

---

## SKILL DEVELOPMENT IS A NEW VISTA FOR CUSTOMISING INFORMATION RETRIEVAL THROUGH SEARCH ENGINE

Puspanjali Jena

### Abstract

*The early days of the WWW changed retrieval tools in many ways. Retrieving searching and filtering massive uncontrolled database of the Web pages presents new challenges. The role of traditional libraries has been changed by the introduction of Internet access. So the users now have access to more information than before; that they must rely on search engine to retrieve and rank the results of their searching. Thus this paper highlights the customer-based information retrieval through search engine by classifying the need of customer into five categories such as functional, hedonic, innovative, aesthetic and sign needs of customer.*

**Keywords :** Skill Development, Search Engine and Information Retrieval

### 1. Introduction

Web search technology is an emergent means in the field of customer oriented services over the past few years. In less than ten years since the World Wide Web came into being, web search has become quite sophisticated and is continuously improving to meet the demands of variety customers. A web search is a process of accessing information available in the global information system of the World Wide Web, where search engines are essentially query generators that use key words and logical relationship to find web documents or files to meet the search requirement and create a virtual document containing links to the documents found in the search.

A search engine on search services is a program designed to help and find information stored on a computer system named as World Wide Web inside a corporate or proprietary network or a personal computer. Through search engine, web provides an enormous amount of information which, on one hand offers an inexhaustible source for searching material, search engine allows one to ask for content meeting specific criteria and retrieve a list of reference that match those criteria. Search engine usually refers to a web search engine, which searches for information on the public web. Other kinds of search engine which search individual personal computers and mobile search engine. All those stated search engines operate in different relevant criteria. In a nutshell search engine is a tool that to find and retrieve information from the web.

The search engine tool has become an instrumental in connecting the wide away of the web sites, have gained sophistication over time through essentially cybernetic techniques. Through the Boolean search where searches of databases link the occurrence of one word of phase with another and

---

then those sources that have both these elements in their contents, is at the base of the web search engines. With the exponential growth of web sites the capacity to search the entire web has made the search engines some of the nodal hubs of the entire web in essence weight station or web traffic.

## 2. History of Search Engine

The early days of the World Wide Web has changed its retrieval tools in many ways. Retrieving, searching and filtering a massive, uncontrolled database of web pages presented new challenges.

Web crawler, the World Wide Web worm along with others began "crawling" the web to gather parts of the web pages usually titled URLs, and index them. Gradually they applied algorithms to increase precision and displayed them to users. Due to increase of web in size and scope, the retrieval problem became more complex and search engines increased their capacity, intention and sophistication for handling them accordingly.

The ability to conduct search engine of the Internet came about in the early 1990s when the World Wide Web was developed by CERN. A study of W3C (1992) reports that WWW was originally developed to allow information sharing within internationally dispersed terms of scientists and to provide dissemination of information by support groups primarily within high-energy physics community. Later on WWW rapidly spread the other areas and attracted much Internet in user support, resources discovery and collaborative work efforts.

In 1992 the first portable browser software was released by CERN as freeware, then after parallel efforts were made in other web browser including Gopher system developed by the University of Minnesota

Gronov (2000) expressed that by 1995 Microsoft Internet explorer and Netscope 'Navigator' were competing for the browser market and a number of other Internet search engines were being developed.

Search engine on the web has a short and concise history because it is of recent origin. The first web search engine was wandex, a new defunct index collected by the World Wide Web wanderer; a web dawdler developed by Mathew Gary at MIT in 1993 launched these engines as a way to demonstrate hardware capacity. In 1993 Aliweb search engine also appeared.

1994- the first "full text" Crawler-based search engine web crawler came out. It is the most popular search engine which let users search for any word in any web page and became the standard for each search engine since then. It is most widely used search engine by public.

1994- (Lycos) : Lycos search engine is the product of Carnegie Mellon university, which is a major commercial endeavour. Lycos indexed the first 100-word page of a web page, and open text began as an attempt at field searching.

1994- (Infoseek) : This search engine proved to be effective for web search storage method, and was adopted by Inktomi.

1995- Altavista and Excite are the two important search engines. They specify primarily language, geographic location and subject of the site. These search engine usually spend their infancy in academia, as they grow they emerge as a profit making agency and launched.

Thus in 1996 Dog pile, Inktomi, Ask Jeeves, 1997 Northern light, 1998 Google, 1999 besides other web link Baidu, 2000, Singing fish, Teoma, Vivisimo, 2003 object search, 2004 Yahoo services MSN search, 2005 MSN search launched 2006 Quaero and Ask.com were launched.

Thus search engine were also known as some of the brightest stars in the Internet before the advent of the web. There were search engine for other protocols or users such as the Archie search engine for anonymous FTP sites and the Veronica search engine for Gopher protocols. Some of search engines are:

**Google :** In the year 1998 Google search engine emerged but around 2001 the Google search engine rose to prominence. It is based on two concepts (i) Link popularity (ii) Page rank. Through the page rank, the other web sites and web pages link to a given page. The page rank of the linking page and the number of links on these pages contribute to the page rank of the linked page. Google and most other search engine utilities not only page rank but more than 150 criteria to determine the relevancy of page rank, which is based on citation analysis. It was developed in the 1950 by Eugene Garfield at the university of Pennsylvania Web link analysis was first developed by Dr. Jon Kleinberg and his team while working on the CLEVER project IBM's Almaden Research centre. Google is eventually the most popular search engine.

**Yahoo / Search:** - Yahoo acquired Inktomi in 2002 and over tune in 2003, which owned all the web and Alta Vista. Though it has own search engine but initially it used Google to provide users with search results on its own main web site yahoo.com, but in 2004 yahoo launched its own search engine based on the combined technologies of its acquisitions and providing a service that gave pre-eminence to the web search engine over the directory.

**Microsoft** - Recently the most popular search engine is MSN search, owned by Microsoft, which is relied on others for its search engine. In 2004, it brought out its beta version, powered by its own web crawler (called msnbot). Since 2005 it started showing its own results live.

**Ask.com** - In Feb 2006, Ask Jeeves was renamed on Ask.com, Ask.com is now algorithmic engine. Ask.com uses the ranking algorithmic which was originally developed for Jeeves and continues to increase in size. In 2005 it was reported that Ask.com did a great deal of work to reduce the number of advertisement on a result page.

---

Search engine do not have a client base link other commercial database producers/ vendors such as avid, Lexis-Nexis, and other. Usually some end users are not willing for a retrieval service on the web. So search engine started to function with an advertising or entertainment model. Banner advertisement partnership with companies such as Amazon, and selling search keywords for a fee were all attempts to generate search engine fundamentals.

### 3. Definition of search engine

In the present phenomenon search engine has been define by different computer technologists in the following ways:

*Peter and Olson* (1996) have viewed on information search behaviour, which comes from the exploration of consumer behaviour. In such discussion, consumers are exposed to some marketing information through their own internal, goal-direct search for relevant marketing information to help complete purchasing behaviour.

*Gordon and Pathak* (1999) have commented that in recent years, the World Wide Web (WWW) has become incredibly popular at homes and offies alink. There is no question that the number of ww users will swell significantly in the near future. Through many internet-enable applications and services are available today, the primary use of the internet (other than e-mail) is for information retrieval.

*Bradlow and Schmittlein* (2000) studies have shown that no single search engine indexes more that about 16 percent of the web, making much of information on the web effectively in -accessible. A significant issue is that search engine do not index all sits eually, and new pages may not be indexed for months after they are placed on the web.

### 4. Search engine fundamentals

Search engine is giant applications that find information, which has really three different applications but they work together to find and retrieval information. The three fundamentals search engines are:

*An indexer* - This deals with back end application that is finds and indexes pages for incorporation in a database, other names for this type of application includes spider, crawler and robot.

*A database* - Its application to store the indexed references to web page.

*A quing interface* - This application handles the queries submitted by the users.

Today there are many search engines available. Some have more complex logic operators, and present their results in a more: User friendly" manner. There are improved versions of older browser or search software, and new software is being introduced.

The powerful search engines such as Lycos, AltaVista and Google provide the powers to do searches of materials throughout the web. The web has transformed the personal computer into a smarter machine. The search technologies, which have been instrumental in connecting the wide array of web sites, have gained in sophistication over time through essentially cybematic technologies. The Boolean search, where searcher database link the occurrence of one word or phase with another and displays those sources that have both of these elements in their contents is at the base of the web search engine.

### 5. Features of search engine

The fundamental purpose of a search engine is to index web sites in a way that allows customer to use keywords to find web pages internet to him. To do this search engine rely on a computer called an indexer, spider or erawler to refer out the customer pages to use key words to find web pages that interest to customer. To do this search engine rely on a computer to ferret out the pages at one's site and then create indexed references to those pages in the search engines database. After the pages are indexed any user can use the front end search process to find the pages. The search engine believes the first document listed in the best match for one's search.

The following are the steps followed in webpage:

- In case of connecting a site say *www.excite.com* user will find that the main page has an area called search.
- Search contains an input field for entering keywords of phases one wants to search for. When the customer click on the search button, the search engine uses the parameters, one has entered to find matching references in the info seek database.
- When the user use search using any keywords and click on the search button, one gets a list of result. Typically the result of search is displayed according to their relevance to the search parameters that one has entered.
- Most search engines display references to the top 10 to 20 pages that match one's search parameters. Successive group of matching pages are also available, but the user has to follow a link to another result page.
- Next result button found at the bottom of the results page to see additional pages might match the search. Often the matching pages are described using the page title and a brief description taken from the page itself.

Some search engine lets one customize the search many different ways that one can show or hide summery information, view results by website and a whole lot more.

## 6. Need of Uniform Resource Locator (URL)

In traditional postal communication system when a sender wants to send a letter he or she has to address the receiver who will receive the letter. Similarly, in case of web world a client wants to access documents needs an address. In order to facilities the access of document described throughout the world, HTTP uses the concept of locators. The uniform resource locator (URL) is a standard for specifying any kind of information on the internet. The URL defines four things (i) method (ii) host computer (iii) port (iv) path.

- ‡ Method - The method is the protocol used to retrieve the document.
- ‡ Host - The host is the computer where the information is located. Web pages are usually stored in computers and computer are given different name which usually being with the character www.
- ‡ Port - The URL optionally contains the port number of the server. If the port is included it should be inserted between the host and the path and it should be separated from the host of a colon.
- ‡ Path - It is the path name of the file where the information is located. Path can itself contain slashes that in the UNIX operating system, separate the directories from sub-directories and files.

### Structure of URL

Method: //host:port/path  
 http:www.planetsearch.com/

### Operation of search engine

Search engine works in three spheres.

- ‡ Web crawling (ii) Indexing (iii) Searching

#### (i) Crawling the web

Crawling the web is the most fragile application since it involves interacting with hundreds of thousand of web service in various names. Servers which are all beyond the control of the system Google have first distributed crawling system where a single URL server services lists of URLs to a number of crawlers. Both the URL server and the crawler are implemented in python. Basic features of crawlers are as follows:

Each crawler roughly 300 connections open at one. It is necessary to retrieve web pages at a first enough pace. At peak speeds, the system can crawl over 100 web pages per second using four crawlers.

Each crawler maintains its own DNS cache, so it does not need to do a DNS look up before crawling each document.

Each of the hundreds of connections can be in a number of different states.

(a) Looking of DNS, (b) Connecting of host, (c) Sending request, (d) Receiving response.

Each of the crawlers uses asynchronous IO to manage events and a number of queues to move page fetches from state to state. It by running a crawler it connect more than half a million servers and generates tens of millions log entries and generates a fair amount of e-mail and phone calls. Because of the immense variation in we page and server, it is virtually impossible to test a crawler without running it on larger port of the internet.

#### **(ii) Indexing**

After each document is parsed it is encoded into a number of barrels. Every word is converted into a word ID by using an in-memory hash table the lexicon. New addition to the lexicon hash table are logged to a file. Once the words are converted into word ID, their occurrences in the current document are translated into hit lists and are written into the forward barrels. The main difficulty of the indexing phase is that the lexicon needs to be shared. Instead of sharing the lexicon, one takes the approach of writing a log of all the extra words that were not found in lexicon, which one fixed at 14 million words. That way multiply indexer can run in parallal and then the small log file of extra words can be possessed by one find indexer.

#### **(iii) Searching**

The goal of searching is to provide quality search results efficiently.

**Web controlling** - Web search engines works by storing information about a larger number of web pages, which they retrieve from www themselves. Pages are retrieved by a web crawler that is an automated web browser and follow every link it sees.

The contents of each page are analyzed to determine how it should be indexed.

Data about web pages is stored in an index database for use in later queries. Some search engines, like Google stores all or part of the source page as well as information about the web pages, whereas some store every word of every page as it finds the Alta Vista.

The catch page always holds the actual search text since it is the one that was actually indexed. So it is very helpful if the content of the current page has been updated and search terms are no longer in it.

When a user comes to a particular search engine and makes a quarry by the keywords, the engine looks up the index and provides a list of best matching web pages according to its criteria with

---

short summary containing the documents title and some times part of the text. Most search engine supports the use to Boolean terms AND OR and NOT to further specify the search query.

The credit of search engine is judged on the relevance of the result set, it sends back. Because there are millions of web pages, which include a particular word or phrase, some pages may be more relevant, popular or authoritative than others. So through the different search engine relevant information are ranked to provide the best result. On the basis of techniques of different search engines it decides which pages are the best matches and what order the results should be shown.

#### **Problems faced by the search engines**

The explosion of electric information retrieval and communication is evident in many ways. Where the web provides an enormous amount of information on the other hand, offers an inexhaustible source for searching learning material, but cause an excessive amount of problem in the search results. At the present, search engines are widely used for searching but there are unsolved problems related to their effectiveness.

*Antonius* (2004) states that keyword-based search engine present serious problems related to the quality of search results. It is happened as relevant pages are to index by a traditional search engine and in this case important information can be reached if its specific internet address is known. Semantically similar queries can return different result on searches based on keywords and not to its meaning.

Retrieving a large body of information by using a search engine is a very time-consuming task for the user who has to perform it manually. This is happened because the result of a search engine is a single web page, and to retrieve information it is necessary to perform several queries.

*Guha* (2003) has viewed about the structure of the web, where he has opined that web is composed of a huge pool of documents and links between them. More and more often the web is not used only by the people but software agent communities who are becoming users of the web too.

*Mc Guinness and Harmelen* (2001) have viewed on semantic web, which is to improve the existing web with a semantic layer that allows machine to understand it, or better to enable software programme to process information efficiently. To achieve its aim, the semantic web is based on the relationship between several layers, each of which has a specific role.

*Bary and Paoli* (1997) comment on layer which provides surface syntax for structured documents, but imposed no semantic constrains on the meaning of these documents.

*Lassila and Swick* (1999) view on RDF which provides a data-model for objects and relationships between them, thus providing a simple semantic, the relative RDF.SCHEMA is a vocabulary for describing properties and classes of RDF resources, with a semantic for generalization hierchies of such proportion and classes.



- Due to growth of web major search engine becomes slower to index new page.
- Many web pages are updated frequently which forces the search engine to revisit them periodically.
- The queries are made on limited searching for keywords, which may result in false positives. Better results can be achieved by using proximity search option with a search bracket to limit matches within a paragraph of phrase, rather than matching random words.
- Many dynamically generated web sites are not indexable by search engines, this phenomena is known as invisible web.
- Some search engine output the search result not by relevance of the term but by the amount of money the matching web sites pay.

Some search results are being polluted with linkspam or built-and switch pages which contain little or no information about matching phrases due to tricks of hundred generated web sites to manipulate search engine to display them in the higher results for numerous keywords.

Dynamically generated sites may be slow or difficult to index or may result in excessive results; perhaps generating 500 to more web pages than average.

#### **Success of search engine operation**

The process of search engine has gained popularity in recent year as a means to reach the target audience through improve website. The following tips are some basic best practices to considered full utilization of web site through search engine.

##### **(i) Conceptualizing target search terms**

The first step to search engine optimization is to identify the search item i.e. keywords, referred key terms for which the user want to visited page through search engine. The target term must be at least two words in length and of course, be relevant to the content of site page. The key term search depends one one's own intuition and team brainstorming. In validating one's chose and researching search term possibilities, two popular tools are used. (a) Word tracker ([www.wordtracker.com](http://www.wordtracker.com)) is the standard tool used for search term selection research. (b) Overture ([www.overture.com](http://www.overture.com)) offers term suggestion and can be utilized for researching target search terms.

##### **(ii) Integration of term**

After selection of the search term the next step is to integrate the terms into one's site page to make them relevant. Initially the selected page focuses content relating to the terms that one can optimize as many page, but each page should focus on not more than one or two of your target terms.

---

**(iii) Inserting target terms throughout page copy**

Once the terms have identified, needs to incorporate the terms into the copy of one's site pages so that when an engine arrives at one's page it can identify the page as being relevant for the search term. Most importantly, it tries to integrate search term into one's page copy in a natural fashion so that the terms make sense in context and complement the overall message of the page content. In order to use search engine in more successful way, one has to take care of the following points.

*Make use of head tags* - Search tend to weight word placed within page head tags (h1, h2, h3, h4 etc.) more heavily than do generate text. H1 tag the top of the page ensures that pages get maximum search engine exposure.

*The page title page* - Search engine weight the title of one's page quite heavily in determining relevance.

**(iv) Linking structure and strategy**

One of the primary ways that search engine find site page is by following the links from one's home page to other inside pages (this process is known as spidering or crawling)

Thus effective search engine optimization is not about "*tricking*" the search engine into ranking one's site favorably. Rather search engine optimization is designed to deliver relevant information to the user seeking that information. The better one is to optimize title page content to suit the needs of one's target visitors.

**Search engine differs from directory lists**

Search engine and directory list are different level. When the user looks for information with a search engine, one uses keywords. When one look for information at directory listing one searches by following links to page within the directory site. The best known directory list in Yahoo (www.yahoo.com).

**Yahoo.com**

When one visited the yahoo home page, the user is greeted by a listing of the top level categories of information available at the site. Under the top level categories the more focused categories of information are available.

If one select news and media category, one jumps to different broad categories of news. Here by selecting another link one can get more narrowly focused category such as business news or technology news.

Yahoo! also makes use of search engine. When the LYCOS (www.lycos.com) search engine lets search the web, the yahoo search engine find payers within yahoo! Yahoo's own web sit contain references to the information one is seeking.

Search engine can't find or index the web sit without a little help that is why search engine helps to find user web sit through a link to one of the pages from someone else's site.

After a search engine finds one's web sit, user the links within one's pages to find additional pages at one's site.

So when website one links at a time adds the information from one website into the database. Search engine don't store all the data in one's page. Search engine create indexes references to one's page.

#### Customer based information search through search engine

There is a new approach towards the integration of search engine with information seeking habit. This approach is categorized into five groups (i) Functional needs (ii) Hedonic needs (iii) Innovation needs (iv) Aesthetic needs (v) Sign needs. Its functional needs focuses products knowledge, uncertainty, phenomenology, experimental, sensory and emotional (iii) innovation needs include novelty seeking, variety seeking and creating (iv) Aesthetic needs include imagery and fantasizing, and at last sign needs include symbolic expression and social interaction.

**(i) Functional needs** - Here the information needs of product knowledge are related to constructing a bank of potentially useful information. Then related information search is done through keyword search of classification directory. The flow of work of this structure is as follows:

i	Product knowledge	Construct a bank of potentially useful information	Search engine Focus on completeness of related information with key word search or classification directory
Ii	Uncertainty	Find a strategy of certain risk reduction effort	Search engine focuses on the presentation of the more reliable or credible information
Iiii	Utility	Try to better serve its	Search engine helps to translate intended search goal to related information
iv	Efficiency	It helps to identify useful data at the time of exposure	Search engine focuses on the function of perform in the least wasteful manner

**(ii) Hedonic needs of customer:** The second information needs focuses on verity seeking which are relating to rotating choices among known products and brand in order to relieve low simulation situation. Where the following three strata are used. The HOD of work of this structure as follows:

i	Phenomenology	Which regards consumption as a subjective state of internal cognition	Search engine focuses on the experience seeking of consumer behaviour
ii	Experiential	Provide pleasure benefits which is obtained from prepurchase activity and product usage	Search engine helps on entertainment arts and leisure, which tend to exhibit high levels of internet and involvement among their users.
iii	Sensory	It provides intrinsic gratification through emotions	Search engine focuses on providing feeling good, enjoyment, excitement, happiness and gratification

**(iii) Innovation needs of customer**

The main purpose of this category of information is to develop creativity, which is related to the idea that highly creative consumers are better able to decide whether to adopt a novel product based on a more competent new product evaluation. So, we need to provide more competent new product evaluation models in order to help the consumers. It deals with the following aspects:

i	Novelty	To be quite diverse in the knowledge	Search engine focus on satisfying to what extent individuals can vary their choices
ii	Variety	Bring out rotation among known product in order to relive low simulation situation.	Search engine helps through different indicator in verity of ways

**Need of Skills development**

- ⌘ Knowledge on web browser - That it is a software application, which provides a graphical user interface (GUI) so that the user can navigate the internet easily by clicking on menus, icons or buttons instead of learning different keyword commands. The librarian must have the idea on many web browsers available having different features. However, every browser shares some of the basic features like handling requests for HTML files, interpreting links, saving a collection of pages, connecting to an e-mail programe and so on.
- ⌘ Knowledge on content area - That it is a portion of the window that holds the document page or other resources as the browser present it. Any text, image, animation, links or any other application files is shown in this area.

- (iii) *Knowledge on status bar* - Knowledge on status bar helps to know that it displays the current state of activity of the web pages. Apart from this, other information appears on the status bar includes the size of the web page, percent of the web page that has been downloaded, error in the web page being downloaded and so on.
- (iv) *Knowledge on navigation or surfing the web* - It includes the Knowledge on the activities of looking through information on the Internet by repeatedly scanning and selecting, usually with the help of hyperlinks. Which is known as navigation or surfing the web.
- (v) *Knowledge on Internet explorer* - Knowledge on internet explorer use a feature called favorites to put website on visits afloat within easy reach. That way, a website can be opened by a simple click without the user having to remember or type anything.
- (vi) *Knowledge on Electronic mail* - This Knowledge helps to know as the process of exchanging message electronically, via a communication network, using the computer. E-mail allow user to communicate with each other in less time and at nominal cost as compared to traditional phone or mail services.
- (vii) *Knowledge on standard protocol* - It helps to know the Knowledge on the connection to the Internet and access to e-mail service, which forwards the mail. The standard protocol used for sending Internet e-mail is called Simple Mail Transfer Protocol (SMTP).
- (viii) *Knowledge on netiquettes* - It helps to have Knowledge about rules that should be followed while communicating over the Internet. Disobeying netiquettes while sending mail can create trouble.
- (ix) *Knowledge on e-mail client* (known as mail user agent or MUA). Which is a computer programme that runs on a personal computer and enable the user to organize filter and sort all the sent and recorded/ e-mail. It is called client because server architecture.
- (x) *Knowledge on news group* - It is a virtual place where people with common interests can ask question and get answer on just about any imaginable topic ranging from sports to space.
- (xi) *Knowledge on search engine*- Which is a tool for searchable database of internet files collection by a computer programme called a wanderer, crawler or spider. It allows the user to enter keywords, relating to particular topics and retrieval information about internet sites containing the keywords.
- (xii) *Knowledge on instant messaging* - It is a web based service to exchange messages in real time between two or more people over the internet. It is a combination of e-mail and chat room. Usually, user sends their message in text format. However, while sending instant message, a user can also communicate through web cam [video chatting, as well as through speakers and microphone (voice chatting)].

### Skill on refining the search

The newly employed professional must have the knowledge on the major search engine which allows the user to choose whether to search for the exact typed phrase, all the words in the phrase, any of the words in a phrase, and so on. However, a few search engines do not provide such options. In this case, the user can define the search by adding one or more words or symbols to the search topic.

**AND** : It is used to search for two or more terms on the same page. Type the word AND between the terms for example poverty AND crime or put a plus sign right before the second term (example knowledge + management) .

**OR** : It is use to search for either two (or more) terms on the same page. Type the word OR between the terms for example (digital or electronic environment) .

**NOT** : It is used to search for pages that include the first term not the second. Type the word NOT between the terms (for example- cat NOT dogs) or put minus sign right before the second term for example (cat - dog) .

" " : Double quotes are used to search for an exact phrase.

( ) : Parenthesis are used to group parts of the search phrase for example, type desktop AND (thorns OR wallpapers) . This return page with the words desktop and themes or both the words themes and wallpapers.

### The Plus model for new user in searching Internet

The PLUS information skills model, which is well suited to school use, has been developed in Scotland James Herring. Who is an authority on information literacy based at Queen Margaret University College, Edinburgh.

PLUS is an acronym that both pupils and teachers will find easy to remember. It breaks information skills into four main parts, as shown in the box below.

- P Purpose** : Identifying the purpose of an investigation or assignment.
- L Location** : Finding relevant information sources related to the purpose
- U Use** : Selecting and rejecting information and ideas, reading for information, note-taking and presentation.
- S Self** : How pupils evaluate their performance in applying information skills to the assignment and what evaluate they learn for the future.

**Purpose :**

- Identifying an information need.
- Learning to frame realistic research questions
- Planning a piece of research using diagrams or headings
- Identifying keywords

**Location :**

- Selecting suitable information media.
- Locating information using library catalogues, indexes, databases, CD-ROMs or search engine.

**Use :**

- Evaluating quality/ relevance of information retrieval
- Slimming and scanning text for information
- Taking notes
- Presenting and communicating information
- Writing a bibliography.

**Self-evaluation :**

- Reflecting on what has been learnt and being able to come to a conclusion based on information found.
- Carrying out personal information audits.
- Identifying successful information skills strategies.

**7. Conclusion**

Undoubtedly it is fair to say that Internet based information retrieval would collapse if search engine were not available. Without search engine, searchers would be about as successful negotiating the Internet as someone trying to look up a phone number in an unsorted phonebook. Search engine provides three chief facilities (i) search engine helps to gather information together conceptually, a set of web pages which forms the universe from which a researcher can retrieve information, (ii) search engine represents the pages in the universe in a fashion that attempts to capture their content, (iii) They allow searchers to issue queries, employ information retrieval algorithms that attempt to find most relevant pages from the universe. Search engines differ somewhat from each other along all these dimensions. Thus World Wide Web (www) is a reposting of information spread all over the world and linked together. The www is a unique combination of flexibility, portability and user friendly through the role of search engine. The www today is a distributed client-server services, in which a client using a browser can access a service using a server.

---

**References**

- 1 G. Antoniou, F. van Harmelen. (2004) A semantic web primer. The MIT press chap.no.1
- 2 R. Guha, R. McCool. (2003) TAP: A semantic web platform. *Computer networks: The International journal of computer & telecommunications networking*, 42,5
- 3 D.L. Mc Guinness and F. van Harmelen. Web ontology language. <http://www.w3.org/2001/sw/webont/>.
- 4 T.Bray, J. Paoli, C.M. Sperberg-McQueen. (1997). Extensible markup language, world wide web journal, 2, 4p.29-66
- 5 O.Lassila and R. Swick. Resources description framework (RDF) model and syntax specification. <http://www.w3.org/TR/1999/REC-rdf-syntax-19990222/>.
- 6 Assael, H. (1984). *Consumer behaviour and marketing action*. Boston, Kent.
- 7 Bloch, P.H. Sherrell, D.L. and Ridgway, N.M. (1986). "Consumer search: an extended framework", *Journal of consumer research*, vol.13, Jan, pp.119-26
- 8 Bradlow, E.T. and Schmittlein, D.C. (2000). "The title engines that could: modeling the performance of World Wide Web search engines", *Marketing science*, vol.19 No.1, winter, pp.43-62.
- 9 Cabena, P., Hadjinian, P., Stadler, R. Verchees, J. and Zanasi, A. (1997). *Discovering data mining: From concept to implementation*, IBM, Prentic- Hall, Upper Saddle River, NJ.
10. Chen, L.D., Sakaguchi, T. and Frolick, M.N. (2000). "Data mining methods, application and tools", *Information Systems Management*, vol.17 no.I, Winter, pp.65-70
11. Cox, K.C., Eich, S.G. and Wills, G.J. (1997), "Visual data mining: recognizing telephone calling frauds", *Data mining and knowledge discovery*, vol.1, pp.225-31.
12. Fawcett, T. and Provost, F. (1997), "Adaptive fraud detection", *Data mining and knowledge discovery*, vol.1, pp.291-316.
13. Fayyad, U.M. (1996). "Advances in knowledge discovery: making sense out of data". *IEEE Expert*, vol.11 no.5 October, pp.22-30.
14. Feelders, A., Daniels, H. and Holsheimer, M. (2000) "Methodological and practical aspect of data mining", *Information Management*, vol.37 no.5, August, pp.271-81.
15. Goedon, M. and Pathak, P. (1999). "Finding information on the World Wide Web: the retrieval effectiveness of search engine", *Information Processing and Management*, vol.35, pp.241-80.
16. Greening, D.R. (2000), "Data mining on the web: There's gold in that mountain of data", available at: [www.webtechniques.com/archi/20001011Greening](http://www.webtechniques.com/archi/20001011Greening).



17. Gronov, G.R. (2000) , " The roads and crossroads of internet history", available at: [www.internetvalley.com/intval14.html](http://www.internetvalley.com/intval14.html).
18. Ha, S.H. and Park, S.C. (1998), "Application of data mining tools to hotel mart on the internet for database making", *Expert Systems with Application*, vol.15, pp.1-31.
19. Hui, S.C. and Jha, G. (2000), "Data mining for consumer service support", *Information and Management*, vol.38 no.1, October, pp.1-13.

#### **ABOUT AUTHOR**

**Dr. (Mrs.) Puspanjali Jena**, Senior Reader, P.G. Department of Library & Information Science, she has contributed more than thirty papers to different conferences and Seminars. She possesses twenty years teaching experiences and activity engaged in guiding research works.