**Problem description:**

The project that is being developed is not present in the market as if so, but there are similar projects but those can only work on Linux and other operating systems;The project designed are on different platforms and language because of the efficiency provided by the other languages such as robust, OOP, Platform independency,security etc. So, the goal is to design the project without the use of those characteristics.

The problem with the projects developed on other language is speed. As we know, C language is considered as the fastest output processing language because of the fact that it works on machine level. As every other project developed on other language needs to travel through levels until it reaches the machine level; so, because of the C language the speed is not compromised and we get the output quickly as compared to others.

And the project will not be only consisting of a search engine, but it will have some extra features such as assistant, account login for users, administrators and newusers providing safety and notification for the designed system.

**Domain Knowledge:**

Abstract - The generic search engines existing today are using ‘search as you type’ technique while searching the information in order to retrieve information faster as compared to keyword search. But generic search engine fails to distinguish different users. Each user is unique and has got his/her unique interests. The generic search engine showssame results to different users without being bothered of their needs/interests. Most of the personalized search enginesretrieve the relevant information but use only the keyword search technique. The project proposes domain knowledgedriven personalized web search by using the typing technique in order to retrieve pertinentinformation and compare theresults with that of keyword search.

Let’s take analysis of different developed search engines methodologies for efficient search results and the proposed method for crucial web page results. Different search engine approaches are applicable for efficient prediction of web page results. Some of the search engines have their own working model with special features. Google has features like android specific application where as Microsoft Bing has application specific to windows platform and yahoo is lagging far behind them, but their main focus is at services like Yahoo Mail, Yahoo answer. Every search engine has its own advantages and disadvantages.

Domain Knowledge Driven Personalized Web Search Engine allows user to traverse easily by suggesting his domain related interested pages at runtime with help of its browsing history. Some popular search engines comparison based on crucial points to cover variety of features provided by them is gathered.

**Scope of project:**

Web search engine is a data structure based system that is designed to search for information on the World Wide Web.The results obtained after searching are generally presented in a line of results often referred to as search engine result pages. The information may be a mixture of web pages, images, and other types of files. Some search enginesperform mining of the data that is available in databases or open directories. Unlike most of the web directories, which are maintained only by human editors, search engines also perform the task of maintaining real-time information by running an algorithm on a web crawler. Web search engines perform the work by storing information about several web pages, and later retrieve them from the HTML mark-up of the pages. These pages are retrieved by a Web crawler which follows every link on the site. Search engine then analyses the contents of each web page and determines howindexing should be performed (for example, words can be extracted from the titles, large page content, headings, or other special fields called meta tags).Data regarding web pages are stored in an index database for the use in later queries. A query from a user can be a single word or group of words. The index helps find information relating to the query as quickly as possible. This cached page always holds the actual search text since it is the one that was actually indexed, hence it can be very fruitful when the content of the current page has been updated and the search terms are no longer in it. This paper proposes architecture for constructing search engine using domain knowledge and user history.

**Objective of the project:**

Web search engines commonly provide search results without considering the user interests or context. It proposes +a personalized search methodology that can easily extend a conventional search engine on the client side as well as on the server side. The mapping framework automatically maps a set of known user interests onto a group of categories in the database and takes advantage of manually edited data that is available in database for training the text classifiers, thereby categorizing and personalizing search resultsaccording to user interests. The objectives of the project are:

1. More accurate the data more accurate the result.

2. Objective in proposed is to provide efficient algorithm for page hits calculation

3. Objective here is not only to have correct prediction forweb pages but to make algorithm generalized

4. Correctly detected bounced results & false entriesrectification will give us the correct predictionwhilesearching.

**Operating Environment:**

Processor-Pentium Dual core processor.

Operating System-Windows XP/Vista/7/8/10.

Memory-1 GB.

Hard disk-500 GB.

Database-Windows File System.

C:\Users\exam\Downloads\Untitled Diagram (1).png

**Phases**

**Days**

SYSTEM STUDY:

2.1 software  requirement specification:

Microsoft Word 2010

**Turbo C** **Version 2.0**

**DOSBox**

2.2 EXISTING SYSTEM

Now a days, the World Wide Web has contributed a lot in

searching information.But still there is room for improvement

because current search engines do not consider the specific

user’s interest and serves each user equally. For the generic

search engine, it become difficult to identify what the user

actually want.

LIMITAIONS:

1. It is difficult for the user to get relevant content.

2. Typical search engine returns same result for

different users hitting same query

2.3 PROPOSED SYSTEM

Framework for proposed system

Usually, with the help of present web search engines users

often miss the goal of their searching or receive the

ambiguous results. But the framework of personalized web

search considers the specific user's interest and suggests the

relevant pages of his/her interest. We have proposed a simple and efficient model which ensures good suggestions as well as

promises for effective and relevant information retrieval.

FUNTIONALITIES:

1.It will be easy for the user to get relevant content.

2.Unlike typical search engine, result for different

users hitting same query will be according to the user

2.4 software tool used:

Data Structure is a way of collecting and organising data in such a way that we can perform operations on these data in an effective way. Data Structures is about rendering data elements in terms of some relationship, for better organization and storage.

In search engine project we are using following tools:-

SEARCH

The simplest, most general, and least efficient search structure is merely an unordered sequential [list](https://en.wikipedia.org/wiki/List_(computing)) of all the items. Locating the desired item in such a list, by the [linear search](https://en.wikipedia.org/wiki/Linear_search) method, inevitably requires a number of operations proportional to the number *n* of items, in the [worst case](https://en.wikipedia.org/wiki/Worst_case_complexity) as well as in the [average case](https://en.wikipedia.org/wiki/Average_case_complexity). Useful search data structures allow faster retrieval; however, they are limited to queries of some specific kind. Moreover, since the cost of building such structures is at least proportional to *n*, they only pay off if several queries are to be performed on the same database. Static search structures are designed for answering many [queries](https://en.wikipedia.org/wiki/Information_retrieval) on a fixed database; dynamic structures also allow insertion, deletion, or modification of items between successive queries.

LINKLIST

A linked list is a linear collection of data elements, in which linear order is not given by their physical

placement in memory. Instead, each element points to the next. It is a data structure consisting of a group of nodes which together represent a sequence. Under the simplest form, each node is composed of data and a reference (in other words, a link)

 to the next node in the sequence. This structure allows for efficient insertion or removal of elements from any position in  the sequence during iteration. More complex variants add additional links, allowing efficient insertion or removal  from arbitrary element references. A drawback of linked lists is that access time is linear. Faster access, such as random access, is not feasible. Arrays have better cache locality as compared to linked lists.

sorting

Sorting refers to arranging data in a particular format. Sorting algorithm specifies the way to arrange data in a particular order. Most common orders are in numerical or lexicographical order.

The importance of sorting lies in the fact that data searching can be optimized to a very high level,

if data is stored in a sorted manner. Sorting is also used to represent data in more readable formats.