**Online Multi Type Documents Search Engine**

**ABSTRACT**

The explosion in the amount of semi-structured data users access and store in online multi type documents search engine systems, there is a critical need for powerful search tools to retrieve data in a simple and efficient way. Existing tools typically support some ranking on the textual part of the query, but only consider structure (e.g., file directory) and metadata (e.g., date, file type) as filtering conditions. We propose a multi-type search engine approach that allows users to perform search engine for structure and metadata conditions in addition to keyword conditions. Our techniques individually score each document and integrate the multi-type search into a meaningful unified score. We also show that our online multi type

Documents search strategies perform and scale well, making our search engine approach practical for every day usage.

**EXISTING SYSTEM**

Documents search tools have been developed to perform keyword searches and locate search document stored in file systems, such as the commercial tools Google Desktop Search .However, these tools usually support some form of *ranking* for the textual part of the query , but only file directory and date, file type as *Filtering* conditions. Recently, the research community has turned its focus on search over to data collections and documents .However, as is the case with commercial file system search tools, these works focus on keyword queries and use other system information only to guide the keyword-based search engine

**PROPOSED SYSTEM**

In this project, we propose a novel approach that allows users to efficiently perform document searches across three different dimensions: file directory, date, and document name. We describe individual approaches for each document and present a unified scoring framework for multi-type queries over search engine file systems. We also present new data structures and index construction optimizations to make finding and storing document, files…

### IMPLEMENTATION

Implementation is the stage of the project when the theoretical design is turned out into a working system. Thus it can be considered to be the most critical stage in achieving a successful new system and in giving the user, confidence that the new system will work and be effective.

The implementation stage involves careful planning, investigation of the existing system and it’s constraints on implementation, designing of methods to achieve changeover and evaluation of changeover methods.

**MODULE DESCRIPTION**

**Main Modules:**

1. **User Login Module:**

In Login Form module presents site visitors with username and password fields. If the user enters a valid username/password combination they will be granted access to additional resources on website. Which additional resources they will have access to can be configured separately.

**2. Online Upload &Download Documents:**

The admin encrypts the input file using the specified file id. It then store the file date and the documents onto the database. In our implementation, upload the document ,the file is stored by using database.

**3. File Date Creation& Document:**

In this module we are creating a metadata for all the system files.

The Module is going to save all file names in a database in addition to that, it also saves some information from the text file. This mechanism is applied to avoid the long run process of the existing system.

**4. Searching Files or Document:**

In this module the user going to enter the text for searching the required file .The searching mechanism is differing from the existing system. Whenever the user gives their searching document. It is going to search from the database .At first search is based on the filename .After that, it contains some related document name. Then it collects some of the file documents, it makes another search tools or engine. Finally it produces a search results for corresponding related document for the user.

# System Configuration:-

**Hardware Requirements:**

* System : Pentium IV 2.4 GHz.
* Hard Disk : 40 GB.
* Floppy Drive : 1.44 Mb.
* Monitor : 15 VGA Colour.
* Mouse : Logitech.
* Ram : 512 Mb.

**Software Requirements:**

* Operating System :Windows95/98/2000/XP
* Application Server : Tomcat6.0.18
* Front End : HTML, Java, Jsp
* Scripts : JavaScript.
* Server side Script : Java Server Pages.
* Database : mysql 5.0
* Database Connectivity : JDBC.

**CONCLUSION**

We defined multi-type document and file relaxations and proposed based scoring approaches for keyword, date, file directory and structure query conditions. This uniformity of scoring allows to be search engine easily aggregated. We have also designed indexing structures and dynamic index construction and query processing optimizations to support efficient evaluation of multi-type search engine queries.

We implemented and evaluated our scoring framework and search engine query processing techniques. Our evaluation show that our multi-type score aggregation technique preserves the properties of individual document scores and has the potential to significantly improve ranking accuracy.

We also show that our indexes and optimizations are necessary to make multi-type search tools.