

Web Search Strategy/Behavior

By

Meenu Sharma

Librarian

Canadian Institute for International studies

*C-2, Phase-1, Industrial Area,
Mohali*

E-mail: meenusharma982@yahoo.com

ABSTRACT

“A library is a place where the right information is provided at right time, at right place and in a right personal way” by Dr. Ranganathan still holds true with the coming of new technology. The resources are shifting from print to electronic forms and the horizons of information age have expanded considerably than ever before, thanks to Internet and World wide web making the whole world a global village, thereby leading to such huge information maze which is beyond imagination. But, in order to make effective use of this mine of information, it becomes equally important to understand it and devise tools & techniques to utilize it effectively. Besides the analytical skills to identify the needs of users, knowledge of search tools has started becoming a speciality. Where search engines, subject’s guides and meta search engines help identify and locate the information, search logic helps to conceptualize the queries. However, the librarians who are excited to use wonderful web resources are digging out and mastering techniques to understand it’s behavior keeping in mind the fact that the value of a resource, and its ability to satisfy the information requirements, will vary for different audiences. However, as the Internet resources are dynamic and keep on changing, so resource evaluation has becomes more difficult than ever before leading to non-credibility of the results. This magnificent tool for librarians and informational professionals if properly understood and utilized can drastically change the frame of the today’s libraries and librarians building new relationships with the users, satisfying their information needs in a way they could never have anticipated.

KEY WORD : Search Engine, Meta Search Engine, Search tools, World Wide Web.

0. INTRODUCTION

World wide web is the most popular term among the librarians as well as information professionals these days for retrieval & dissemination of information, but somehow it is so weird that there are no particular guidelines available to understand and search effectively this universe of information paradigm.

Developed by Tim Berners Lee, World wide web came into being in 1991 and continues to grow every day, almost tripling in size over the past two years. In order to make effective use of this storehouse of knowledge, let’s first realize that the web is a very friendly system of finding information. It is organized into web pages identifiable by their URL’s (Universal Resource Locator) and can be viewed through browsers like Internet Explorer or Netscape navigator. But in spite of all this, the question arises as to what to do if we want to go to a particular web site whose’s address is lost or is not known?

The answer lies in understanding the web behavior and searching effectively.

1. SEARCHING THE WEB

Talking of searching effectively, some strategies are to be applied that include: -

- A. Identifying main concepts of the topic
- B. Search logic: the manner in which the terms would be combined
- C. Information of search tools
- D. Evaluation of the retrieved results

A. Identifying main concepts of the topic

The first step and obviously an important one is to identify the kind of information need whether it is business, education or society related, as the context of the information has to be conveyed in a logical manner to WWW. Secondly, identification of the keywords is important. Suppose we wish to know about "Role of computers in libraries", then the keywords become role, computers, and libraries eliminating the prepositions. The kinds of spellings for our query or keywords are also important, as they are many variant forms of the same spellings in the same language like colour and color, labour and labor etc.

On the Internet as postings are by different people using different vocabularies, various kinds of synonyms are also used for a single term and the irony is that the search tool is only going to look for terms that would be specified without really helping out what we want. Like for a word "disease", it may search for web pages using words ailment, infection, illness or sickness. So there is need to practice with different search words we choose to look, as there are different ways of expressions of idea on the same topic.

B. Search logic

The golden rule of searching on Internet is to first understand the relationship between different pieces of information.

Use of search tools can be improved if we know the logic to retrieve best results from various search engines. The question or query must be changed into a particular search statement. The popular strategies, which are being used to do searching from various tools, are as follows: -.

• **Boolean Logic or Operators:** - These are the most popular & useful guides to construct search queries. They include AND, OR, NOT and proximity operator NEAR.

In case, we wish to expand our search, then we have to use OR operator like Libraries OR Museums providing links to all those sites, which have either of these two words.

In order to narrow the search, we use the AND operator. Libraries AND Museums retrieve pages, which have both of these words present, and using the operator NOT will exclude items from a search. For example, Libraries NOT Museums will retrieve results, which do not have museums.

Another operator NEAR ensures that keywords are found within 10 words of one another eg. Rural development NEAR Bihar. It will retrieve only pages where rural development is found near the keyword Bihar. It is one of the most powerful operators for narrowing context-based searches.

- **Truncation (use of wildcards):** Are special symbols, which allow to search the terms arising from the same root. For example, comput* will retrieve results computers, computing, computing and computerization. They are especially useful for retrieving variant spellings (like colour/color) and words with common root (psychological, physiology, physiologists).

- **Phrase searching:** searches for a particular phrase like “role of computers in Libraries”: will retrieve results talking only about role of computers in libraries instead of looking for keywords- role, computers and libraries separately.

- **Case Sensitive Searching:** It allows searching for words in the case, the specified letter is entered and retrieve results when that word appears exactly as typed e.g. AltaVista is case sensitive, if you type Fruits, it will only retrieve results with keyword Fruits but not fruits.

A search command are never universal and differs from one search tool to another e.g. goggle does not support OR operator.

C. Search Tools

Search tools are basically the sites on the Internet, which help to find or locate the information on WWW. These include search engines, metasearch engines, subject directories etc. Each type of search tool works in different manner and has its own special ways of displaying and ranking results.

- **Search Engines:**

The most popular search tools are the search engines These generally index terms or words that are believed to be significant or at the top of the document like on the title or header or repeated several times throughout the document. Usually a search engine makes use of special software robots called as spiders, which crawl through the whole World Wide Web and index the list of the words found on websites. The most popular examples are google, excite, altavista, hotbot etc. These search engines differ in size of their database (as to how many web pages they index), frequency of updating their database and criteria of looking for keywords and displaying results Search engines give basically two types of searching facilities: - Basic and Advanced.

Basic involves the input of keywords or words with + (plus) and – (minus) combinations, double quotes (phrase searching) and wildcard searching (Truncation).

Advanced searching helps the use of Boolean operators, phrase searching, field searching and truncation. It also allows using the date and language features just like in Goggle and also helps to control the ranking. The advanced feature options are different with different search engines like Excite has a “zoom in” feature, which allows refining the subject before we search. AltaVista allows searching within a certain date range by entering the beginning and ending dates. Search engines also have help files (providing

information on how the system works), searching tips and FAQ's (Frequently asked questions). Another useful feature found on various search engines is "similar pages option". These are additional web pages that are similar to the ones displayed e.g. while searching information on "librarians" a web page on librarian association can be found in "similar pages" option which can be a good point to start research on librarians.

• **Subject specific search Engines**

These are specialized search engines, which focus on a particular subject, or area or particular reference. These engines thrive on other general search engines for their basic data and filter the information according to subject specific contexts resulting in fewer and appropriate hits. These are helpful where the intention is basically comprehensive study of particular area or field rather than overall coverage.e.g. AltaVista Canada (regional Canada), monster.ca (jobs with reference to Canada), Medline plus (Medicine). The time in which the results are retrieved, number of retrieved meaningful records and the description of the retrieved results along with relevancy rankings differs in different subject specific search engines.

• **Metasearch engines**

Multithreaded engines search several engines at once These engines do not crawl the web themselves but they just act a middle agent by passing the query to many other search engines and retrieving the best matches. For example: metacrawler searches 7 search engines including Lycos, Info seek, WebCrawler, Excite, AltaVista, and Yahoo and uses Boolean logic; phrase searching etc. Metasearch engines do not have a database of their own but they pass the search queries to several other search engines all at once, thereby displaying the results by ranking intelligently. The advantage of metasearch engines over others is that they handle simple searches quite efficiently and save lot of time to go to individual search engines. It is quite helpful to those people who are looking for an overview study on a subject or something unique.

There are certain disadvantages to these metasearch engines. Besides being slow, these only spend a little time on each search engine and thereby it may be possible that they only retrieve about 10-20% of the results which each search engine shows.

• **Subject Directories**

Imagine you want to know about a particular subject Let's say trip to Goa. Naturally you would like to know about the information of holidaying in Goa as a part of tours and travels. Obviously, a search engine would not be the best bet, so where to go? Well, here come the Subject directories.

They will browse through list of websites by subject in search of related information. They are comparatively smaller than search engines and most of the times provide links to the homepages of the retrieved results. The information is generally compiled by someone and may include a search engine for searching. Yahoo is the largest and the most popular subject directory. So, if you are clear about the topic or main theme of your query, the best way is to go to the directory rather than a search engine.

They're also a helpful way for exploring a topic and learning about related topics and also provide the platform to join subject related online communities for exchange of ideas.

D. Evaluation of the retrieved results

Now, having received the results on screens, it is a big confusion how to determine as to which is the most reliable and useful

information, as all the information seems to be correct. World wide web is a place where anybody can publish anything just by having a computer and Internet connection without any editorial directions. Just because the huge information is so easily and speedily available it leaves us all mesmerized. Before going on to the conclusions, critical thinking must be applied to evaluate the authenticity of the web pages retrieved. Evaluation of web resources is based to varying degrees on the criteria related to evaluating the source, content, and the format

The main questions which are to be asked to ourselves while evaluating the information is whether the link or the results or the information found in it matches our purpose or not. Is the information up to date, relevant and focused? Being a global system, it may contain information generally applicable but not geographically acceptable.

Before believing on the information retrieved, it's also important to determine the accuracy and reliability by assessing the authorship, updation, reference & sponsoring agency. However sometimes it can be difficult to determine the sites providing serious information, and sites selling a product or just entertaining. Some search engines, like Yahoo, will sort search information into categories, separating business and entertainment sites from other types of information. It is then up to the researcher to determine which categories to examine but it cannot be true in all the cases. However, check for any bias evidence, which can harm the credibility of the information. Like whether the information is inclined towards a particular political, religious, ethnic or sexual group?

Last, but not the least, it's important to acknowledge that the links provided by the search tools should be up to date. Sometimes the search results may not exist any more. Normally when a web site is removed from www, the address is not automatically deleted from search engine.

Internet resources are dynamic and keep on changing & updating, so resource evaluation becomes more difficult than might be the case

2. CONCLUSION

World wide web is much wider than these search tools open it to us. It keeps on changing and the irony is, we cannot even entrér the whole net and can only access some fraction at a given time. But not to worry, search tools help us to give a way out when caught in this web in spite of the fact that none of it is comprehensive. It is very interesting to watch this dynamic web behavior and with new and new technologies coming up, the time is not far when the proper classification systems would be used to index WWW to make it more user friendly

Happy Surfing!

REFERENCES

Print Sources: -

1. Light on Internet: essentials of Internet & World Wide Web, Wendy G. Lehnert, 1999
2. Internet in a Nutshell: a desktop complete reference, Valerie Quercia, 1998

3. **Microsoft Internet Exp**
4. **lorer 4, Tim Duffey, 1999**

Web resources: -

1. **Librarian's index to Internet: -<http://www.lii.org>**
2. **<http://www.yahoo.com>**
3. **<http://www.lycos.com>**
4. **<http://www.excite.com>**
5. **<http://www.google.com>**



***Ms Meenu Sharma* is the Librarian in Canadian Institute for International studies, Mohali. She holds B Sc in Medical and is the gold medallist in B L I Sc and M L I Sc from Panjab University. She has attended the Training Program at Georgian College, Ontario, Canada.**