Index

AND Connector

Use the AND connector in a search request to connect two expressions, both of which must be found in any document retrieved.   For example:

> *apple pie and poached pear* would retrieve any document that contained both phrases.

> *(apple or banana) and (pear w/5 grape)* would retrieve any document that (1) contained either *apple* OR *banana*, AND (2) contained *pear* within 5 words of *grape*.

Alphabet Customization

dtSearch uses the Windows character set to decide how to index and search text.   If you need more control over the way dtSearch handles characters, choose **Index Options** in the Options menu, then click on the **Edit Alphabet** button next to the alphabet file name.

The Edit Alphabet dialog box displays a list of all of the characters and how dtSearch classifies each one. dtSearch classifies characters into four categories: **letter**, **space**, **hyphen**, and **ignore**.

letter	A searchable character. All of the characters in the alphabet (a-z and A-Z) and all of the digits (0-9) should be classified as letters.
space	A character that causes a word break. For example, if you classify the period (".") as a **space** character, then dtSearch would process *U.S.A.* as three separate words: *U, S* and *A.*
hyphen	Usually only two characters get classified as **hyphen**: the ASCII hyphen ("-") and the underscore ("_"). By default, dtSearch indexes a hyphenated word as four different words to cover all possible ways to search for the word. *Full-text* would be indexed as *full*, *text*, *fulltext*, and *full-text*. To change hyphen handling, choose **Index Options** in the Options menu.
ignore	A character that is disregarded in processing text. For example, if you classify the period as **ignore** instead of **space** then dtSearch would process *U.S.A.* as one word: *USA.*

For characters that are letters, you can specify whether the character is a lower case or upper case letter, whether or not it has an accent, and the lower case or unaccented equivalent.

Automatic Indexing

You can use the System Agent in Microsoft Plus! for Windows 95 or other program schedulers to automate indexing functions.   To set up dtSearch to update an index automatically:

1. Create the index.

2. Choose **Create Batch Script** in the Index menu.   The Create Batch Script dialog box will appear.   It looks just like the Update Index dialog box, but when you click OK instead of indexing it will create a text file describing the indexing job you have set up.

3. In the scheduling program, there will be a space for the command to execute.   Enter the location of the dtSearch program (DTSWIN.EXE) and, using the /SCRIPT command-line option, the name of the script you created.   For example, if you installed dtSearch in C:\DTSEARCH and created a script named EXAMPLE.JOB, the command to execute would be:

```
c:\dtsearch\dtswin.exe /script c:\dtsearch\example.job
```

4. Set up the time to run the indexing job in the scheduling program.

Notes

1. See "Batch Indexing and Searching" for more information about dtSearch batch scripts.

2. Indexing documents uses a lot of system resources and will slow other programs noticeably. Therefore, it is strongly recommended that you schedule indexing for times when you are certain that you will not be using your system for other purposes.

Batch Indexing Commands

Creating an Index

To create an index, set up a script like this:

```
DTSEARCH SCRIPT
SET INDEXNAME = "MyIndex"
SET INDEXPATH = C:\DOCINDEX
CREATE
```

There are two options you can use with CREATE: CASE and ACCENT. Use CASE if you want the index to be <u>case-sensitive</u>. Use this ACCENT if you want the index to be <u>accent-sensitive</u>. To use one of these options, simply list them on the same line as the CREATE command.

WARNING: If you create an index that already exists, dtSearch will delete the old index and create a new, empty index in its place. In interactive mode, you will get a warning message before this happens, but in batch mode dtSearch assumes that you really want to create a new index.

Adding Documents to an Index

To add documents to an index in a batch script, you need to select an index to add to, list the directories to index, and list the file name filters you want to use. Use either the name (SET INDEXNAME) or the path (SET INDEXPATH) to specify the index. Use SET DOCPATH = to tell dtSearch the directory in which the documents are located. Add a <+> to the end of a directory name to index subdirectories of a directory:

SET DOCPATH = C:\DOCS<+>

Use quotation marks around directory names that include spaces. If the list takes more than one line, end each line except for the last with a backslash ("\"). To specify file name filters, use SET FILTER =. Use SET EXCLUDE = to specify files that you do not want to index. After you have provided the information dtSearch needs to index your documents, use ADD to add the documents to the index. Example:

```
DTSEARCH SCRIPT
SET INDEXNAME = "MyIndex"
SET DOCPATH = C:\DOCS C:\DOCS2 "C:\Long Directory Name<+>"
SET FILTER = *.DOC *.TXT *.WPF
SET EXCLUDE = A*.DOC
ADD
```

Compressing an Index

To compress an index, use SET INDEXNAME or SET INDEXPATH to specify the index to compress and then use the COMPRESS command. Example:

```
DTSEARCH SCRIPT
SET INDEXNAME = "MyIndex"
COMPRESS
```

Batch Indexing and Searching

dtSearch batch scripts allow you to automate indexing and searching tasks.   A script is a text file that lists a series of indexing commands that dtSearch will carry out.

Automatic Generation of Batch Scripts

The easiest way to create a script is to use **Create Batch Script** in the Index menu.   When you select this option, a dialog box will appear that is almost identical to the Update Index dialog box.   Fill out the box exactly as you would if you were updating an index, then click OK to create the script.

Running a Batch Script

Interactively: To execute a script from within dtSearch, choose **Execute Batch Script** in the Index menu.

Windows 3.x: In Program Manager, choose Run from the File menu, type C:\DTSEARCH\DTSWIN /SCRIPT [scriptName], and click OK.

Windows 95:   Choose Run from the Start menu, type DTSWIN /SCRIPT [scriptName], and click OK.   In command-line mode, you can use the START command in a DOS batch file, like this:

```
START /W DTSWIN /SCRIPT [scriptName]
```

dtSearch will record any errors encountered during the job in a file with the same name as the script but with the extension LOG.   Errors and informational messages will not be displayed on the screen, since dtSearch assumes that batch jobs will run without someone watching them.   After a batch indexing job, you should check the log file to see if any errors occurred.

Batch Indexing and Searching

Batch Language Reference

Words in capital letters are keywords.   Items in braces -- { and } -- are optional.   The > filename parameter means that output from the command can be redirected to a file.   Use >> instead of > if you want the output to be appended to an existing file.

If a filename or directory name contains a space, use quotation marks around the name.   Example:

```
SET DOCPATH = "C:\1994 Documents"
```

If a command is too long to fit on one line, use the backslash ("\") to continue the command on the next line.   Example:

```
SET REQUEST = apple or orange or grape or \
   pear or banana
```

ADD
Begin adding documents to an index.   Use SET INDEXNAME or SET INDEXPATH to specify the index, SET FILTER to specify filename filters, and SET DOCPATH to specify directories.

COMPRESS
Compress an index.   Use SET INDEXNAME or SET INDEXPATH to specify the index to compress.

CREATE {CASE} {ACCENT}
Create a new index.   Use SET INDEXNAME to specify the name of the index and SET INDEXPATH to specify the directory.   Options:

CASE   Create a case-sensitive index.

ACCENT   Create an accent-sensitive index.

LIST FILES {MATCH "filter"} > filename
List files in an index.   Use SET INDEXNAME or SET INDEXPATH to specify the index.   The optional filter, if present, will select the files to be listed.

LIST INFO > filename
List information about an index (the information that is displayed in the Index Information dialog box). Use SET INDEXNAME or SET INDEXPATH to specify the index.

LIST WORDS {COUNT} {MATCH "filter"} > filename
List words in an index.   Use SET INDEXNAME or SET INDEXPATH to specify the index.   The optional filter, if present, will select the words to be listed.   The filter can include wildcards (* and ?), fuzzy (%) and phonic searching (#), and stemming (~).

LIST TEXT filename1 > filename2
List the text of filename1

OPTION option-text
Change a user option in the dtSearch OPTIONS.DAT file.   The text following the word OPTION is treated as if it were a line in OPTIONS.DAT.

QUIT
Exit dtSearch.

REMOVEDELETED
Check each file in an index to see if it has been deleted and, if it has, remove it from the index.   Use SET INDEXNAME or SET INDEXPATH to specify the index.

REPORT ResultsName ReportName WordsOfContext
Create a search report in ReportName based on search results in ResultsName.

SEARCH {COUNT} {HITS} > filename
Perform a search.   Use SET INDEXES to specify the indexes to search, SET REQUEST to specify the search request, SET RESULTS to specify the name of the search results file, and SET SORT to specify how the results should be sorted.   SEARCH > filename will list the full name of each retrieved document. COUNT adds the number of hits in each file to the listing.   HITS adds a list of the word offsets of the hits in each file to the listing.

SET DOCPATH = path1 path2 ...
Specify the directories to be indexed.   If the list extends over more than one line, each line (except the last) should end in a backslash ("\") so dtSearch knows that the list continues on the next line.   Add <+> to a directory name to index the directory and any subdirectories.   Directory names with spaces in them must be quoted.

SET EXCLUDE = filename filters
Specify files not to be indexed.   More than one filter can be included (use spaces to separate them).

SET FILTER = filename filters
Specify files to be indexed.   More than one filter can be included (use spaces to separate them).

SET HEADER = YES|NO
A search listing will contain a header with your search request, the date of the search, and the number of files retrieved.   To suppress this header, use SET HEADER = NO.

SET INDEXES = list of indexes
Select indexes to be searched.   List the names (in quotation marks) of each index that you want to search.

SET INDEXNAME = name
Specify the name of an index to update or create.

SET INDEXPATH = path
Specify the directory of an index to update or create.

SET LIB = filename
Select a new index library.   If the library does not exist, it will be created.   SET LIB should precede any references to indexes contained the library to be selected.

SET REQUEST = search request
Enter a search request for a SEARCH command.

SET RESULTS = filename

Specify the name of the Search Results file to create in a search.   If you view this file in dtSearch using **View File** in the File menu, it will act exactly like the Search Results files created in an interactive search.

SET SORT = HITS|DATE|NAME|NONE ASCENDING|DESCENDING

Specify how you want the results of a search to be sorted.

WINDOW MAX|MIN|NORMAL

Change the appearance of the dtSearch main window.

Batch Listing Commands

The LIST command will list words in an index, files in an index, information about an index, or the text of a file.   By default, the output of a LIST command will go into your LOG file. To redirect it to another file, use the same > and >> symbols used in the SEARCH command, like this:

```
LIST WORDS > words.lst
```

LIST WORDS

LIST WORDS will create a list of all of the words in the currently selected index.   Example:

```
SET INDEXNAME = MyIndex
LIST WORDS > words.lst
```

LIST WORDS COUNT will list, next to each word, the number of times it occurs in the index.

To selectively list words in an index, use a filter, like this:

```
LIST WORDS MATCH "APPL*" > words.lst
```

This will list all words that match *appl**.   The filter can contain wildcards like * and ?, the phonic search character #, the stemming character ~, and the fuzzy search character %.   For example,

```
LIST WORDS MATCH "AP%%PLICATION"
LIST WORDS MATCH "#APPLICATION"
```

The first command would list all words that match a fuzzy search for *application*.   The second command would list all words that sound like *application*.

LIST FILES

LIST FILES works exactly like LIST WORDS except that it lists filenames instead of words.   The > and >> symbols work the same way, and you can use the same types of filters.   For example,

```
SET INDEXNAME = MyIndex
LIST FILES MATCH "SMITH*.DOC" > smith.lst
```

This would list all files in MyIndex whose name matches SMITH*.DOC.

LIST INFO

LIST INFO lists the information about an index that appears in the Index Information dialog box.   This includes the date the index was created and last modified, the number of documents in the index, the size of the index, etc.   Example:

```
SET INDEXNAME = MyIndex
LIST INFO
```

LIST TEXT

LIST TEXT provides an easy way to extract the text from a file that is in a format that dtSearch recognizes.   Example:

```
LIST TEXT smith.doc > smith.txt
```

If SMITH.DOC is a WordPerfect file, SMITH.TXT would contain the text of that file as it would appear in dtSearch.   Another example:

```
SET INDEXES = MyIndex
SET REQUEST = apple w/5 pear
SET RESULTS = results.dat
SEARCH
REPORT results.dat report.dat 30
LIST TEXT report.dat > report.txt
```

This script would search MyIndex for *apple w/5 pear*, write the results to RESULTS.DAT, prepare a search report based on RESULTS.DAT, and then extract the text of the search report to REPORT.TXT.

Batch Scripts

A script file is any text file that dtSearch can read.   You can create a script using the **Create Batch Script** option in the Index menu, or you can create a script by hand.   Before you try to write your own scripts, you may want to create an example using the Create Batch Script function so you can see what a batch script looks like.   After you create the script, choose **View File** in the File menu to read it.   Using **Edit File**, you can easily edit the script to create a new script.

You can also create a script outside of dtSearch.   A script can be a plain text file or it can be in one of the word processor formats that dtSearch recognizes, such as WordPerfect.   The first line of a script must consist of the words DTSEARCH SCRIPT.

A script consists of a series of SET statements that specify the documents to be indexed and the index to use, and commands such as ADD and COMPRESS.   dtSearch ignores any line beginning with a *, so you can include comments in a script.   Here is an example:

```
DTSEARCH SCRIPT
* Create the index
SET INDEXNAME = "MyIndex"
CREATE
* Add documents from C:\DOCS
SET DOCPATH = C:\DOCS
SET FILTER = *.DOC *.TXT *.WPF
ADD
* Compress the index
COMPRESS
```

This script would: (1) create an index called "MyIndex," (2) add to the index documents in the directory C:\DOCS, and (3) compress the index. (You can combine several different indexing and searching operations in a single script.)   To run this script, put it in the file UPDATE.JOB in the dtSearch directory and execute the command:

```
DTSEARCH /SCRIPT UPDATE.JOB
```

from the dtSearch directory.   dtSearch will then execute the script, logging any error messages to the file UPDATE.LOG.

Batch Searching Commands

Searching

The following is an example of a batch script that would perform an indexed search:

```
DTSEARCH SCRIPT
SET REQUEST = apple and pear and banana
SET INDEXES = "Index1" "Index2" "Index3"
SET SORT = HITS
SET RESULTS = c:\dtsearch\search1.dat
SEARCH
```

SET REQUEST tells dtSearch what you want to search for.

SET INDEXES tells dtSearch which indexes you want to search.

SET SORT specifies how the search results will be sorted. Options are:

HITS Sort by hits.
NAME Sort by name.
DATE Sort by date.
NONE Unsorted.

You can use ASCENDING or DESCENDING after the type of search to specify an ascending or descending sort. For example:

```
SET SORT = NAME DESCENDING
```

By default, name sorts are ascending and hit and date sorts are descending.

Search Results

After your search, you will want to be able to see the results.   To store the results of your search in a dtSearch Search Results file, use SET RESULTS to tell dtSearch the name of the file to create:

```
SET RESULTS = c:\mysearch.dat
```

This tells dtSearch to write the results of the search to C:\MYSEARCH.DAT.   After your search, you can choose **View File** in the File menu to view C:\MYSEARCH.DAT.   Your search results will appear in the same way as other search results in dtSearch.

A second way to view your search results is to tell dtSearch to create a listing file, which is simply a list of the retrieved documents in ANSI text format.   If you view this file in dtSearch, it will look like a list of names and will not act like a Search Results window.   Listing files are useful if you want to use another program to process the results of your search.   To create a listing file, add a > and the name of the file to create after the SEARCH command, like this:

```
SEARCH > listing.txt
```

To append your search results to an existing file, use >> instead of >, like this:

```
SEARCH >> listing.txt
```

The listing file will contain a header with your search request, the date of the search, and the number of files retrieved.   To suppress this header, add SET HEADER = NO to your script before the SEARCH command.

Ordinarily, a listing file will contain only the names of the retrieved files.   If you want the number of hits in each file to be included in the file, add COUNT after the SEARCH command, like this:

```
SEARCH COUNT > listing.txt
```

Search Reports

The REPORT command will create a search report based on search results.   Example:

```
REPORT results.dat report.dat 30
```

This would create a search report in REPORT.DAT based on the search results in RESULTS.DAT.   The report would contain 30 words of context surrounding each hit retrieved.

Indexes

Button Bar Options

The button bar is a row of icons across the top of the dtSearch main window.   Each of the icons has a brief label explaining its purpose. For more information about a button, position the mouse cursor over it. A description of the button's function will appear at the bottom of the dtSearch main window.

The button bar can be displayed as a row of icons, icons and text, or just text.   To change the button bar display, click the right mouse button anywhere in the button bar.   A menu will pop up listing the three ways to display the button bar.   To select the one you want, click on it.

The menu also lists an option to change the font used in the text labels for the button bar.

Using the Search Dialog Box

First, tell dtSearch where you want to search

1. For index searches, click the **Select** button next to **Indexes to search** to select the indexes you want to search.

2. For unindexed searches, click the **Select** button next to **Directories to search** to select the directories you want to search.

3. To limit your search by filename, file size, or file date, click the **Filters** button to access the Search Filters dialog box.

Next, tell dtSearch what you want to find

1. Enter a search request in the space provided.   There are two basic types of search requests:

    A structured search request consists of a group of words, phrases or macros linked by search connectors such as AND and OR to precisely indicate the relationship between them.

    A natural language search request consists of an unstructured natural language or "plain English" query.   In a natural language search request, words such as AND and OR are disregarded.

2. To enter a structured search request, enter words or phrases linked by connectors.

    To have dtSearch enter the connectors for you, click the **Compose Request** button.   In the Compose Search Request dialog box, you can enter the words or phrases you want to search for and click on the connector buttons to describe how the words are related.

    From the Search dialog box, select the **Search features** you want to use in your search.

    Stemming searches other grammatical forms of the words in your search request.   For example, with stemming enabled a search for *apply* would also find *applies*.

    Phonic searching finds words that sound similar to words in your request, like *Smith* and *Smythe*.

    Fuzzy searching sifts through scanning and typographical errors.   Fuzziness adjusts from 1 to 10 depending on the degree of misspellings.   (Try starting with 3.)

    Use thesaurus tells dtSearch to use a thesaurus that you have created to find synonyms of words in your search request.

3. To enter a natural language search, click **Natural language** under **Search features** and type in a plain English search request: *Get me Sam's memo on the 1994 takeover of CorpX.*   Natural language searching is only available in index searches.

    A natural language search can rank retrieved files from most to least relevant according to the density and rarity of matching words in your documents.   For example, in the above search request, if *1994* appears in thousands of your files, and *CorpX* only appears in two, then the *CorpX* files would receive a higher relative ranking.

4. Click OK to start the search.

Search Tools

Use the buttons at the top of the search dialog box to select any of the search tools below.

Word List
A scrolling list of the words in the index you have selected.   Next to each word is a number, which is the number of times the word occurs in the index. As you type in a search request, the list will scroll to the word you are typing.

If you have selected more than one index to be searched, the highlighted index will be the one listed.   To change the index displayed in the Word List click on the index you want to see in the **Indexes to Search** box.

Hints
The basics of creating search requests.

History
A list of your most recent search requests.   Next to each request is a pair of numbers, like "35/283," which tells you the number of files and hits that the request found.

Macros
A list of macros you have defined.

To insert a word, previous search request or macro into a search request, click on the item to highlight it, then click the **Insert** button.

Composing a Search Request

Click the **Compose Request** button in a search dialog box to access the **Compose Search Request** dialog box.   This dialog box helps you to compose a search request.

A search request consists of a group of words, phrases, or macros linked by connectors such as "AND" and "OR" that indicate the relationship between them.   To enter a search request in the Compose Search Request dialog box, enter the words you want to search for and press one of the connector buttons (and, or, not, etc.) to describe how the words are related.

Click the **Lookup** button to see fuzzy, phonic, stemming and wildcard variations on a word.   Highlight and click Insert to put the word in your search request.

Click the **Numeric range search** button to add a numeric range search term to your search request.

Compressing an Index

When you reindex a document that you have previously indexed, dtSearch marks the information about the old version of the document as "obsolete" but does not remove it from the index.   Compressing an index can take a while (dtSearch completely reconstructs the index) but it makes the index smaller and makes searches faster.

There are two ways to compress an index.

1. Check the Compress Index option in the Update Index dialog box.

2. Choose Compress Index in the Index menu, select the index you want to compress, and click OK.

Contents

Contents

[Indexing](#)

[Searching](#)

[Viewing Search Results, Retrieved Documents, and Images](#)

[Setting up dtSearch](#)

[Advanced Features](#)

Contents

Indexing

Searching

Viewing Search Results, Retrieved Documents, and Images

Setting up dtSearch

Advanced Features

Contents

Indexing

Searching

Viewing Search Results, Retrieved Documents, and Images

Setting up dtSearch

Advanced Features

Contents

Contents

Indexing

Searching

Viewing Search Results, Retrieved Documents, and Images

Setting up dtSearch

Advanced Features

Contents

Copying Text

To copy text from a retrieved document

1. Mark the text to copy.

2. Click the **Copy** button in the button bar, press Ctrl-C, or choose **Copy** from the Edit menu.   The Copy Marked Text dialog box will appear.

3. Choose whether you want to copy the text to the Windows clipboard or directly to a file.   Copying text to a file lets you append text to an existing file.   If you append text to an existing file, it must be a plain text file, not a word processing document.

4. Enter any annotations you want to add to the copied text.

    Include filename.   Check this box if you want the name of the file the text is copied from to appear at the top of the text.

    Include page and line number.   Check this box to add the page and line where the text appeared in the original document.

    Annotation.   In this box, you can enter a brief note to be copied with the text.

5. Click OK.

Copying a File

To copy a file

Choose **Copy File** in the Edit menu to copy the file you are viewing. In the Copy File dialog box, enter the name of the copy you want to create and click OK.   If the filename you enter already exists, dtSearch will ask if you want to overwrite the existing file.   To suppress this message, clear the **Confirm Overwrites** check box.   *Tip:* **Copy File** will unzip a file if it is in a ZIP archive.

To copy search results or a search report

You can use **Copy File** to save search results or search reports for later viewing.   To see the copied search results or search report, use **View File** in the File menu.

To copy multiple files in a search results list

To copy multiple files in a Search Results list, select the files you want to copy (press SPACEBAR to select a file) and then choose **Copy File.**   Since you will be copying more than one file, enter a directory name instead of a filename in the dialog box.

Copying an Index

Choose **Copy Index** in the Index menu to copy an index (or click the **Copy** button in the <u>Index Manager</u>), then select the index to copy.   You must enter a name and directory to copy the index to that is different from the original index.   The directory that you copy an index to should not already exist.

Creating an Index

Before you create an index

File Formats   dtSearch automatically recognizes major word processor files, DBF files, ANSI files, and ZIP files.   For a complete list of the file formats that dtSearch supports, see **Supported File Types** in the on-line help.

EXE and other Program Files   By default, dtSearch will index and search everything in a selected directory.   To avoid indexing files that are not documents, you can use filename filters in the Update Index dialog box.

Other indexing options   Choose **Index Options** in the Options menu if you want to change the default setting for handling of hyphens, numbers, etc. in your indexes.   If you change these options after building an index, you will need to rebuild the index.

Create Index (Basic)

Choose **Create Index (Basic)** in the Index menu to create a new index. Enter a name for the index (up to 16 letters or numbers) and click OK to create the index.

If you "create" an index that already exists, you will clear the previously existing index (the documents will, of course, be unaffected).   dtSearch will ask you to confirm that you really want to do this.

Create Index (Advanced)

Choose **Create Index (Advanced)** in the Index menu if you want to use the advanced options listed below:

Path Enter the directory where dtSearch will store the index.   If you leave this blank, dtSearch will create the index in a subdirectory of the dtSearch directory.

Case sensitive index   Check this box if you want dtSearch to take capitalization into account in indexing words.   In a case sensitive index, *APPLE*, *Apple*, and *apple* would be three different words. For most users, this would be a very bad idea since most users would like to retrieve a document containing *Apple* in a search for *apple*.   It will also greatly increase the relative size of an index.   This option is useful mainly for programmers who want to index case sensitive source code.

Accent sensitive index   Check this box if you want dtSearch to take accents into account in indexing words.   Again, for most users this is not a good idea, since this option increases the chance that you will miss retrieving a document because an accent was omitted in one letter.   This option will also increase the size of an index.

Customizing the Display

Search Results Format

Choose **Search Results Format** in the Options menu to modify the format dtSearch uses to display search results.   In the Search Results Format dialog box, select the items you want displayed and click OK.

Document Font

Click **Font** in the button bar to change the font used to display documents or search results.

Tab Width

Word processor documents define tab settings internally, but in ANSI or ASCII text files a tab size must be assumed.   To change the tab setting for these files, choose **Preferences** in the Options menu and enter the width (in characters) under **Tab size.**

Button Bar

To change the button bar appearance, click the right mouse button anywhere in the button bar.   A menu will pop up listing the three ways to display the button bar.   The button bar can be displayed as a row of icons, icons and text, or just text.   To select the one you want, click on it.   The menu also lists an option to change the button bar text font.

DDE API

dtSearch responds to a DDE-based API based on batch scripts.   To enable the dtSearch DDE interface, run DTSWIN /DDE.   To execute a script, connect to the service "dtSearch" with the topic "script," then send a DDE Execute message with the name of the script to execute.

Deleting an Index

Deleting an index does not affect the original documents.   It just removes the index from your system. To delete an index, choose **Delete Index** from the Index menu (or click the **Delete** button in the <u>Index Manager</u>), select the index to delete, and click OK.   dtSearch will ask you to confirm that you really want to delete the index.

Editing a File

Choose Edit File from the File menu to access the Windows notepad editor.   The Edit File dialog box will appear.   Enter the name of the file that you want to edit and choose Open.   The Windows Notepad editor will pop up over dtSearch with the file you specified.   See your Windows documentation for information about using Notepad.

Notepad cannot edit files in word processor formats such as WordPerfect or Ami Pro.   It will only edit plain text files.   The editor provides limited editing capabilities, intended mainly for jotting down notes to accompany text clipped from documents.   (To clip text from a document, use the Mark and Copy functions, which are explained in the chapter on "Viewing Search Results.")

Environment Variables

If you include an environment variable in a batch script command, dtSearch will substitute it with the value of the variable in the current environment.   For example, if the variable %DTSDIR% has the value C:\DTSEARCH, then a script could include the line:

```
SEARCH > %DTSDIR%\RESULTS.LST
```

dtSearch would interpret this to mean:

```
SEARCH > C:\DTSEARCH\RESULTS.LST
```

Error Messages

*Note: if you want dtSearch to "beep" every time it displays an error message, check **Beep on messages** under **Preferences** in the Options menu.*

"Application table is full"   You can define up to 7 applications in dtSearch.   If you attempt to add more, this message will appear.

"Cannot copy an index onto itself"   You have asked dtSearch to copy an index, and either the directory to copy it to is the same as the directory the index currently occupies, or the name of the index to copy is the same as the new name for the copied index.   When you copy an index, the name and directory for the new index must be different from the name and directory of the existing index.

"Cannot overwrite protected file type"   This message will appear if you attempt to use Edit File to edit a document that dtSearch recognizes as a word processing document.   To protect against accidental destruction of original documents, dtSearch considers all files with special formatting codes that it recognizes (such as WordPerfect files) to be "protected."   dtSearch will not overwrite such files.

"Directory xxx already contains index yyy"   A directory can only hold one index.   If you try to use Create Index to create a second index using the same directory name, this message will appear.

"Error launching application"   dtSearch tried to execute a program and the program returned an error code.   This means that something went wrong in the launched program.

"Error(s) in search request -- CONTAINS without a TO"   A field search has the form *beginning TO end CONTAINS contents*.   The *beginning TO end* part defines where dtSearch should search, and the *contents* part is what dtSearch should look for.   Without the *beginning TO end* part, the request cannot be interpreted.

"Error(s) in search request -- Illegal use of the TO connector"   The TO connector is used to define a field for a field search. The format for a field definition is *beginning TO end*, where *beginning* marks the beginning of the field and *end* marks the end.   Both parts of the field definition may consist only of words, phrases, and the OR connector -- other connectors are not permitted in field definitions. Thus, *(apple or pear) to (grape or fish)* is legal; *(apple and pear) to (grape or fish)* is not.

"Error(s) in search request -- Illegal use of W/N connector"   Some search requests using the W/N connector are ambiguous or impossible to evaluate.   See "W/N Connector."

"Error(s) in search request -- Missing words"   This message appears when you have left one or more words out of a search request.   Example: *apple and pear and*

"Error(s) in search request -- Parentheses do not balance"   You have left out an opening or closing parenthesis in a search request.

"Error reading WordPerfect file"   When dtSearch reads a WordPerfect document, it checks for errors in the file and displays this message if it finds any.   You can generally fix a corrupted WordPerfect document by getting the most recent release of WordPerfect and retrieving the document into WordPerfect, scrolling to the end of the file, and saving it.

"Error running script"   This message will appear after you have executed a batch script which caused an error to occur.   Check the log file created during execution of the script to see what happened.   (The log file will have the same name as the script file but it will end in .LOG)

"File xxx is not a script file"   You have attempted to execute a file using the /SCRIPT option, and dtSearch determined that the file was not a batch script.   A dtSearch batch script must begin with the

words `dtSearch Script.`

"Index already exists"   If you try to create an index that already exists in a different directory, this message will appear.   If you are trying to create a new index, try using a different name for the new index, or (if you do not want the old index) deleting the old index.   If you are trying to re-create an existing index, make sure the directory you have selected is the same one that you used to create the original index.

"Index has version 1.x format," "Index has version 2.x format," "Index has version 3.x format"
You have asked dtSearch to search or update an index created by an old version of dtSearch.   dtSearch version 4 cannot use these indexes.

"Index library is full"   An index library can hold a maximum of 100 indexes.   If you try to create, copy, or recognize an index when you have reached this limit, you will get this error message.   Delete any indexes that you do not need, or create a new index library to use for the new index.

"Index too large to fit on drive"   You have asked dtSearch to copy an index and there is not enough space for the index on the drive where it is to be copied.

"Macro is circular"   A circular macro is one that references itself, directly or through other macros. Such a macro would expand infinitely if used in a search request.

"Macro not found"   dtSearch assumes that any search term beginning with the @ character is a macro name.   If it does not find a macro with this name, this message appears.

"Missing file"   One of the files that dtSearch needs is missing.   Try re-installing dtSearch to restore the missing file.

"No application to launch"   You have tried to launch an application with a file, and dtSearch was unable to match the file with an application.   When you launch an application with a file, dtSearch checks the type and name of the file you launched against the applications you set up in your applications table. If dtSearch cannot find a match, this message will appear.   Choose Applications in the Options menu to make sure the application was set up correctly.

"No files to index"   This means that you tried to add some documents to an index, but either you did not specify any files to be added or all of the files were already in the index.

"Not enough space to compress or update index"   dtSearch did not find enough disk space available to update or compress an index.   Try removing some files from the disk or moving the index to a disk drive with more space available. Also, try indexing smaller groups of files.

"Not found"   You have used Find Text to search for text in a document, and the text was not found.

"To create an index, you must enter a name for the index"   You have used Create Index to create a new index and did not enter a name for the index.

"Too many words retrieved in index"   Your search request contained extremely broad search terms, such as *(r* w/5 s*) and t**, that matched several hundred words in an index. Try using more specific search terms, turning off phonic searching if it is on, or reducing the level of fuzziness if fuzzy searching is enabled.

"Unable to access drive"   You have asked dtSearch to access a disk drive and there is some problem with the disk drive.   If you were trying to access a floppy disk drive, make sure that there is a disk in the drive, that the disk is correctly inserted, and that the drive latch is closed.   If you were trying to access a hard disk drive, the drive you requested may not exist or there may be a problem with your hard disk.

"Unable to access ...", "Unable to read ..."   dtSearch cannot find the file or directory you specified. Check that the file or directory name is correct and that you are using the right disk or disk drive.

"Unable to create ...", "Unable to write ...", "Unable to save ..."   dtSearch cannot write to or create the file or directory you specified.   Check that the file or directory name is correct and that you are using the right disk or disk drive.   These messages may also appear when a disk is full.

"Unable to access index"   dtSearch was unable to find the index where it was supposed to be.

"Unable to use directory.   Use the /dir command-line option to specify a different directory."   On a network, each user must have a private directory to use for temporary files.   You have asked dtSearch to use a directory that another user is using.   See "Using dtSearch on a Network."

"You must create an index before using this command"   You have tried to execute a command that cannot be used unless you already have an index to work with.   Choose Create Index (Basic) in the Index menu to create an index.

File Types

dtSearch recognizes most file formats automatically.   If you are indexing only files such as word processor documents that dtSearch supports and can automatically recognize, you can disregard this section.

If you are indexing other types of files, dtSearch provides a way to specify, using filename filters, how you want dtSearch to process the files.   For each filter, you can specify a "File Type" that tells dtSearch how you want the file to be handled.   The file types fall into two groups:

Older word processors   If you are indexing files that use older file formats that cannot be recognized automatically, use these file types to tell dtSearch how to recognize the files.   These formats are: WordPerfect 4.2, WordStar versions before 4.0, and XyWrite.

ASCII and ANSI   Windows programs use the "ANSI" character set instead of the "ASCII" character set used in DOS.   The two character sets are mostly the same but the ANSI character set handles accented letters differently from the ASCII character set.   This is not a problem for word processor files because dtSearch can detect the file format and handle the file correctly.   However, for plain text files created with a Windows program such as Notepad, dtSearch has no way to know whether the file is ANSI or ASCII.   By default, dtSearch assumes that text files use the ANSI character set.   If you have plain text files created using DOS that include accented letters, use the ASCII file type to tell dtSearch how to recognize these files.

Before using the file type information, dtSearch will try to detect the format itself.   Therefore, no matter what file type specifications you enter, dtSearch will recognize formats such as WordPerfect 5.1 that it can detect automatically.

dtSearch checks the filename filters in the order that you created them and uses the first one that matches.

To set up a file type specification

1.   Choose File Types in the Options menu to access the File Types dialog box.

2.   To set up a new file type, select "<add new item>" and choose the Edit button.
     To edit a file type that you set up previously, move the cursor to it and choose the Edit button.

3.   In the Setup File Type dialog box, enter the following information:

Type   Choose the file type you want to set from the list in the dialog box.

Filename Filter: Enter a filename filter that will identify files having that format.   This filter is similar to usual filename filters (*.DOC, *.*, etc.), except that it can include drive and directory specifications, such as: *\WP42*, which matches anything in a directory named WP42, or C:*.DOC, which matches any filename ending in ".DOC" on drive C:.   You must enter a filename filter for each file type you want to set up, even if the filter is just "*"

Filename Filters

A filename filter is a pattern, like *.*, *.DOC, or FILE??10.DAT, that you can use to tell dtSearch which files you want it to search or index in a directory.   For example, if you have a directory of word processing files, and want to search all of the files beginning with SMITH and having the extension DOC, you would use the filter SMITH*.DOC.

A "?" matches any single character, so SMITH?.DOC would match SMITH1.DOC, but not SMITH123.DOC.   An asterisk ("*") in a file name filter matches any number of characters, so SMITH*.DOC would match SMITH001.DOC, SMITHAAA.DOC, etc.

To use more than one filter, just list the filters, separated by spaces.   Example:

        SMITH*.DOC JONES*.DOC *.DAT

If you do not specify a filename filter, dtSearch will index all of the files in the directories selected.

"Exclude" filters work the same as ordinary filters.   They just specify files that you do not want to include.

Finding Text in a Window

Click the Find button or press F2 to search for any text in a displayed document.   dtSearch will ask you for the text to search for.   Type in the text you want to find and choose OK.   Find is like the search function in your word processor -- you can search for words, spaces, punctuation, etc.   Capitalization is ignored, and the * and ? characters are searchable like any other characters (they are not treated as wildcards).

Find starts searching after the line that the cursor is on, so you may want to move to the top of the document before beginning a search.   Find will bring you to the beginning of the line containing the text searched for.

Fuzzy Searching

Fuzzy searching will find a word even if it is misspelled.   For example, a fuzzy search for *apple* will find *appple*.   Fuzzy searching can be useful when you are searching text that may contain typographical errors, or for text that has been scanned using optical character recognition (OCR).   There are two ways to add fuzziness to your searches:

1. Check **Fuzzy searching** in the search dialog box to enable fuzzy searching for all of the words in your search request.   You can adjust the level of **fuzziness** from 1 to 10.

2. You can also add fuzziness selectively using the % character.   The number of % characters you add determines the number of differences dtSearch will ignore when searching for a word.   The position of the % characters determines how many letters at the start of the word have to match exactly. Examples:

`ba%nana`	Word must begin with *ba* and have at most one difference between it and *banana.*
`b%%anana`	Word must begin with *b* and have at most two differences between it and *banana.*

Getting Started

How to Order dtSearch

Toll-free Orders:	1-800-IT-FINDS (1-800-483-4637)
Phone:	(703) 413-3670
Fax:	(703) 413-3473
Web site:	www.dtsearch.com
Internet e-mail:	sales@dtsearch.com
CompuServe:	72607,3323
Address:	2101 Crystal Plaza Arcade, Suite 231 Arlington, Virginia 22202 United States
Technical Support:	(703) 827-0793

Distributors

United Kingdom	ElectronArt Design (+44) 181 983 8686 (voice & fax) uksales@dtsearch.com
Australia	Indigo Pacific Pty Ltd 61 2 9955 8000, Fax 61 2 9955 8511 australia@dtsearch.com
Canada	Gatierf Publications Ltd. 1-888-IT-FINDS, Fax (403) 293-6232 wright@gatpub.com

Hypertext Documents

dtSearch supports two hypertext formats: a proprietary "dtSearch Hypertext" format, described below, and HTML.   For most applications, HTML is recommended.

HTML

dtSearch supports hypertext links in HTML documents as well as basic formatting (bold, underlining, italics, indentation, centering, highlighting hits, etc.).   Hypertext links will appear in dtSearch in green and underlined.   To follow a hypertext link, double-click on it.   A link may connect to another document, an image file, or a document on the internet, in which case dtSearch will launch your web browser to access the document.

dtSearch Hypertext

A dtSearch hypertext document is a text file that begins with the words `dtSearch Hypertext` on the first line.   This text is not displayed in dtSearch and is used only to signal that the rest of the document should be interpreted as hypertext.   dtSearch can view, index, and search hypertext documents like any other type of document. Hypertext links will appear highlighted when a document is viewed. Hypertext links can point to another document or can launch another program (such as a graphics viewer).

Hypertext links have the following format:

```
Sample text. @Highlighted text@^other.doc^ More text
```

The text between the @ marks is highlighted and the text between the ^ marks is the hypertext link associated with the highlighted text.   Pressing ENTER while the cursor is on the highlighted text would tell dtSearch to display OTHER.DOC.

A hypertext link can be a simple filename, as in the above example, or it can be more complex.   More complex links contain a series of the following items:

*F name   Name of the file to view

*W n   Position cursor after the nth word in the file.

*E command   Execute the command.

*C   Close the current window.   If a link does not contain a *C, a new window will pop up over the current window when the user chooses the link.

If a link does not contain a full pathname, dtSearch will assume that the file is in the same directory as the file being viewed.   For example, if C:\JOE\DOCS\FIRST.DOC contains a link to SECOND.DOC, dtSearch will assume that SECOND.DOC is in C:\JOE\DOCS.

Examples:

```
^*F smith.doc *W 500 **^
```
View SMITH.DOC with the cursor positioned after the 500th word.   Do not close the current window when the user selects the link.

```
^*F smith.doc *C **^
```

Close the current document and open a window displaying SMITH.DOC.

```
        ^*E c:\viewer\view.exe image1.pcx **^
```

Launch VIEW.EXE and have it display IMAGE1.PCX..

Since dtSearch uses ^ and @ to recognize hypertext links, these characters cannot be used in document text.   If you need to use ^ or @ in document text, you can use the LinkMark and TargMark commands to specify different marker characters.   Example:

```
        .LinkMark #
        .TargMark ~
```

LinkMark is @, the character used to mark highlighted text.   TargMark is ^, the character used to indicate the "target" of a hypertext link.

The .title command in a document changes the way the title of the document is displayed in dtSearch. Normally, the title of the window displaying a document is the filename.   If a hypertext document contains a .title command, the supplied title will replace the filename.   Example:

```
        .title Instruction Manual -- Chapter 1
```

Imaging and Scanning

Imaging Options

Choose **Imaging Options** in the Options menu to access the Imaging Options dialog box.   The Imaging Options dialog box enables you to customize the way dtSearch associates documents with image files and to specify an image viewer other than the built-in dtSearch image viewer for viewing image files.

Image Filename Extensions   dtSearch recognizes image files by their extension.   By default, dtSearch expects files with the following extensions to be image files: BMP, EPS, GIF, IMG, JPG, PCX, TIF, TGA, WMF, WPG.   To add or remove extensions from this list, edit the list of extensions under Image Filename Extensions.

Image Directories   Directories listed under Image Directories are used to <u>link images to documents</u>. When you click the <u>Image</u> button while viewing a document, dtSearch will look for a file with the same name as the document and one of the filename extensions listed under Image Extensions.   dtSearch will check first in the directory where the document is located and then, if the image is not found, in the other directories listed under Image Directories.

Image Viewer   dtSearch has a built-in image viewer, but you can specify your own viewer in place of the dtSearch image viewer in this space.

Index Libraries

An index library is a collection of up to 100 indexes.   dtSearch uses index libraries to record the names and locations of the document indexes that you create.   When you select indexes to search, or pick an index to update, compress, etc., the list of indexes displayed comes from the current index library.   If you do not need to use more than 100 indexes, you do not need to worry about index libraries.   dtSearch starts out with a library called IXLIB.ILB that will hold any indexes that you create.

If you need to use more than 100 indexes, choose **Create Library** in the Index menu to create a new library.   A dialog box will appear allowing you to enter the name of the new library.   Enter a name for the new library and choose OK.   When you create a library, that library becomes your current library.   The new library will be empty.

To change the current library, choose **Change Library** in the Index menu and select the index library you want to use.

Index Manager

Choose **Index Manager** in the Index menu to access the Index Manager dialog box.   The Index Manager enables you to get information about each index you have created.   To see information about an index, move the cursor to it.

Buttons in the Index Manager let you <u>create</u>, <u>update</u>, <u>recognize</u>, <u>delete</u>, <u>copy</u>, <u>rename</u>, <u>verify</u>, or <u>list</u> the contents of an index.

Index Menu

Choose Index in the main menu to get into the Index menu.   This menu lists options to

<u>create</u> an <u>index</u>,
create an index using advanced options,
<u>add documents</u> to an index,
<u>compress</u> an index,
create or execute a <u>batch script</u>
<u>rename</u>, <u>copy</u>, <u>delete</u>, or <u>recognize</u> an index,
display <u>information</u> about an index,
<u>verify</u> that an index is not damaged,
<u>list the words</u> in an index, and
create or change <u>index libraries</u>.

Index Searches

The Index Search dialog box enables you to search using indexes.   Choose **Index Search** in the **Search** menu to access the Index Search dialog box.

First, tell dtSearch where you want to search

1. Click the **Select** button next to **Indexes to Search** to select indexes to search.

2. To limit your search by filename, file size, or file date, click the **Filters** button to access the Search Filters dialog box.

Next, tell dtSearch what you want to find

1. Enter a search request in the space provided.   A search request consists of a group of words, phrases or macros linked by connectors such as "AND" and "OR" that indicate the relationship between them.   To have dtSearch enter search connectors for you, click **Compose Request** to access the **Compose Search Request** dialog box.

2. Under Search Features, choose the features you want to use in this search.

   Natural language searching searches for documents that match a "plain language," or unformatted, search request.   dtSearch will rank retrieved files from most to least relevant.   Natural language searching is only available in index searches.

   Stemming searches other grammatical forms of the words in your search request.   For example, with stemming enabled, a search for *apply* would also find *applies*.

   Phonic searching searches for words that sound similar to words in your request, like *Smith* and *Smythe*.

   Fuzzy searching sifts through scanning and typographical errors.   Fuzziness adjusts from 1 to 10 depending on the degree of misspellings.   (Try starting with 3.)

   Use thesaurus tells dtSearch to use a thesaurus that you have created to find synonyms of words in your search request.

   If you want to see how stemming, phonic searching, fuzzy searching, or synonym searching will affect your search, click **Lookup**.

3. Click **OK** to start the search.

Search Tools

At the top of the dialog box is the Search Tools panel, which displays the following:

Word List	A scrolling list of the words in the index you have selected. Next to each word is a number, which is the number of times the word occurs in the index. As you type in a search request, the scrolling list will display the words in the index that are closest to the word you are typing.
	If you have selected more than one index to be searched, the highlighted

index will be the one listed.   You can select the index displayed in the Word List by clicking on the index you want to see in the Indexes to Search box.

Hints   The basics of creating search requests.

History   A list of your most recent search requests.   Next to each request is a pair of numbers, like "35/283," which tells you the number of files and hits that the request found.

Macros   A list of macros you have defined.

A row of buttons across the top of the Search Tools panel selects the search tool that is displayed.   To copy text from the Word List, History list, or Macros list, click the Insert button.

Indexing Documents

To add documents to a new index

1. After you create an index, dtSearch will ask if you want to add documents to the index.   Click **Yes**.

4. In the Update Index dialog box, click **Select** to select the directories you want to index.

5. (Optional)   Under **Filename Filters**, enter filters (e.g., *.DOC, *.TXT, etc.) to select documents to add.   If you leave this blank, dtSearch will index all of the files in the directories you selected.   Under **Exclude Filters**, enter filters (such as *.EXE) for any files you do not want to include in the index.

6. Click OK to start indexing the documents.

To update an existing index

1. Choose **Update Index** in the Index menu to access the Update Index dialog box.

2. Click on the name of the index you want to update.

3. If you want to add new directories to the index, click **Select**.

4. Check the **Index new or modified documents** box if it is not already checked.

5. If you have deleted some files that were in the index and you want to remove them from the index, check the **Remove deleted documents from index** box.

6. If you have updated the index several times, you may want to check the **Compress index after adding documents** box.   Compressing an index removes obsolete document information from an index.   It can take a while (dtSearch completely reconstructs the index) but it makes the index smaller and makes searches faster.

7. Click OK to start updating the index.

To rebuild an index from scratch

To tell dtSearch to rebuild an index from scratch, check the **Clear index before adding documents** box, check the **Index new or modified documents box**, and click OK to start rebuilding the index.

Notes

Disk Space   To build an index, dtSearch needs free space on the drive containing the index at least equal to the size of the documents to be indexed.   (This is because dtSearch creates temporary files during indexing.)   Once complete, an index is usually about a fourth the size of the original documents, though this can vary depending on the number and type of documents.   In general, the more documents in an index, the smaller the index will be as a percentage of the documents.

Indexing Documents on Floppy Disk   When an index contains documents stored on floppy disk, make sure that **Remove deleted documents from index** is not checked when you update the index.   *Tip:* You may find it useful to store the documents on each disk in a subdirectory named after the disk.   For example, if you have disks labeled SMITH and JONES, move the documents on the SMITH disk into a directory called SMITH, and move the documents on the JONES disk into a directory called JONES.

This will help you to locate the documents after a search.   You can see which disk has the documents you want by looking at the directory name in search results.

Index Options

Choose **Index Options** in the Options menu to change how dtSearch handles any of the following options.   If you change any of the options, you will need to rebuild your indexes to reflect the options that you select.

Binary Files   A binary file is a file that has a format dtSearch cannot recognize and that does not appear to be a plain text file.   See "Unrecognized File Types" for a detailed discussion of this topic.

Hyphenation   By default, dtSearch indexes a hyphenated word such as *first-class* in four different ways: *first*, *class*, *first-class*, and *firstclass*.   This covers all of the possible meanings of a hyphen character in text.   This option gives you the choice of selecting one of the possible treatments to use in all cases.

Index Numbers   If your documents contain a lot of numbers and you do not expect to want to search for them, clear this checkbox to make dtSearch exclude numbers from your index.   This will make your indexes smaller and will speed indexing.

Alphabet File   The alphabet file determines how dtSearch interprets the characters in your documents. The default alphabet file included with dtSearch is ENGLISH.ABC, which is designed to work with text, including accented letters, in the English-language version of Windows.

To modify the alphabet file (for example, to make a character such as + searchable) click the **Edit Alphabet** button to access the Edit Alphabet dialog box.

If you have a non-English version of Windows, click the **New Alphabet** button to generate a new alphabet file using the default alphabet for your version of Windows.   Enter a name for the new alphabet and click OK.

Installing dtSearch

1. Insert Disk 1 in drive A (or drive B).   (If dtSearch is on one disk, insert that disk in the drive.)

2. In Windows 95, click the **Start** button and choose **Run**.

   In Windows 3.1 or Windows NT, open **Program Manager** and choose **Run** from the File menu.

3. Type A:SETUP (or B:SETUP) and click OK.

4. Follow the directions on the screen to complete installation.

Kwic View

The "KWIC" ("Key word in context") function will display only the part of a document that contains terms matching your search request.   When you are in a search results window, move the cursor to the name of the document that you want to see and choose KWIC in the Search menu. dtSearch will ask how many words of context you want.   This is the number of words surrounding each hit that you want to see. Enter the number of lines that you want and choose OK.

A window will pop up over the search results window displaying a KWIC view of the document you selected.   You can use most of the same browsing and menu commands in a KWIC window as you can in a retrieved document.   Click the right mouse button in the KWIC view window to see the local menu listing these commands.

You can also see all of the document by moving the cursor to the name of the document and pressing ENTER or by double-clicking on the document name with the mouse.   In addition, a file displayed in KWIC mode contains highlighted line and page markers to show where in the document the text appeared.   These highlighted markers act as hypertext links into the full text of the document.

For example, suppose that you are viewing a document in KWIC mode and see a block of text with a marker "[Page 10 Line 15]."   If you move the cursor to the marker and press ENTER (or double-click with the mouse), a window displaying the full document will pop up over your KWIC view, positioned to line 15 of page 10.   To return to the KWIC view, press ESCAPE.

Launching Applications

To open a document in the application that created it, click the **Launch** button while you are viewing the document.   dtSearch will attempt to locate and execute the application that created the document.   If the application is already running, dtSearch will try to make the application load the document.   For example, if you are viewing an HTML document and click the Launch button, dtSearch will launch your web browser with the document.

Windows maintains "associations" to link documents to their applications based on the extension at the end of a filename.   For example, files ending in .TXT are generally associated with Windows notepad. For users who only use one word processor and who always use the default filename extension associated with that word processor (usually .DOC), this method of linking documents and applications is sufficient.

If this method is not sufficient, dtSearch provides a way to link documents and applications based on the file format of the document.   For example, you can tell dtSearch to always use WordPerfect to launch a WordPerfect document, regardless of the name of the document.   To set up an application this way, choose **Applications** in the Options menu.

Launching Applications

Legal Notices

Linking Images and Documents

Linking of documents and images is useful for associating one or more image files with documents that you have indexed.   It is also useful if you are using optical character recognition (OCR) to scan documents and want to be able to retrieve a document in a search and then view the image of the retrieved document.

Linking an Image to a Document

dtSearch links image files to documents by looking for files with matching names.   When you click the **Image** button while viewing a document, dtSearch attempts to find and display the image file associated with the document.   To find the image file, dtSearch looks for a file with the same name as the document and an extension that indicates it is an image file.   (These extensions are listed under **Image Extensions** in the Imaging Options dialog box.)   dtSearch will first check in the directory where the document is located and then, if the image is not found, in the other directories listed under **Image Directories** in the Imaging Options dialog box.

Linking Multiple Images to a Document

Some scanning software stores the pages of a document in multiple files with names ending in 001, 002, 003, etc.   For example, a three page file named EXAMPLE.DOC would have image files EXAMP001.TIF, EXAMP002.TIF, and EXAMP003.TIF.

If dtSearch does not find an exact match to the document filename, dtSearch will look for and collect image files containing truncated versions of the filename.   It will display the first of these files as "page 1" of the image, the second as "page 2," etc. Use the PgUp and PgDn keys to get from one page to another.

List Words in Index

Choose **List Words** in the Index menu (or click the **List Words** button in the Index Manager) to generate a list of all of the words in an index.   Select the index you want to list and choose OK.   A second dialog box will appear asking for the name of a file to store the words in.   Enter the name of a file and click OK.

Local Menu

In a document window, search results, an image file, a search report, or a KWIC view, click the right mouse button anywhere to pop up the "Local" menu for the window you are in.   The Local menu lists all of the things you can do with the window you are in.   The Local menu is different for each type of window.   For example, the Local menu for a search results window includes an option to generate a search report.   The Local menu for a retrieved document includes an option to skip to the next document retrieved in the search.

Lookup Word

Click **Lookup Word** in the search dialog box to access the Lookup Word in Index dialog box.   This dialog box lists the words in an index that match a word that you enter, using any combination of <u>wildcards</u> and <u>fuzzy</u>, <u>phonic</u>, <u>stemming</u>, or <u>thesaurus</u> search options.   To see a list of matching words:

1. Type in the word you want to look up.   The word can contain the wildcards * or ?.

2. Choose the index you want to look the word up in.

3. Choose the search options you want to use (fuzzy searching, phonic searching, stemming, or none).

4. Click **Lookup**.

To insert one of the words you found into your search request, click **Insert.**   To save the list of words in a file, click **Save List.**

Marking and Saving Text

NOT and NOT W/N

Use NOT in front of any search expression to reverse its meaning. This allows you to exclude documents from a search.   Example:

```
apple sauce and not pear
```

NOT standing alone can be the start of a search request.   For example, *not pear* would retrieve all documents that did not contain *pear*.

If NOT is not the first connector in a request, you need to use either AND or OR with NOT:

```
apple or not pear
not (apple w/5 pear)
```

The NOT W/ ("not within") operator allows you to search for a word or phrase not in association with another word or phrase.   Example:

```
apple not w/20 pear
```

Unlike the W/ operator, NOT W/ is not symmetrical.   That is, *apple not w/20 pear* is not the same as pear not w/20 apple.   In the *apple not w/20 pear* request, dtSearch searches for *apple* and excludes cases where *apple* is too close to *pear*.   In the *pear not w/20 apple* request, dtSearch searches for *pear* and excludes cases where *pear* is too close to *apple*.

Natural Language Searching

A natural language search request is any combination of words, phrases, or sentences.   After a natural language search, dtSearch can sort retrieved documents by their relevance to your search request. Weighting of retrieved documents takes into account: the number of documents each word in your search request appears in (the more documents a word appears in, the less useful it is in distinguishing relevant from irrelevant documents); the number of times each word in the request appears in the documents; and the density of hits in each document.   Noise words and search connectors like NOT and OR are ignored.

Because natural language searching uses word information stored in an index, it is impossible to perform an unindexed natural language search.   Therefore, natural language searching is not available in unindexed or combination searches.

Noise Words

A noise word is a word such as *the* or *if* that is so common that it is not useful in searches.   To save time, noise words are not indexed and are ignored in index searches.   To modify the list of words defined as noise words, edit the file DTSEARCH.NOI.   You can use the **Edit File** function (in the File Menu) to edit DTSEARCH.NOI, or you can use your word processor. If you use your word processor, make sure that you save the file as a DOS text file (with the name DTSEARCH.NOI in the directory where dtSearch is installed) so that dtSearch can read it.

The words in DTSEARCH.NOI do not have to be in any particular order, and can include wildcard characters such as * and ?.   However, noise words may not begin with wildcard characters.

Search connectors such as *and* and *or* should always be listed as noise words since it is not possible to search for them.

When you create an index, the index will store its own copy of the noise word list.   Changes you make to DTSEARCH.NOI will be reflected in future indexes you create but will not affect existing indexes.

Numeric Range Searching

A numeric range search is a search for any numbers that fall within a range.   To add a numeric range component to a search request, enter the upper and lower bounds of the search separated by ~~ like this:

```
apple w/5 12~~17
```

This request would find any document containing *apple* within 5 words of a number between *12* and *17*.

Notes

1. A numeric range search includes the upper and lower bounds (so *12* and *17* would be retrieved in the above example).

2. Numeric range searches only work with positive integers.

3. For purposes of numeric range searching, decimal points and commas are treated as spaces and minus signs are ignored.   For example, *-123,456.78* would be interpreted as: *123 456 78* (three numbers).   Using <u>alphabet customization</u>, the interpretation of punctuation characters can be changed.   For example, if you change the comma and period from **space** to **ignore**, then *123,456.78* would be interpreted as *12345678*.

OR Connector

Use the OR connector in a search request to connect two expressions, at least one of which must be found in any document retrieved.   For example, *apple pie or poached pear* would retrieve any document that contained *apple pie*, *poached pear*, or both.

Options

Other Batch Commands

QUIT terminates execution of a script and also causes dtSearch to exit.

WINDOW controls the appearance of the dtSearch main window.

WINDOW MIN reduces it to an icon.
WINDOW MAX maximizes it to fill the screen.
WINDOW NORMAL displays the window at its default size.

OPTION allows you to change a dtSearch setup option.   What follows the OPTION command is treated as if it were a line in the dtSearch OPTIONS.DAT file.

Page Setup

The Page Setup dialog enables you specify how you want documents to appear when printed.   Enter the margins either in inches or in centimeters, and enter any header or footer you want to appear on each page.

You can use the following special codes in headers and footers:

&f   Filename
&p   Page number
&d   Current date
&r   Right justify
&l   Left justify
&c   Center

Example: The header "&r&f Page &p" would put the filename and page number at the top right of each page.

Phonic Searching

Phonic searching looks for a word that sounds like the word you are searching for and begins with the same letter.   For example, a phonic search for *Smith* will also find *Smithe* and *Smythe*.

To ask dtSearch to search for a word phonically, put a # in front of the word in your search request. Examples: *#smith, #johnson*

Check **Phonic searching** in the **Search features** section of the search dialog box to enable phonic searching for all of the words in your search request.   Phonic searching is somewhat slower than other types of searching and tends to make searches over-inclusive, so it is usually better to use the # symbol to do phonic searches selectively.

Preferences

Choose Preferences in the Options menu to change the following options:

Tab size   The is the width of a tab in a file displayed in a document window. For most word processors, this will be specified in the document. However, for ASCII or ANSI text files, the value specified here will be used.

Beep on messages   Check this box if you want a "beep" every time dtSearch displays an information or error message.

User Name   Enter your name here if you will be using dtSearch on a network.   If your network supports use of an environment variable such as %USER% or %USERNAME% to identify users, you can enter that environment variable here.

Printing a File

To print a file you are viewing, click the **Print** button in the button bar.   The Print dialog box will appear. The Print dialog box enables you to specify how you want documents to appear when printed.   Enter the margins either in inches or in centimeters, and enter any header or footer you want to appear on each page.   You can use the following special codes in headers and footers:

    &f   Filename
    &p   Page number
    &d   Current date
    &r   Right justify
    &l   Left justify
    &c   Center

Example: The header "&r&f Page &p" would put the filename and page number at the top right of each page.

To select a different printer, click the **Printer** button.

Quick Start

dtSearch can search megabytes of text in a second.   It does this by building an index that stores the location of words in your files.   An index does not store the text of documents, so to view retrieved documents you will still need to keep the documents accessible.

Indexing Documents

1. Choose **Create Index (Basic)** from the Index menu.

2. Enter a name for the index and click OK.

3. dtSearch will ask if you want to add documents to the index.   Click **Yes** and the Update Index dialog box will appear.

4. In the Update Index dialog box, click **Select** to select the directories you want to index.   When you finish selecting directories, click OK.

5. Click OK to begin adding documents to your index.

Note 1: **Disk Space**.   To build the index, dtSearch needs free space on the drive containing the index at least equal to the size of the documents to be indexed.   (This is because dtSearch creates temporary files during indexing.)   Once complete, an index is usually about a fourth the size of the original documents, though this can vary depending on the number and type of documents.   In general, the more documents in an index, the smaller the index will be as a percentage of the documents.

Note 2: **File Formats.**   dtSearch automatically recognizes major word processor files, database files (32-bit version through ODBC), ANSI text files, ZIP files, and HTML files.   For a complete list of the file formats that dtSearch supports, see **Supported File Types** in the on-line help.

Note 3: **EXE and other Program Files.**   By default, dtSearch will index and search everything in a selected directory.   To avoid indexing files that are not documents, use filename filters in the Update Index dialog box.

Searching using the Index

1. Click **Search** on the button bar (or choose **Index Search** from the Search menu).   The Index Search dialog box will appear.   At the top of the dialog box is a box that can display any of the following search tools:

Word List	A scrolling list of the words in the index you have selected. Next to each word is a number, which is the number of times the word occurs in the index. As you type in a search request, the list will scroll to the word you are typing.
	If you have selected more than one index to be searched, the highlighted index will be the one listed. You can change the index displayed in the Word List by clicking on the index you want to see in the Indexes to Search box.
Hints	The basics of creating search requests.

History	A list of your most recent search requests. Next to each request is a pair of numbers, like "35/283," which tells you the number of files and hits that the request found.
Macros	A list of the <u>macros</u> you have created.

To select one of the tools, use the buttons at the top of the Index Search dialog box.   To insert a word, previous search request or macro into a search request, click on the item to highlight it and click the **Insert** button.

2. Click the **Select** button to select indexes to search.

3. Enter a <u>search request</u> in the space provided.   There are two basic types of search requests:

   A <u>natural language</u> search request consists of an unstructured natural language or "plain English" query.   In a natural language search request, words such as AND and OR are disregarded.   A natural language search can rank retrieved files from most to least relevant according to the density and rarity of matching words in your documents.

   A <u>structured</u> search request consists of a group of words, phrases or macros linked by search connectors such as AND and OR to precisely indicate the relationship between them.

   To enter a <u>natural language</u> search request, click **<u>Natural language</u>** under **Search features** and type in a plain English search request: *Get me Sam's memo on the 1994 takeover of CorpX.*

   To enter a <u>structured</u> search request, enter words or phrases linked by connectors.

AND	both words must be present
OR	at least one of the words must be present
NOT	the next word must not be present (use with AND or OR)
W/5	the words must occur not more than 5 words apart (you can also use W/10, W/20, etc.)
*	matches any number of letters
?	matches any single letter

   Examples:
   ```
   apple and pear
   *apple w/6 pear
   apple sauce w/27 grape ju?ce
   apple and (pear w/5 bana*a)
   apple and not pear
   ```

   To have dtSearch enter the connectors for you, click the **<u>Compose Request</u>** button to access the Compose Search Request dialog box.   In the Compose Search Request dialog box, you can enter the words or phrases you want to search for and click on connector buttons to describe how the words are related.

4. Select **Search Features** to use in the search.

   <u>Stemming</u> searches other grammatical forms of the words in your search request.   For example, with stemming enabled a search for *apply* would also find *applies*.

   <u>Phonic searching</u> finds words that sound similar to words in your request, like *Smith* and *Smythe*.

Fuzzy searching sifts through scanning and typographical errors.   Fuzziness adjusts from 1 to 10 depending on the degree of misspellings.   (Try starting with 3.)

To see how stemming, phonic searching, fuzzy searching or wildcards will affect your search, click the **Lookup** button.

5. Click OK to start the search.

Viewing Search Results

1. After a search is complete, a dialog box will appear listing options for sorting search results (file name, file date, or number of hits).   Pick one of the options, click OK and a <u>Search Results</u> window will appear.

2. To select a document to view from the Search Results window, double-click on it or move the cursor to it and press ENTER.   The document will appear in a new window overlaying Search Results.

3. To jump to the next hit in a document window, click **Next Hit** on the button bar or press SPACEBAR.

4. To search for text in a retrieved document, click the **Find Text** button in the button bar, enter the text you want to find, and click OK.

5. To switch between a document window and search results, press TAB.   To close a document window, press ESCAPE.

6. For a <u>menu</u> of all options you can use in a document window or Search Results, click the right mouse button anywhere in the window.

7. Click the **Launch** button to open a document in the application associated with it.   For example, an HTML document would be launched in your web browser.   See "<u>Launching Applications</u>" in the manual or on-line help for more information on associating documents and applications.

8. Click the **Image** button to view an image associated with a document file.   See "<u>Linking Images and Documents</u>" in the manual or on-line help for more information on associating documents and images.

9. Hypertext links in HTML documents will appear in dtSearch in green and underlined.   To follow a hypertext link, double-click on it.   A link may connect to another document, an image file, or a document on the internet, in which case dtSearch will launch your web browser to access the document.

Copying Retrieved Text

To copy a block of text from a document to the Windows clipboard or to a file, mark the text (click and drag with the mouse) and click the **Copy** button in the button bar.

To copy an entire file, choose **Copy File** in the Edit menu while you are viewing the file.   If a file is stored in a ZIP archive, Copy File will extract it from the archive.

To copy multiple files from Search Results, press SPACEBAR to mark the documents you want to copy (a "*" will appear next to each selected document) and then choose **Copy File**.   You can also use SPACEBAR to mark multiple files to print, delete, or include in a Search Report (see below).

Create a Quick Summary of Your Search Results

An easy way to see all hits in all retrieved documents is to build a <u>Search Report</u>.   A Search Report shows all hits along with the amount of context that you request.   It also contains hypertext links to the hits in your files.

1.   Choose **Search Report** from the Search menu.   The Generate Search Report dialog box will appear.

2.   Enter the number of words of context that you want dtSearch to include in your Search Report and click OK to generate the report.

3.   In the Search Report, double-click on the highlighted hypertext links to jump to the location of hits in your documents.   Search Reports contain two types of hypertext links: highlighted page and paragraph numbers like `[Page 4 Paragraph 22]` which take you right to the location of a hit, and highlighted document names, which take you to the top of the document.

Scanning

If you have TextBridge® OCR by Xerox,® you can scan documents from within dtSearch.

1.   Install TextBridge if it is not already installed and test it by scanning a document using TextBridge directly.   Once you have TextBridge set up correctly, dtSearch will automatically recognize your settings when you scan from within dtSearch.

2.   In dtSearch, choose **Scan** from the File menu to access the <u>Scan</u> dialog box.

3.   Under **Document to create**, enter the name of the file you want to create and select the file type of the document you want to create.

4.   Check **Save image of document** if you want to be able to view the document as an image.

5.   Click the **Scan** button to start scanning the document.

6.   After you have scanned a document, it will appear in the **Scanned Documents** list in the dialog box. To index the documents listed, click the **Index** button.

7.   If you checked **Save image of document** when you scanned a document, you can click the **<u>Image</u>** button when you are viewing the document to see the scanned image of the document.

Recognizing an Existing Index

Recognize Index enables you to add an existing index to your index library, making it accessible for searching or indexing.   This can be useful on a network if you want to be able to search an index that another user created on the network.

To recognize an index, choose **Recognize Index** in the Index Menu (or click the **Recognize Existing** button in the <u>Index Manager</u>).   Use the Recognize Index dialog to locate one of the files in the index you want to recognize and choose OK.   (dtSearch index files have names like INDEX_R.IX.and INDEX_F.IX. They always begin with INDEX and end with .IX)   dtSearch will look in the directory for the index, extract the information it needs to recognize the index, and add the index to the list of indexes in the current index library.

Relative Paths

A relative path is a way to specify the location of a directory relative to another directory.   For example, suppose your documents are in I:\CHRON\DOCS and your index is in I:\CHRON\INDEX.   The absolute path for the documents would be I:\CHRON\DOCS.   The relative path from I:\CHRON\INDEX to I:\CHRON\DOCS would be ..\DOCS.

By default, dtSearch uses relative, rather than absolute, paths in indexes and index libraries.   Paths in an index are stored relative to the location of the index. The table of indexes in an index library stores index directories relative to the location of the index library. The relative path is more flexible because, if you move the entire I:\CHRON directory to another location such as J:\CHRON or I:\SMITH\CHRON or H:\CHRON, the relative paths to the documents would remain valid.

Relative paths are useful for shared indexes on networks when a physical drive uses a different drive letter for different users.   (For example, some users might see a drive as S: while other users see it as T:. On a peer-to-peer network, one user's C: drive might be another user's K: drive.)   As long as documents reside on the same drive as their index, dtSearch can use relative paths to prevent such drive mapping problems.

Relative paths are also useful for text database publishers who want to distribute a package with text and an index.   If the index is built using the relative paths option, users will be able to install the package on any drive and in any directory without affecting the validity of the index.

The only disadvantage to using relative paths is that, if you move an index to a different location and do not move the documents with the index, dtSearch may no longer be able to locate the documents because the relative path information will point to the wrong directories. If you need to disable the relative paths option, edit your OPTIONS.DAT file to include the line "RelativePaths = No".   Indexes created with relative paths disabled will store complete path information for every document.

Renaming an Index

To rename an index, choose **Rename Index** from the Index menu (or click the **Rename** button in the Index Manager), select the index to be renamed, enter the new name for the index, and click OK.   Note that the name of the directory in which the index is stored will not be affected.

Scanning Documents

If you have TextBridge® OCR by Xerox® (version 3), you can scan documents from within dtSearch.   To scan a document:

1. Install TextBridge if it is not already installed and test it by scanning a document using TextBridge directly.   Once you have TextBridge set up correctly, dtSearch will automatically recognize your settings when you scan from within dtSearch.

2. In dtSearch, choose **Scan** from the File menu to access the <u>Scan</u> dialog box.

3. Under **Document to create**, enter the name of the file you want to create, and select the file type of the document you want to create.

4. Check **Save image of document** if you want to be able to view the document as an image.

5. Click the **Scan** button to start scanning the document.

6. After you have scanned a document, it will appear in the **Scanned Documents** list in the dialog box. To index the documents listed, click the **Index** button.

Note: It is more efficient to index documents in larger batches, so if you have only scanned a few documents you may want to accumulate more documents before you index them.

Scanning Options

See your TextBridge manual for details on these options.

Preview	View a preview of a scanned page before OCR
Verify	Correct errors in a scanned page before saving the document.
Treat as fax	Document being scanned is a fax.
Language	Choose the language (English, German, etc.) of the document being scanned.
Page orientation	Select whether the document orientation is **Portrait** (longer from top-to-bottom than from side-to-side) or **Landscape**. If the document is mixed, select **Auto**.
Auto page segmentation	Check this box for documents with more than one column of text.
Ignore photos/Halftones	Check this box to speed processing of documents that contain pictures.
Brightness	Select the amount of light used to scan your documents. For HP scanners with HP AccuPage, choose **Auto**.
Resolution	Select the number of dots per inch (higher resolutions produce larger image files and take longer to process but result in more accurate character recognition).
Automatic document feeder	Check this box to instruct TextBridge to use the scanner's automatic document feeder, if available.

Troubleshooting

Scan menu option is grayed in the File menu. The **Scan** option is available only if dtSearch detects that TextBridge version 3 is installed.

Other scanning problems. If you are having difficulty getting scanning to work from within dtSearch, try using TextBridge directly. Once you have set up TextBridge to work correctly, you should be able to scan from within dtSearch using the same options. dtSearch will automatically detect the options you selected in TextBridge.

Search Filters

Click the **Filters** button in the search dialog box to access the Search Filters dialog box.   The Search Filters dialog box enables you to limit a search to files with a certain name, modification date, or size. The Search Filters dialog box contains the following filter options:

Name matches   Enter a filename filter like *.DOC.

Name does not match   To exclude documents enter a filter like *.EXE.

File size   Enter the maximum and minimum file size range (in bytes) for your search.

File date   Enter the starting and ending date ranges for your search.

You can leave any of these fields blank.   To clear all of the fields, click **Clear all**.

Search Macros

To edit or create a macro, choose **Macros** in the Options menu and click the **Edit** button.

Macros can be useful for abbreviating long names or phrases that you use frequently, or abbreviating field definitions in field searches.   A macro can contain anything that can be part of a search request.

A macro has two parts: a **Name** (up to eight letters), which you use to refer to the macro in search requests, and the **Expansion**, which is what the macro is expanded to.   A macro name must begin with the @ character.

For example, if you define the macro *@IRC* to mean *internal revenue code*, and then search for *standard deduction w/3 @IRC*, dtSearch will search for *standard deduction w/3 internal revenue code*.

Searching

The Search menu offers three ways to search for documents:

Choose **<u>Index Search</u>** to search using one or more document indexes.   An index search is the fastest way to search for documents. Before you can do an index search, you must first build an index.

Choose **<u>Unindexed Search</u>** to directly search files in one or more directories.

Choose **<u>Combination Search</u>** to combine an index search with an unindexed search.

The **Search** button on the button bar will take you directly to the type of search you used most recently.

The Search menu also lists options to view previous search results or previous search reports.   Choose **Previous Search** or **Previous Search Report** in the Search menu to see a list of all of the searches dtSearch has saved.   Select the one you want to see and choose OK.

Search Reports

A search report lists each hit found in each of the documents retrieved in a search with a few lines of context surrounding it.   It is like a combined KWIC view of every document retrieved in a your search.

To create a search report from a Search Results window, choose **Search Report** from the Search menu, enter the amount of context (words or paragraphs) you want surrounding each hit in your report, and click **OK**.   To include selected files in a search results list, move the cursor to each file that you want to include and press SPACEBAR to mark the file.   A * (asterisk) will appear next to each selected file.

In a search report, the highlighted document names are hypertext links to the full documents, and the highlighted page and line number markers are hypertext links to the places in the document where hits appear.   To use one of these links, double-click on it or move the cursor to it and press ENTER.   To get back to the search report, press ESCAPE.

Search Requests (Overview)

A search request consists of a group of words, phrases, or <u>macros</u> linked by connectors such as AND and OR that indicate the relationship between them.   Examples:

apple <u>and</u> pear	both words must be present.
apple <u>or</u> pear	either word can be present.
apple <u>w/5</u> pear	*apple* must occur within 5 words of *pear*.
apple and <u>not</u> pear	only *apple* must be present.
@addr <u>contains</u> 123	the field *@addr* must contain *123*.
apple w/5 <u>xfirstword</u>	*apple* must occur in the first five words.
apple w/5 <u>xlastword</u>	*apple* must occur in the last five words.

If you use more than one connector, you should use parentheses to indicate precisely what you want to search for.   For example, *apple and pear or orange juice* could mean *(apple and pear) or orange*, or it could mean *apple and (pear or orange)*.

<u>Noise</u> words, such as *if* and *the,* are ignored in searches.

Search terms may include the following special characters:

?	matches any character
*	matches any number of characters
%	<u>fuzzy search</u>
#	<u>phonic search</u>
~	<u>stemming</u>
&	<u>synonym search</u>
~~	<u>numeric range</u>

To enable fuzzy searching, phonic searching, synonym searchhing, or stemming for all search terms, check the boxes under **Search features** in the <u>search dialog box</u>.

Search Requests

Search Results Format

Choose **Search Results Format** in the Options menu to modify the format dtSearch uses to display search results.   In the Search Results Format dialog box, select the items you want to be displayed and click OK.   (Either the name of the retrieved file or its "title" (the first few words in the file) must be displayed.)

Search Tools

The **Search Tool** box in a search dialog box can display any of the four search tools described below. To select one of the tools, use the buttons at the top of the **Search Tool** box.   To insert a word, macro, previous search request or macro from the Search Tools into a search request, click on the item to highlight the it, then click the **Insert** button.

Word List
A scrolling list of the words in the index you have selected. Next to each word is a number, which is the number of times the word occurs in the index. As you type in a search request, the scrolling list will narrow in on the word you are typing.

If you have selected more than one index to be searched, the highlighted index will be the one listed.   You can select the index displayed in the Word List by clicking on the index you want to see in the Indexes to Search box.

Hints
The basics of creating search requests.

History
A list of your most recent search requests.   Next to each request is a pair of numbers, like "35/283," which tells you the number of files and hits that the request found.

Macros
A list of macros you have defined.

Searching for Documents

Field Searches

A field search is a search request that must be satisfied within a certain defined portion of a document.   A field search looks like this:

```
(beginning to end) contains (something)
```

The *beginning TO end* part defines the boundaries of the field.   The CONTAINS part indicates the words or phrases you are searching for in the field.   The only connector allowed in the beginning and end expressions in a field definition is OR.   Examples:

```
(name to address) contains john smith
(name to (address or xlastword)) contains (oak w/10 lane)
```

If a document contains more than one instance of a field, dtSearch will search each instance separately for text matching the search request.   The field boundaries are not considered hits in a search.   Only the words being searched for (*john smith, oak, lane*) are marked as hits.

Selecting Directories to Index

The Select Directories to Index dialog box allows you to pick the directories you want dtSearch to index. To add a directory to the list of directories, type in the directory name (or browse for it using the "Directories'" and "Drives" lists and click **Add to List**.   To remove a directory from the list, highlight it and click **Remove from List**.   To index a directory and its subdirectories, check the **Index subdirectories** box.

Selecting Directories to Search

The Select Directories to Search dialog box allows you to pick the directories you want dtSearch to search in an unindexed search.   To add a directory to the list of directories, type in the directory name (or browse for it using the "Directories'" and "Drives" lists and click **Add to List**.   To remove a directory from the list, highlight it and click **Remove from List**.   To search a directory and its subdirectories, check the **Index subdirectories** box.

Selecting Files to Copy or Print

In a Search Results window, you can select multiple files to print, copy, or summarize in a search report.

To select a file, move the cursor to it and press SPACEBAR (or choose **Select File** in the Edit menu). A * (asterisk) will appear next to the file.   To clear the *, press SPACEBAR again.

To select all of the documents, choose **Select All** in the Edit menu.

To clear all selections, choose **Clear Selections** in the Edit menu.

After selecting files, you can print, copy, or prepare a search report based on your selections.

Click the **<u>Print</u>** button to print the selected files.

Choose **<u>Copy File</u>** in the Edit menu to copy the selected files to a directory.

Choose **<u>Search Report</u>** in the Search menu to prepare a search report from the selected files.

Setting up an Application

Choose **Applications** in the Options menu to associate files with applications for launching.   (Note: You do not need to do this if the file associations maintained by Windows are sufficient for your purposes.)

To define a new application, select **<add new item>** and click the **Edit** button.   To edit an application that you set up previously, move the cursor to it and click the **Edit** button.   A dialog box will appear allowing you to specify the following for each application:

Name   Enter the name of the application.

File to Execute   Enter the name of the file to run to start this application.   For example, if you have WordPerfect 6.1 in a directory called C:\OFFICE\WPWIN you would enter C:\OFFICE\WPWIN\WPWIN.EXE (the name of the WordPerfect program).

Directory to run in   Enter a directory here if you want dtSearch to run the application in a specific directory.   If you leave this blank, dtSearch will not change directories before running the application.

Command line options   Enter any command line switches (e.g., /S, -D, etc.) that you want dtSearch to use when launching this application.   This is optional.

Window   Choose how the program should appear when launched (maximized, minimized, or normal).

Filename Filter   Enter a filter to match files with this application.   If you leave this blank, all files will be included.   This filter is similar to usual DOS filename filters (*.DOC, *.*, etc.), except that it can include drive and directory specifications, such as:

>   *\WP42*   Matches anything in a directory named WP42
>   C:*.DOC   Matches any filename ending in .DOC on drive C:

File type   Choose the file format that is associated with this application.   If you select **Any**, all files matching the filename filter will be associated with the application.   If you select WordPerfect 6.x, all WordPerfect 6.x files matching the filename filter will be associated with the application.

Setup Files

dtSearch stores user setup information in the following files:

APPS.DAT	Applications (for launching)
ENGLISH.ABC	Alphabet definition file for English-language text
FILETYPE.DAT	File type specifications
MACROS.DAT	User-defined macros
OPTIONS.DAT	Preferences
SRCHDIRS.DAT	List of directories to search (unindexed search)
SRCHIX.DAT	List of indexes to search (indexed search)
STEMMING.DAT	Stemming rules
WPCHARS.DAT	WordPerfect character set information

OPTIONS.DAT is a file that dtSearch uses to store your preferences about fonts, search options, etc. Most users will never need to access the file directly, but some advanced configuration options can only be changed by editing the OPTIONS.DAT file.   See OPTIONS.DOC for more information about the contents of the OPTIONS.DAT file.

Sorting Search Results

After a search is complete, the Sort Search Results dialog box will appear, listing several different ways to sort the results of your search:

Name   Sort by file name, in alphabetical order.

Hits   Sort by number of hits.

Date   Sort by file date.

Unsorted   Display documents in the order they were found.

Select the way you want to sort your search results and choose **OK**.

After a natural language search, dtSearch can sort retrieved documents by **Relevance**.   Weighting of retrieved documents takes into account: the number of documents each word in your search request appears in; the number of times each word in the request appears in the documents; and the density of hits in each document.

Stemming

Stemming extends a search to cover grammatical variations on a word.   For example, a search for *fish* would also find *fishing*.   A search for *applied* would also find *applying*, *applies*, and *apply*.   There are two ways to add stemming to your searches:

1. Check **Stemming** under **Search features** in the search dialog box to enable stemming for all of the words in your search request.   (By default, the box is checked.)   Stemming does not slow searches noticeably and is almost always helpful in making sure you find what you want.

2. If you want to add stemming selectively, add a ~ at the end of words that you want stemmed in a search.   Example: *apply~*

The stemming rules included with dtSearch are designed to work with the English language.   These rules are in the file STEMMING.DAT.   If you need to implement stemming for a different language, or if you want to modify the English stemming rules that dtSearch uses, you can create a new set of stemming rules to be used in place of STEMMING.DAT.   See the STEMMING.DAT file for more information about how to do this.

Supported File Types

Word Processor Formats

dtSearch 4.0 can automatically recognize, index, and display documents in the following formats:

    Ami Pro
    HTML
    Microsoft Rich Text Format
    Microsoft Word for DOS
    Microsoft Word for Windows versions 1, 2, 6, and 7
    Microsoft Works
    Multimate Advantage II
    Multimate version 4
    PKZIP versions 1.x and 2.0 (each file in a ZIP archive will be indexed as a separate document)
    WordPerfect and WordPerfect for Windows versions 5.0, 5.1, 5.2, 6.0, 6.1, 7
    WordStar versions 4, 5, 6
    WordStar 2000

Using 32-bit ODBC, dtSearch can also index and display records in Access, Paradox, FoxPro, and dBase databases.   Each record is treated as a separate document.

dtSearch can index and display, but cannot automatically recognize, documents in the following formats:

    WordPerfect 4.2
    WordStar versions before 4
    XyWrite

Choose File Types in the Options menu to tell dtSearch how to recognize these types of files.

Even if dtSearch does not support a file format, you can still index and search it.   See Unrecognized File Types for information about using dtSearch with these files.

Image Formats

dtSearch 4.0 can display images in the following formats:

    BMP
    EPSF
    GIF
    IMG
    JPEG
    PCX
    TIFF
    Targa
    WMF
    WPG (WPG version 1.0 only)

Synonym Searching

Synonym searching will find synonyms of a word that you include in a search request, using a <u>thesaurus</u> that you define.   A synonym group is a group of words that dtSearch treats as equivalent when performing a search.   For example, if you define a synonym group to include *improve*, *ameliorate*, *amend*, *better*, and *help*, then a search for *improve* would also find any of the other words in the ring.

Synonym searching works in combination with other search features like stemming.   If you enable both synonym searching and stemming in the above example, a search for *amending* would also find *improving*, *helped*, etc.

There are two ways to use synonym searching in a search:

1. Check **Use thesaurus** in the search dialog box to enable synonym searching for all of the words in your search request.

2. You can also add fuzziness selectively by adding the & character after words in your request. Example: *improve& w/5 search*

The File Menu

Thesaurus

Choose **Thesaurus** in the Options menu to create synonym groups to use in searching.   A synonym group is a group of words that dtSearch treats as equivalent when performing a search.   For example, if you define a synonym group to include *improve*, *ameliorate*, *amend*, *better*, and *help*, then a search for *improve* would also find any of the other words in the ring.

Synonym searching works in combination with other search features like stemming.   If you enable both synonym searching and stemming in the above example, a search for *amending* would also find *improving*, *helped*, etc.

To create a synonym group:

1.   Click the **Add** button in the Setup Thesaurus dialog box.

2.   Enter the words in the synonym group you want to create, separated by spaces or on different lines, and click OK.

To edit an existing group:

1.   Click on any of the words in the group in the **Words** list in the dialog box.   The synonyms of the word you clicked on will appear in the **Synonyms** list.

2.   Click the **Edit** button.

3.   Edit the list of words, adding or deleting words as needed, and click OK.

dtSearch
User's Manual

Copyright 1996 DT Software, Inc.

How to Contact DT Software, Inc.

Orders:	1-800-IT-FINDS (1-800-483-4637)
Technical Support:	703-827-0793
Internet:	www.dtsearch.com sales@dtsearch.com tech@dtsearch.com
CompuServe:	72607,3323
Telephone:	(703) 413-3670
Fax:	(703) 413-3473
Address:	2101 Crystal Plaza Arcade, Suite 231 Arlington, Virginia 22202 United States

Distributors

United Kingdom	ElectronArt Design (+44) 181 983 8686 (voice & fax) uksales@dtsearch.com
Australia	Indigo Pacific Pty Ltd 61 2 9955 8000, Fax 61 2 9955 8511 australia@dtsearch.com
Canada	Gatierf Publications Ltd. 1-888-IT-FINDS, (403) 285-1772, Fax (403) 293-6232 wright@gatpub.com

Legal Notices

be correct, DT Software will replace the defective diskettes or documentation.

Disclaimer of Additional Warranty.   Except as provided in the above Limited Warranty, dtSearch is provided AS IS.   DT Software makes NO OTHER WARRANTY OF ANY KIND, EXPRESS OR IMPLIED, INCLUDING WITHOUT LIMITATION, ANY WARRANTIES OF MERCHANTABILITY AND/OR FITNESS FOR A PARTICULAR PURPOSE.

Limitation of Liability.   To the extent permitted by law, the total liability of DT Software for any claim or damage arising out of the use of the licensed program or otherwise related to this license shall be limited to replacement of defective diskette(s) or documentation.   IN NO EVENT SHALL DT SOFTWARE BE LIABLE FOR ADDITIONAL DAMAGES, including lost profits, lost savings or other incidental or consequential damages arising out of the use of or inability to use the licensed program, even if DT Software has been advised of the possibility of such damages.   IN NO EVENT WILL DT SOFTWARE'S LIABILITY FOR ANY DAMAGES EVER EXCEED THE LOWER OF SUGGESTED LIST PRICE OR ACTUAL PRICE PAID FOR THE SOFTWARE.

Unisys LZW Notice.   LZW decompression is licensed from Unisys Corporation under U.S. Patent No. 4,558,302 and foreign counterparts.   Unisys requires the following license terms: "This software is restricted to use only on personal computers and workstations for the purpose of reading and/or writing GIF-LZW and TIFF-LZW images.   A separate purchase of this software is required for each additional personal computer or workstation which will use the software, either directly or via a server.   Customer may not modify or copy this software or make or develop any derivatives thereof, except that one copy may be made solely for backup or archival purposes."

Contents

+++++

Unindexed Searches

In an unindexed search, dtSearch directly searches every file that you ask it to search, looking for documents that match your search request. Unindexed searches are much slower than index searches. Choose **Unindexed Search** in the **Search** menu to access the Unindexed Search dialog box.

First, tell dtSearch where you want to search

1. Click the **Select** button next to **Directories to Search** to select directories to search.

2. To limit your search by filename, file size, or file date, click the **Filters** button to access the Search Filters dialog box.

Next, tell dtSearch what you want to find

1. Enter a search request in the space provided.   A search request consists of a group of words, phrases or macros linked by connectors such as "AND" and "OR" that indicate the relationship between them.   To have dtSearch enter search connectors for you, click the **Compose** button to access the Compose Search Request dialog box.

2. Under Search Features, choose the features you want to use in this search.

     Stemming searches other grammatical forms of the words in your search request.   For example, with stemming enabled, a search for apply would also find applies.

     Phonic searching searches for words that sound similar to words in your request, like Smith and Smythe.

     Fuzzy searching sifts through scanning and typographical errors.   Fuzziness adjusts from 1 to 10 depending on the degree of misspellings.   (Try starting with 3.)

     Use thesaurus tells dtSearch to use a thesaurus that you have created to find synonyms of words in your search request.

If you want to see how stemming, phonic searching, fuzzy searching will affect your search, click **Lookup**.

     Natural language searching is not available in unindexed searches.

Search Tools

A row of buttons across the top of the Search Tools panel selects the search tool that is displayed.   To copy text from the History list or Macros list, click the **Insert** button.

Hints	The basics of creating search requests.
History	A list of your most recent search requests. Next to each request is a pair of numbers, like "35/283," which tells you the number of files and hits that the request found.
Macros	A list of macros you have defined.

Unrecognized File Types

dtSearch recognizes most file types automatically.   If you are indexing only files such as major word processor documents, DBF files, ANSI text,   or ZIP files containing any of the above, you can disregard this section.   (To see a list of the file types that dtSearch recognizes, see **Supported File Types** in the on-line help.)

"Binary" Files

A "binary" file is a document that uses a file format that dtSearch cannot recognize.   By default, dtSearch will index everything in binary files as text.   This ensures that all words in such files will be searchable, but it can also result in indexing a lot of non-text data, which adds garbage words to your index.

Filtering Binary Files.   To tell dtSearch to index, search, and display only the text in binary files, choose **Index Options** in the Options menu and click the **Filter text** option under **Binary Files**.   **Filter text** treats data as "non-text" if it is not punctuation, an English-language letter, or a number.   This makes indexes smaller and also makes binary files appear cleaner when you view them in dtSearch.   To see how **Filter text** will affect your files, first view a sample file using **View File** in the File menu, then check **Filter text**, then view the file again.

Excluding Binary Files.   To avoid indexing files such as *.EXE and other program files, you can: (1) keep text files in separate directories and only index those directories, (2) use filename filters in the Update Index dialog box to exclude these files, or (3) check the **Do not index** box under **Binary Files** in the Index Options dialog box.

ASCII Text Files and Older Word Processor Files

Windows programs use the ANSI character set instead of the ASCII character set used in DOS.   The two character sets are mostly the same but the ANSI character set handles accented letters differently from the ASCII character set.   For plain text files created with a Windows program such as Notepad, dtSearch has no way to know whether the file is ANSI or ASCII.   By default, dtSearch assumes that text files use the ANSI character set.

XyWrite, WordPerfect 4.2, and WordStar versions before 4.0 have file formats that cannot be detected automatically.

To tell dtSearch how to recognize ASCII text files and these older word processor file formats, do the following:

1.   Choose **File Types** in the Options menu to access the Setup File Types dialog box.

2.   To set up a new file type, select **<add new item>** and click **Edit**.
     To edit a file type that you set up previously, move the cursor to it and click **Edit**.

3.   Enter the following information to set up the file type:

     Type.   Choose the file type you want to set up from the list.

     Filename Filter.   Enter a filename filter that will identify files having that format.   This filter is similar to usual filename filters (*.DOC, *.*, etc.), except that it can include drive and directory specifications, such as: *\WP42*, which matches anything in a directory named WP42, or C:*.DOC, which matches any filename ending in .DOC on drive C:.   You must enter a filename filter for each file type you want to set up, even if the filter is just *.*

Before using the file type information, dtSearch will try to detect the format itself.   Therefore, no matter what file type specifications you enter, dtSearch will recognize formats such as WordPerfect 6.1 that it can detect automatically.

Using dtSearch on a Network

Installing dtSearch in a Shared Directory

1. Install dtSearch in a directory that each user will have read-only access to.

2. Make each of the dtSearch program files read-only.   (This step prevents sharing violations on some networks.)

3. Set up private directories.   Each user will need a private, writeable directory that dtSearch can use for temporary files and to store the user's setup files.   This directory cannot be shared.

   The easiest way to set this up is for each user to use a directory on the user's own local hard disk (such as C:\DTSEARCH) as a private dtSearch directory.   Have each user run dtSearch using the /DIR command-line switch to designate a private directory, like this:

   ```
   DTSWIN /DIR C:\DTSEARCH
   ```

   Alternatively, if you want private directories to reside on the network rather than on each user's local hard drive, and if your network creates an environment variable containing each user's name (%USER% or %USERNAME%), you can use that environment variable with /DIR.   For example, to give each user to have a private directory on the shared H: drive:

   ```
   DTSWIN /DIR H:\%USER%\DTSEARCH
   ```

   An alternative to the /DIR command-line switch is to designate a private dtSearch directory by setting the %DTSDIR% environment variable in each user's AUTOEXEC.BAT file, like this:

   ```
   SET DTSDIR=C:\DTSEARCH
   ```

Shared Setup Files

dtSearch looks for setup files in each user's private directory first and, if it does not find setup files there, checks the shared dtSearch directory.   When dtSearch saves a setup file, it writes the file to the user's private directory. This makes it easy to set up all users with a common initial configuration that they can then customize to suit individual preferences.   To set up the initial configuration, do the following:

1. Run dtSearch directly from the shared network directory without the /DIR option, so all setup files will be stored in that directory.

2. Set up the configuration you want.

3. Exit dtSearch.

4. Make the new setup files read-only.

Shared Indexes and Concurrent Indexing and Searching

dtSearch supports multiuser access to indexes stored on networks.

Drive Mapping.   To avoid possible drive mapping problems, build an index on the same drive as the documents it indexes.   This prevents drive mapping problems since dtSearch, by default, uses relative

rather than absolute paths in indexes.   (See "Relative Paths" in the manual or on-line help.)

Index Libraries.   A user's index library is a file that lists the indexes a user can search or update.   On a network, index libraries can be either shared or private.   If all users share a common library, then each user will automatically be able to search indexes that are added to the library.   If a user has a private index library, the user can use Recognize Index to add shared network indexes to the user's index library.

If a user is using a shared index library that the user has read-only access to and the user creates an index, dtSearch will be unable to add the index to the shared library and will create a private index library for the user.

Read/Write Privileges.   Write and read access to shared indexes is controlled completely by the network's access permission settings.   If an index is stored on a network drive, any user who has write access to the directory containing the index will be able to update the index in dtSearch.   Any user who has read access to the index will be able to search the index or perform other functions (such as Verify Index and Copy Index) that do not require write access.

Concurrent Access.   When a user is updating an index, other users will be able to search but not update the index.   When a user is compressing an index, other users will be unable to search or update the index.

User Names (Optional)

If a network supports use of an environment variable such as %USER% or %USERNAME% to identify users, then users can enter that environment variable in **User Name** in **Preferences** in the Options menu.

Network Troubleshooting

After a search, dtSearch displays search results correctly but users cannot view retrieved documents.

1. Most commonly, this is caused by drive mapping problems.   For example, for the user who indexed the documents, the documents were in C:\DOCS, but for the user doing a search, the documents were in K:\DOCS.   These problems are especially common on peer-to-peer networks, because each user sees each drive differently.   To avoid drive mapping problems, build indexes on the same drive as the indexed documents.

2. Make sure all users have read access to the directory containing the documents.

An index is built in a shared directory but it does not appear in a user's list of indexes.

Have the user use **Recognize Index** in the Index menu to add the index to the user's index library.

When one user is running dtSearch from a shared directory and another user tries to run dtSearch, the second user gets a "sharing violation" message.

1. Make sure every user is using a separate private directory.

2. On some networks, a file cannot be shared unless it is read-only.   Use Explorer (in Windows 95) or Windows' File Manager (in Windows 3.x and Windows NT) to make every file in the shared dtSearch directory read-only.

Verifying an Index

To verify that an index is in good condition, choose **Verify Index** in the Index menu (or click the **Verify** button in the <u>Index Manager</u>).   As dtSearch examines the index, it will list every word, filename, and directory name in the index.   When dtSearch is done verifying the index, it will tell you whether the index has been damaged.

Viewing Images

To view an image file associated with a document that you are viewing, click the **Image** button in the button bar.   You can also choose **View File** from the File menu and enter the name of the image file you want to view.

Click the **Zoom In** button in the button bar, or press +, to magnify an image.   Click **Zoom Out**, or press -, to shrink an image.

Choose **Invert** in the <u>Local</u> menu, or press **I,** to invert the colors on an image.   This can be useful for viewing images of scanned documents, which are sometimes stored as white-on-black instead of black-on-white.

Viewing Multiple Windows

Use TAB or the mouse to switch from one window to the next.   (Use SHIFT-TAB to switch back.)

To close the window you are in, press ESCAPE.

To "cascade" your windows (rearrange them so that they overlap), choose Cascade from the Window menu.

To "tile" your windows (rearrange them so that they do not overlap), choose Tile from the Window menu.

Viewing Search Results

After a search, a Search Results window will appear listing the documents retrieved.   To view a retrieved document, double-click it or move the highlight bar to the document name and press ENTER.

To see a summary of your search results, choose **Search Report** in the Search menu.   dtSearch will generate a summary listing all hits in all documents, along with as much context as you request.   Choose **KWIC** in the Search menu for a summary listing the hits in a single document.

In a search report or KWIC window, the highlighted document names are hypertext links to the full documents, and the highlighted page and paragraph number markers like `[Page 5 Paragraph 2]` are hypertext links to the places in the document where hits appear.   To use one of these links, double-click on it or move the cursor to it and press ENTER.

Getting Around

Click **Next Hit** in the button bar, or press SPACEBAR, to see the next hit in a retrieved document.

Click **Prev Hit** in the button bar, or press BACKSPACE, to see the previous hit in a retrieved document.

Click **Next Doc** in the button bar to see the next document in the search results list.

Click **Prev Doc** to see the previous document in the search results list.

Cursor commands:
Top of document	Ctrl-Home or Home,Home,Up
End of document	Ctrl-End or Home,Home,Down
Page up	PgUp
Page down	PgDn

Press ESCAPE to close a document or Search Results window.

Press TAB (or click with the mouse) to switch between windows.

Finding Text

To search a document for a word or sequence of characters, click **Find Text** in the button bar.   Find Text is like the search function in your word processor -- you can search for words, spaces, punctuation, etc. Capitalization is ignored, and the * and ? characters are not treated as wildcards.

Selecting Files

To select files to print, copy, delete or include in a search report, use the SPACEBAR key in Search Results.   An * (asterisk) will appear next to each file you select.   To clear the *, press SPACEBAR again. To select all of the documents in the list, choose **Select all** from the Edit menu.   To clear all selections, choose **Clear selections**.

Local Menu

In a document window, Search Results, image file, Search Report, or a KWIC view, click the right mouse button anywhere to pop up the Local menu.   The Local menu for a window lists all of the things you can

do with the window.   The Local menu is different for each type of window.   For example, the Local menu for a search results window includes an option to generate a search report.   The Local menu for a retrieved document includes an option to skip to the next document retrieved in the search.

Viewing Search Results

Viewing a File

To view a file, choose **View File** in the File menu.   A dialog box will appear asking for the name of the file you want to view.   Type in the name of the file and press ENTER.

Click the right mouse button to pop up the <u>Local Menu</u>   for the document   This menu provides quick access to all commands that work with retrieved document windows.   All of the options listed below are available in the Local menu.

<u>Finding Text</u>   To search a document for a word or sequence of characters, click the **Find** button or press F2.

<u>Printing Text</u>   Click the **Print** button to print a document.

<u>Launching Files</u>   Click the **Launch** button or press F8 to "launch" the document.   If you have set up an application for the type of document you are viewing (for example, WordPerfect for a WordPerfect document), dtSearch will execute that application with the file you are viewing.

Cursor commands

Top of document   Ctrl-Home or Home,Home,Up
End of document   Ctrl-End or Home,Home,Down
Page up   PgUp
Page down   PgDn

Viewing a Retrieved Document

To view a file retrieved in a search, move the cursor to the file you want to see in the search results window and press ENTER.   The document will appear in a new window overlaying the search results window.   Hits from your search will be highlighted.

Click the right mouse button to pop up the <u>Local Menu</u>   for the document   This menu provides quick access to all commands that work with retrieved document windows.   All of the options listed below are available in the Local menu.

<u>Finding Text</u>   To search a document for a word or sequence of characters, click the **Find** button. Find is like the search function in your word processor -- you can search for words, spaces, punctuation, etc.   Capitalization is ignored, and the * and ? characters are searchable like any other characters (they are not treated as wildcards).

<u>Printing Text</u>   Click the **Print** button to print the document.

Finding Hits
Click the **Next Hit** button or press SPACEBAR to search the document for the next hit in your search request.
Click the **Prev Hit** button or press BACKSPACE to go back to the previous hit in the document.
Click the **Next Doc** button to see the next document in the search results list.
Click the **Prev Doc** button to see the previous document.

<u>Launching Files</u>   Click the **Launch** button to "launch" the document.   If you have set up an application for the type of document you are viewing (for example, WordPerfect for a WordPerfect document), dtSearch will execute that application with the file you are viewing.

Cursor commands

Top of document	Ctrl-Home or Home,Home,Up
End of document	Ctrl-End or Home,Home,Down
Page up	PgUp
Page down	PgDn

W/N Connector

Use the W/N connector in a search request to specify that one word or phrase must occur within N words of the other.   For example, *apple w/5 pear* would retrieve any document that contained *apple* within 5 words of *pear*.   The following are examples of search requests using W/N:

```
(apple or pear) w/5 banana
(apple w/5 banana) w/10 pear
(apple and banana) w/10 pear
```

Some types of complex expressions using the W/N connector will produce ambiguous results and should not be used.   The following are examples of ambiguous search requests:

```
(apple and banana) w/10 (pear and grape)
(apple w/10 banana) w/10 (pear and grape)
```

In general, at least one of the two expressions connected by W/N must be a single word or phrase or a group of words and phrases connected by OR.   Example:

```
(apple and banana) w/10 (pear or grape)
(apple and banana) w/10 orange tree
```

If you enter an ambiguous search request, dtSearch will display a message warning you of the error.

dtSearch uses two built in search words to mark the beginning and end of a file: *xfirstword* and *xlastword*.   The terms are useful if you want to limit a search to the beginning or end of a file.   For example, *apple w/10 xlastword* would search for *apple* within 10 words of the end of a document.

What is a Document Index?

A document index is a database that stores the locations of all of the words in a group of documents except for <u>noise</u> words such as *but* and *if*.   A document index can contain up to 1,000,000 documents.

Once you have built an index for a group of documents, dtSearch can use it to perform very fast searches on those documents.   An index does not store the text of documents, so to view retrieved documents you will still need to keep the documents accessible.

To build an index, dtSearch needs free space (on the drive containing the index) at least equal to the size of the documents to be indexed.   This is because dtSearch creates temporary files during indexing.

A document index is usually about one fourth the size of the original documents, although this may vary considerably depending on the number and kinds of documents in the index.   In general, the more documents in the index, the smaller the index will be as a percentage of your original documents.

Wildcards (* and ?)

A search word can contain the wildcard characters ***** and **?**.   A **?** in a word matches any single character, and a ***** matches any number of characters.   The wildcard characters can be in any position in a word. For example:

appl*   would match apple, application, etc.
cipl   would match principle, participle, etc.
appl?   would match apply and apple but not apples.
ap*ed   would match applied, approved, etc.

Use of the ***** wildcard character near the beginning of a word will slow searches somewhat.

Words and Phrases

You do not need to use any special punctuation or commands to search for a phrase.   Simply enter the phrase the way it ordinarily appears. You can use a phrase anywhere in a search request.   Example:

```
apple w/5 fruit salad
```

If a phrase contains a noise word, dtSearch will skip over the noise word when searching for it.   For example, a search for *statue of liberty* would retrieve any document containing the word *statue*, any intervening word, and the word *liberty*.

Punctuation inside of a search word is treated as a space.   Thus, *can't* would be treated as a phrase consisting of two words: *can* and *t*.   *1843(c)(8)(ii)* would become *1843 c 8 ii* (four words).   (To customize the way dtSearch handles punctuation in text, choose **Index Options** in the Options menu and click the **Edit Alphabet** button.)

Working with Indexes