

SEARCH ENGINES: TOOLS FOR LIBRARY

Balwan Singh

Assistant Librarian
CCS Haryana Agricultural University
Hisar-125004

P. C. Kapila

Assistant Librarian
CCS Haryana Agricultural University
Hisar-125004

Discusses search engines giving definitions taken from different sources like dictionaries, etc. Explains the working of search engines and their effective use in searching information on Internet on the basis of boolean operators— AND, OR, NOT and NEAR. Highlights the role of search engines in libraries in locating pin pointed information. Suggests that library staff should have sufficient knowledge of search engines to make maximum use of Internet. Lists some of the general and specialised national and international search engines, which are frequently used to access information on Internet.

INTRODUCTION

Information in the present day context has become one of the most important necessities of man. Every one today wants to get pin-pointed information at short notice. For this, amongst many sources of information, Internet is one of the options from where anyone with knowledge on how to use the Internet, can retrieve the required information. For retrieving information quickly from Internet, one needs to know from where the required information can be obtained, as the Internet in itself is vast. A user of the Internet can retrieve the information in two ways: either directly from the web site if the URL is known or through search engines. To use search engines one should know the web address of the search engine. It is very important that when a user wants to search information on a specific field, search engines usually search by keywords and they build the index on the keywords found in the databases or on the World Wide Web.

DEFINITIONS

According to Computing Dictionary [1], "Search engine is a program that allows users to locate

specified information from a database or mass of data. Search engine sites are extremely popular on the World Wide Web because they allow users to quickly sift through millions of documents on the Internet."

Dictionary of Computing and Digital Media [2] defines search engine as, "A database front that allows a user to seek information on the Internet by keyword. Search engines may look for titles of documents, URLs, headers, or text."

In Que's Computer & Internet Dictionary [3], search engine is defined as, "Any program that locates needed information in a database, but especially an Internet-accessible search service that enables you to search for information on the Internet."

According to Beiser [4], "A search engine enables a user of electronic data resources to quickly locate the specific information desired from within a large volume of mostly unrelated extraneous information."

WORKING OF SEARCH ENGINES

Every search engine creates its own copy of the World Wide Web called an index. The size of a search engine's index varies from search engine to search engine, but it is always smaller than the World Wide Web as a whole. For example, it is estimated that the Web currently consists of about one billion pages, whereas the largest search engine index is currently about 250 million pages.

The search engine builds a list of pages to add to its index using a special piece of software known as a spider. The spider crawls across

the Web, adding pages it visits to the list of pages to index. The spider is capable of reading a Web page and finding links to other pages to visit. In this way, the spider can travel across the Web finding pages to add to the search engine index.

Some time after a page has been "Spidered" (visited by a spider), search engine software adds a copy of the page to the search engine index. When a user enters a query into a search engine, search engine software searches the index to find pages that match the query. It then sorts those pages into ranking order. Each search engine uses its own "secret recipe" to find and rank pages, but most base their recipe on the frequency and location of the search term on the page.

EFFECTIVE USE OF SEARCH ENGINES

It is important to choose right key words for effective searching. Similarly, it is essential for Web searching to have right key words, as there is a lot of information and out of it some may be useful and much of it useless. No single body is responsible for arranging or organizing the information. In case proper keyword selection is made, it means that half the battle is won. Thus one should know how to enter, combine keywords and success rate is improved if we learn about these tools and their uses. Of course procedures vary from one search engine to another. Some search engines may have different set of instructions depending upon use of their version such as simple or advanced. However, certain general concepts are applicable in using search engines. If we have basic understanding of the tools then any search engine, which we encounter on Web can be used for searching.

Natural Language Search

It is better for new searcher on the web to use plain keyword search. In this way most of the hits will appear at or near the top of the list. One may be able to get the desired material on the topic being searched for.

It is difficult to get precise information by using natural language search or by using multiple words and phrases. For this, the below mentioned means to search specific keywords can be followed:

Boolean Search

Boolean operators AND, OR, NOT and NEAR can be used to make the search in the following ways:

AND Searches

For specific searching for two or more words, which need to appear in the results, AND may be used. This can be used in many ways:

Using Plus Sign

The simplest way to use AND in a search is to put plus (+) sign in front of each word or phrase that must appear in the results. For example, to get material on classification and cataloguing the search phrase may be formulated as: *classification + cataloguing*.

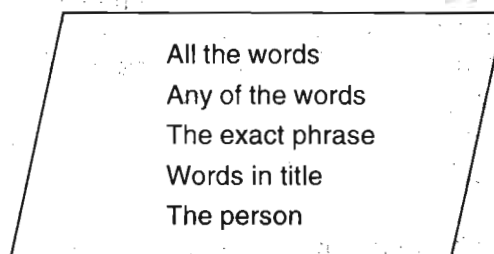
Using 'AND'

Some of the search engines actually use boolean operator AND to combine two words – 'production' AND 'supply'. But all the search engines do not use caps for 'AND'. However it is better to use caps to avoid any confusion.

Menu Option or Radio Button

These are used by some of the search engines to make the AND searching extremely easy. By

Menu Option



following ways, menu option or radio button can be used in search engines:

Radio Button

- ☐ An exact phrase
- ☐ Matches on all words (AND)
- ☐ Matches on any words (OR)
- ☐ A person's name

OR Searches

OR search results in greater hits. In case several words with space in between them are given, the search engine automatically searches using the words separately. The other way is to type 'OR' between the words or phrases. It is better to use caps for the word 'OR'. OR search can also be menu or button based as explained above in case of AND search.

NOT Searches

NOT searches are used to eliminate keywords/subjects from searches. For example, to have information 'Cancer' and exclude 'Tumour', putting minus (-) sign in front of the word that is to be avoided such as *Cancer-Tumour* displays hits containing cancer but not tumor. No space is given between minus sign and the word to be excluded. Some search engines use the word NOT, like 'Cancer' NOT 'Tumour'. Sometimes using NOT search may exclude good materials of interest. So it should be very carefully used.

NEAR Searches

It is at times required to specify the NEAR search where some of the words appear in close proximity to one another. Alta Vista and Lycos offer NEAR searching in part of their advanced search capability but not on their basic search forms. The format is same for both: 'Vajpayee' NEAR "Disinvestments Policy". In Alta Vista search is made within 10 words of each other. Lycos makes within 25 words: *Vajpayee* NEAR/25 "Disinvestments Policy".

Complex Queries

By grouping Boolean operators (AND, OR, NOT and NEAR), search engines also allow to create

complex queries. Statements can be given using parenthesis such as '*Bandra*' NEAR ('*Mumbai*' OR '*Bombay*'). This search is called 'Nested Search' because one statement is "nested" within another. It is however suggested to use brackets in the searches when need be. Professionals always prefer to use nested searches to get better results. Combination of AND, OR, NOT or NEAR can be decided on the search to be made.

Wild Cards

The wild card search is made by asterisk, i.e., to look for variation on a particular word. In case, search is to be made for variation on a particular word like *agric** to find references to 'agricultural', 'agriculture' etc., it will be better to use wild card along with at least one unique keyword such as *Pest* NEAR *agric**. In case the keyword is not used, the search may take long time and provide large list of hits.

There may not be a provision of wild card search in some of the search engines. But variation is automatically managed by the system. This type of search is called 'concept based'. This way similar words or concepts are displayed in the results.

Stop Words

Search engines ignore some words, which are called stop words. Examples of stop words are, 'the', 'to', 'with', 'from', 'of', 'that', 'who', and all boolean operators- 'AND', 'OR', 'NOT', 'NEAR', etc.

ROLE OF SEARCH ENGINES IN LIBRARIES

In the last one decade, Internet is extensively used to access the information because it is a great source providing quick information on any subject. Libraries find it a very useful tool to achieve their objectives in presenting information without losing much time. Even if the libraries may not subscribe journals, which are available on electronic media, they can just get Web connection for these journals and users can access when the need be. Printed journals take long time to reach the library but on the Web as soon as the journal is out, it can be

accessed on the Internet. Moreover, information can be accessed any time, anywhere. The first and the foremost requisite is that one should have Internet connection.

In brief it can be stated that search engines are useful in following ways:

- a) They are ready reference source for current information
- b) Retrieve required information in minimum possible time
- c) Directing towards the location of resources other than searching for, and
- d) Acquaint the users with the latest trends on a topic or subject worldwide.

POPULAR SEARCH ENGINES

Without the address of search engines it will be difficult for a person to access information. To save time it is very essential to have exact address and tool to search the desired information. Moreover it is also helpful for the librarians to keep the addresses of these search engines for providing effective services. Given in following sections are some of the frequently used search engines.

INTERNATIONAL SEARCH ENGINES

Google (<http://www.google.com>)

It is extensively used search engine for finding any type of information required by the librarians for their day-to-day service to the users. Perhaps it is the largest international database.

Yahoo (<http://www.yahoo.com>)

Yahoo is not exactly a search engine as is perceived by users. Launched in 1994, it is web's oldest "directory," a place where human editors organize web sites into categories.

MSN Search (<http://search.msn.com>)

Microsoft launched this Web in 1994 with crawling listing but in 1999 they replaced the crawler-based search with FAST. With the introduction of FAST, the searchers may reach

the last page of the search very quickly. It provides comprehensive search on any topic and makes it widely used search engine in the world.

LookSmart (<http://www.looksmart.com>)

LookSmart is not for public use but provides its results to other search engines. It was launched independently in October 1996, was backed by Reader's Digest for about an year, and then company executives bought back control of the service.

AltaVista (<http://www.altavista.com>)

AltaVista is the oldest crawler-based search engine on the web. It opened in December 1995 and for several years was the "Google" of its day, in terms of providing relevant results and having a loyal group of users that loved the service as improvements have been made, for looking especially good image search. It also has an outstanding news search service.

HotBot (<http://www.hotbot.com>)

HotBot was started in May 1996 and gained a strong following among serious searchers for the quality and comprehensiveness of its crawler-based results, which were provided by Inktomi, at the time.

INDIAN SEARCH ENGINES

Yahoo-India (<http://yahoo-india.com>)

It offers local contents concerning India such as finance, sports, entertainment, news, etc.

India (<http://123india.com>)

Being India's Premier Search Engine:, it provides links to arts, business, economy, books, magazines, computers, news, education, entertainment, food, government, health, medicine, humor, religion, science, technology, sports and tourism.

Appu (<http://www.appu.com>)

It is a comprehensive Indian search engine, which includes information on entertainment, careers, news, sports, government and politics, etc.

Dhundo (<http://www.dhundo.com>)

This search engine includes information on arts and culture, artists, placement, art galleries, business and economy, companies, jobs, finance, real estate, and computers.

Sholey (<http://sholey.com/home.php>)

It is a very good search engine for providing information on finance, real estate, travel, internet, business, health, jobs, entertainment, shopping, etc.

SPECIALISED SEARCH ENGINES

There are number of search engines on every subject/ field. In the sections below are listed some of the important and widely used search engines on Science, Medical, Agriculture, Engineering and Library and Information Science.

SCIENCE SEARCH ENGINES

Scirus (<http://www.scirus.com>)

It is a most comprehensive science specific search engine. With the help of latest search engine technology it searches 150 million science specific pages. It covers scientific, technical and medical data.

BioMedNet (<http://research.bmn.com>)

This search engine is managed and controlled by Elsevier Limited and covers the subject areas such as agronomy, crop and soil science, horticulture, plant science, cell biology, drug discovery, genetics, endocrinology, ecology, evolution, environment, infectious diseases, neuroscience, pharmacology, etc.

MEDICAL SEARCH ENGINES

US National Library of Medicine (<http://www.nlm.nih.gov>)

This is very useful and exhaustive search engine for medical professionals all over the world. Further, MEDLINE, PubMed and MEDLINEplus databases have been created by them for searching medical information.

MedicineNet (<http://www.medicinenet.com/script/main/hp.asp>)

It provides easy-to-read, in-depth, authoritative medical information for consumers via its robust, user-friendly, interactive web site. Since 1996 it has been providing most comprehensive information in the fields of medicine and healthcare.

AGRICULTURE SEARCH ENGINES

AgriNet (<http://agrinet.tamu.edu>)

It is the source of all agricultural information which is created by Texas Agricultural Market Research Center, Department of Agricultural Economics, Texas A&M University, USA.

AgNet (<http://agnet.com.au>)

This links to comprehensive lists of Australian agricultural sites. These lists are compiled and maintained by AgNet as free service to Australian agricultural community.

AgFind (<http://www.agfind.com>)

It is a powerful search engine for agricultural industry, which provides right information very fast. It is useful for agricultural marketing in USA.

ENGINEERING SEARCH ENGINES

ENGINEERING.com (<http://engineering.com>)

It covers the resources on mechanical, architectural, civil, electrical, computer science, nuclear, aerospace, mining engineering, etc. It serves the needs to access engineering data, software, tables, job information etc.

Athenus (<http://www.athenus.com>)

This search engine makes searches on all the branches of engineering. On the request of users new features are added from time to time.

LIBRARY SEARCH ENGINES

Refdesk (<http://www.refdesk.com>)

Since 1995 it has been providing indexing quality of Internet Sites and assisting the users about their navigation.

Librarian's Index to the Internet (<http://lii.org>)

It searches annotated subject directory of more than 12000 Internet resources selected and evaluated by librarians for their usefulness to users of public libraries. It is reliable and efficient guide to Internet resources. It has a free subscription of 15000 subscribers covering 85 countries and functioning since 1990.

CONCLUSION

Today, Internet is considered the biggest source of information and finds an important place in libraries as quickest means to access information at any time. But it requires the help of search engines for the effective and optimum use. Search engine is an aid to find pin pointed

information to save time and energy. Libraries can only take the maximum out of it if their staff have adequate knowledge of search engines. This is very much needed in the present day context.

REFERENCES

1. OVICE (P. C). Computing dictionary. 4th ed.1999. Sand hills, Lincoln, NE. p 247-48.
2. HANSEN (Brad). Dictionary of computing & digital media: Terms & acronyms. 2000.BPB Publications; New Delhi. p 276.
3. PFAFFENBERGER (Bryan). Que's computer & Internet dictionary. 6th ed.1996. Prentice-Hall of India; New Delhi. p 465.
4. BEISER (Karl). Search engines for library websites. *Library technology Reports*. 36, 3; 2000; 2-60.