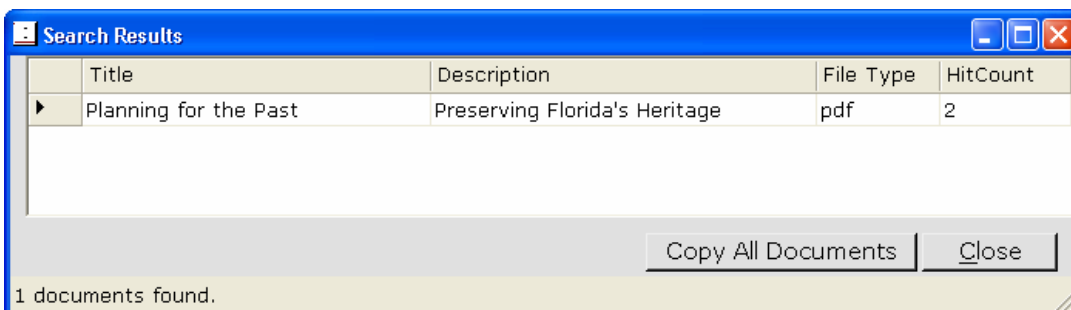
The results are shown in a grid, as follows:



Searching by names can be tricky due to variations as well as middle names and initials. The following search can locate a person's name within a document:

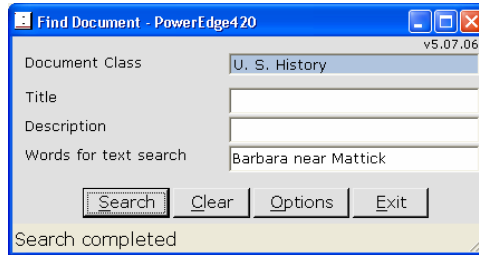


The grid containing the results is shown below:



However, a search for “Barbara Mattick” fails to return any documents. A search for “Mattick” will return documents but the results may include persons named Mattick with other first names.

Several reserved words can be used with text searches, including “and”, “or” and “near”. We could search for “Barbara NEAR Mattick” to find a name without specifying the middle name or initial, as shown below:



The reserved words are defined in the following table:

Reserved Word	Usage
AND	Locate documents containing both words – such as “dog AND cat”.
OR	Locate documents containing either word – such as “dog OR cat”.
NEAR	Locate documents where the words appear within 50 words of each other.

Notes:

1. Text searches using the Windows Indexing Service are usually case-insensitive. However, searching for a name that includes an initial may require capitalizing the initial.
2. Convention index value searches can be combined with text searches. For example, we could search for documents that were added to OurLibrary after January 1, 2005 that contain a specific word or phrase. Find Document limits the search to documents that meet the conventional criteria and then performs a text search using only those documents.
3. When documents are initially added to an *OurLibrary* repository, the documents may not be added to the Indexing Service for several minutes or longer. When the repository computer is busy, Windows suspends the Indexing Service.
4. Technical information regarding the Windows Indexing Service can be found at www.msdn.microsoft.com. You could then search the MSDN site for “Indexing Service” to return a list of articles regarding this service. Note that the Find Document app uses the Indexing Service Query Language Dialect 2 by default.
5. The *OurLibrary* repository server or workstation requires configuration to enable text searching. The Adobe I-Filter must be installed in order to search Adobe PDF files.