The Search Engine project’s components and working is described below:

Packages:

There are 3 packages in the Search Engine Project:

1. ComplxStruct – Logic of Searching the hash map
2. SearchEngine – Basic function calling
3. WebCrawler – Removing html tags and pulling the body content of the webpage in separate words.

Working:

1. The following 5 URLs are inserted into ocList[] array

* “http://espnfc.us"
* "http://www.espnfc.us/scores"
* "http://www.espnfc.us/transfers/news"
* "http://www.espnfc.us/video"
* “<http://www.espn.com/>”

1. Using WordCrawler function the html content of each of the 5 links are pulled and inserted into input[index] array.
2. The user enters a search keyword as a console input. The Scanner Class helps in scanning the keyword throughout the html content retrieved in input array.
3. If the result is obtained the links are displayed, else Not Found! Error is thrown.

The SearchEngine java project can be found in the path “alakshm1/SearchEngine”.

The Testing results can be found in the path “alakshm1/Testing.docx”.