

# oUSE: User's Manual

*By: Andrew Socha, Collin Dobmeyer, and Jimmy Hartnett*

# Table of Contents

Section 1 – Initialization:	Page 1
Section 2 – Querying the Search Engine:	Pages 1-2
Section 3 – Managing the Search Engine:	Page 2

## Section 1 – Initialization:

When first accessing the Search Engine, the user must decide whether to parse the entire file corpus into a single file. Enter 'y' to parse the file set or 'n' if this has already been done. If yes, then you must also enter the file path to the file corpus. The files in the corpus should be renamed with an incrementing file scheme (e.g: 0 (1), 0 (2), 0 (3), ...). Most modern Operating Systems will do this when all files are in one folder; just select all files, and rename to '0' . Next, the user has the choice between using an AVL Tree and a Hash Table to store the Search Engine's index, which is used by the engine to search for specific words. Enter 'AVL' to use an AVL Tree or 'HashMap' to use a Hash Table. Note that this choice only affects the speed and efficiency of the Search Engine and will not affect what results you receive when searching. Next, the user must read an index from a file into either an AVL Tree or a Hash Table, as chosen above, by entering the file path of the index's file. Afterward, the user is brought to the Querying section of the Search Engine.

## Section 2 – Querying the Search Engine:

The user can query the Search Engine with specific keywords and receive a list of documents that include those keywords and are considered most relevant. From this list, the user can choose a specific document to view by entering 'CMDOPEN' which will prompt you for the search result number. Enter that number to view the document.

To query for a single word, just type that word. To query for multiple words and only include documents that have *all* of the words, enter 'AND' followed by each of the words you want to search for. To query for multiple words and include documents that include *any* of the words, use 'OR' followed by each of the words you want to search for. Note that you

cannot use both AND and OR searching at the same time. You can remove files documents your query that have a specific word by using 'NOT' followed by any words you want to remove, placing NOT after all of the words that you are AND or OR searching for. You can also restrict your query to a specific date range using 'DATELT' (date less than) and/or 'DATEGT' (date greater than) at the very end of your query, followed by the date in the form mm/dd/yyyy or mm-dd-yyyy. See the examples below for further assistance.

Examples of Queries:

- Politics DATEGT 01/01/2000 DATELT 01/01/2010
- OR Dallas Houston Austin Texas NOT weather Rangers wildfires
- AND Saint Louis Missouri NOT Cardinals DATELT 07-19-2007 DATEGT 05-22-1998

### Section 3 – Managing the Search Engine:

Managing the Search Engine requires typing specific commands into the query. To view a list all commands, enter 'ENTERCOMMAND'. The manager can switch between using an AVL Tree and a Hash Table with 'SWITCH\_DATA\_STRUCTURE' and then 'y' to confirm when prompted. The manager can write the index to a file by entering 'WRITE\_INDEX\_TO\_FILE'. Then, enter a fully qualified file path when prompted. The manager can clear the index by entering 'CLEAR\_INDEX'. The manager can also add a new document to the index using a file of his own, giving the search engine access to that document during a search, by entering 'ADD\_FILE\_TO\_INDEX'. Then the manager must enter the file path of the new file.