

Name : Sambit Mishra  
UIN : 720002013  
Assignment : Homework 1  
Course : Information Storage and Retrieval – CSCE 670

### **How to use the Search Engine:**

The engine has been build using Python on Ubuntu 12.04 platform. In order to run the code type the following command at the shell :

**python Homework1.py directory\_path**

directory\_path is the path where all the files are present over which the search has to be performed.

Eg.

**python Homework1.py /home/sambit/IR/Homework1/books/books**

In the above example the files were present inside **/home/sambit/IR/Homework1/books/books**.

Once you run the above code , the search engine will build the indexes. It took hardly 2 minutes to build the indexes. Once the index have been build , you will asked to input 1 for Querieng and 2 to exit the search engine. Input the required choice and search.

### **Examples of some search queries:**

**Boolean Queries** : joseph frederick baltimore , mark twain

**Phrase Queries** : “mark twain” “what is man” , “modern structures demanded” [ Note: If you want to search phrase queries use double quotes[“ query term” ] to surround your query term.

**Wild Card Queries** : jose\* ,

**Combination Query** : quixo\* “mark twain”, w\*w x\*x y\*y "crazy" query

Once you have entered the query term, hit enter, you will get the result as following:

### **Sample Result:**

Query : quixo\* “mark twain”

Result :

The query is found in documents  
['416.txt', '245.txt', '76.txt', '543.txt']