

वर्षम्-०५ अङ्क:-०५
जनवरी-दिसम्बर, २०२३

शोधज्योति:
ISSN-2350-0700

Search Engines: An Introduction

- Ravindra S Kale^१

- Dr. Rajeshkumar Gamit^२

Abstract:

Search engines are very much useful as a tool in the library. There are various kind and subject specific search engines available on the www web. The various information available on the web, it's an index of the information available in various locations can helps user and staff to locate the information available on the web. The researcher here provided basic information regarding the search engine, its meaning, concept, how it works, Boolean search, definition, types of search engine, list of search engine and there special feature, Meta search engine etc included in the paper.

Key Words: Search engine; Types of Search engine; Boolean Search;

1. Meaning and concept of Search Engine:

The Internet is a vast digital library. The role of a librarian in the internet era is to provide or assist the user in getting the right information at the right time in the right amount from plenty of information available on the net. A web search engine is an interactive tool to help users locate information available via the internet. Web search engines are databases that contain references to thousands of resources. Users interact with the databases putting a

^१ Librarian, Shree Somnath Sanskrit University, Veraval

^२ Assistant Professor, Department of Libray and Information Soleuce, Gujarat University, Ahmedabad.

query that ask the database if it contains resources that match specific criteria. A web search engine provides an interface between the user and the database. The interface presents the user with an opportunity to type in a search string, which may be a word, a phrase, a date, or some other criterion, and a way to submit the request. (Kale, 2009)

The web houses millions of websites, billions of web pages and tons of data in different data formats running on different technology, platforms, and it continues to grow exponentially. Apparently, it is difficult to remember the precise URLs of websites. A search engine is a computer program that searches for documents containing keywords or phrases of interest to users. The search engine itself is virtually a powerful workstation-class machine that searches a database of information collected from the Internet. Primarily software programs called robots or spiders that crawl through all the files on the Internet, and download Open Access them into a searchable database to gather this information. These work as indexes to the literature available on the Net. (HUSSAIN, 2006)

2. Definition of Search Engines:

Editors of the American Heritage Dictionaries (2022) defines search engines as : “A software program that searches a database and gathers and reports information that contains or is related to specified terms. A website whose primary function is providing a search engine for gathering and reporting information available on the Internet or a portion of the Internet”.

According to Britannica Concise Encyclopedia (Wolff, 2006) define a search Engine is: “A tool for finding information, especially on the Internet or World Wide Web. Search engines are essentially massive databases that cover wide swaths of the Internet. Most consist of three parts: at least one program, called a spider, crawler, or bot, which "crawls" through the Internet gathering information; a

database, which stores the gathered information; and a search tool, with which users search through the database by typing in keywords describing the information desired (usually at a Web site dedicated to the search engine). Increasingly, metasearch engines, which search a subset (usually 10 or so) of the huge number of search engines and then compile and index the results, are being used”.

According to the Dictionary of Computer and Internet Terms (Rigdon, 2016) “A program that searches for keywords in documents, databases, Internet content or other files.” and “A website directory which indexes as many websites as it can and allows you to search its database for sites on particular subjects. Because of the immense size and rate of growth of the world wide web, no one has a complete directory of all sites which exist. The most popular is- Google, so popular that it’s name has become a verb.”

3. Types of Search Engines:

Search Engines are of different kinds, classified based on their type, and some of them are briefly explained below:

a. Free Text Search Engine:

This search engine permits Internet users to search the net simply by using a single word or several words and phrases as well. There is no limit in search; searchers will be free to search according to their interest. They can search for a name of an organization, a quotation, an individual’s address, etc. Examples of Free text search engines are:

Euro seek www.euroseek.com

Look Smart www.looksmart.com

Excite www.excite.com

Northern Light www.northernlight.com

Search Europe www.searcheurope.com

Surf Storm www.surfstrom.com

b. Index-based Search Engines:

These search engines are giving more emphasis on classifying information under a series of major subject headings and then subdividing these into a tree structure of more specific headings. Searchers will be able to use this without having a deeper know-how in the subject. The major drawback of these search engines is that they only index a very small percentage of the published websites and they may not be arranged sensibly with regards to the hierarchy.

Examples of Index based search engines are:

Excite www.excite.com

About www.about.com

Azoos www.azoos.com

Brain Fox www.brainfox.com

Facite www.facite.com

Galaxy www.galaxy.com

Go Hip www.gohip.com

No Search www.nosearch.com

Search King www.searchking.com

Yahoo Search www.yahoosearch.com

c. Natural Language Search Engines:

They take the natural query and run searches on these words and other related words. Instead of using any type of search technique, the user can simply type their queries by using normal language. for example, what is the old name of Mumbai? Who is the President of India? Examples of Natural language search engines are:

Ask Jeeves www.askjeeves.com

C4 www.c4.com

Find Same www.findsame.com

d. Directory-based Search Engines:

These engines are also known as subject collections or as “subject gateways” and are created and maintained by human editors, subject experts, or using some automated indexing process. But they rely mainly on people to identify and group resources. These directories offer access to the information that has been classified into specific categories. The group Internet Websites into categories such as Internet tutorials, universities, museums etc. Yahoos, SOSIG, Look smart, Magellan, Open Directory Project. Snap are an example of such a directory.

e. Meta - Search Engines:

Meta-Search engine also known as Meta-Crawler or Multi-threaded Engine is a search tool that sends search queries simultaneously to several search engines, web directories and sometimes to the invisible (deep) web, which is a collection of online information not indexed by traditional search engines. After collecting the results, Meta-search engines remove the duplicate links and, according to their algorithm, combine/ rank the results into a single merged list. Metasearch engines do not have their databases and do not accept URL submissions. They are not doing any search but they simply display the list of other search engines. i. Single List: Most meta-search engines display multiple engines each resulting in a single merged list from which duplicate entire have been sorted out. ii. Multiple lists: Some meta-search engines do not collate multiple engine search results but display them instead, in separate lists as they are retrieved from each engine. The functionality of these

engines depends largely upon the performance of those engines that create and manage their databases. (Shailendra & Anas, 2006) They are attractive because users do not have to Visit multiple search engines and re-enter their query Meta search engines are useful in obtaining a quick overview of a subject or a unique term. Examples for Meta search engines are:

Meta Crawler www.metacrawler.com

Dogpile www.dogpile.com

Intelliseek www.intelliseek.com

Mamma www.mamma.com

f. Subject-specific Search Engines:

Emphasize a certain subject, geographical area, region etc. as such special search engines provide limited search pages excluding irrelevant information from the result of the particular search. There are many specialized search engines available for many different categories, for example;

Health Care www.achoo.com

Science and Technology www.scrius.com

Law www.findlaw.com

Law www.lawguru.com

Science www.sciseek.com

Biotechnology www.bioweb.com

Management www.Grokker.com

4. How do search engines work?

Search engines use automated software programs known as spiders or bots to survey the Web and build their databases. Web

documents are retrieved by search engine index. When we enter a query at a search engine site, our input is checked against the search engine's index of all the web pages it has analyzed. The best URLs are then returned to you as hits, ranked in order with the best results at the top. Most sites offer two different types of searches--"basic" and "refined" or "advanced." In a "basic" search, you just enter a keyword without sifting through any pulldown menus of additional options. Depending on the engine, though, "basic" searches can be quite complex. Advanced search refining options differ from one search engine to another, but some of the possibilities include the ability to search on more than one word, to give more weight to one search term than you give to another, and to exclude words that might be likely to muddy the results. You might also be able to search for proper names, phrases, and on words that are found within a certain proximity to other search terms. Some search engines also allow us to specify what form you'd like your results to appear in, and whether we wish to restrict our search to certain fields on the internet (i.e., Usenet or the Web) or specific parts of Web documents (i.e., the title or URL). Many, but not all search engines allow us to use Boolean operators to refine your search. These are the logical terms AND, OR, NOT, and the proximal locators, NEAR and FOLLOWED BY.

Boolean AND means that all the terms we specify must appear in the documents, i.e., "heart" AND "attack." We could use this to exclude common hits that would be irrelevant to our query.

Boolean OR means that at least one of the terms we specify must appear in the documents, i.e., bronchitis, acute OR chronic. You might use this if you didn't want to rule out too much.

Boolean NOT means that at least one of the terms we specify must not appear in the documents. You might use this if we anticipated

results that would be totally off base, i.e., nirvana AND Buddhism, NOT Cobain.

Not quite Boolean + and - Some search engines use the characters + and - instead of Boolean operators to include and exclude terms.

NEAR means that the terms we enter should be within a certain number of words of each other.

FOLLOWED BY means that one term must directly follow the other.

ADJ, for adjacent, serves the same function. A search engine that will allow us to search on phrases uses, essentially, the same method (i.e., determining adjacency of keywords).

Phrases: The ability to query phrases is very important in a search engine. Those that allow it usually requires that we enclose the phrase in quotation marks, i.e., "space the final frontier."

Capitalization: This is essential for searching for proper names of people, companies, or products. Unfortunately, many words in English are used both as proper and common nouns--Bill, bill, Gates, gates, Oracle, oracle, Lotus, lotus, Digital, digital--the list is endless. All search engines have different methods of refining queries. The best way to learn them is to read the help files on the search engine sites and practice! (Barlow, 2004)

5. Major Search Engines:

There are so many search engines available on the Net few of them are mentioned below:

1. **Amfibi** (<http://www.amfibi.com/>) : General search engine with an uncluttered interface in a choice of Catalan, English, French, and Spanish. Also has a web directory using the Open Directory.

2. **ChaCha** (<http://www.chacha.com/>) : This search engine is a fusion of computer technology and human intelligence. Live human guides are available to provide additional assistance in finding relevant results.
3. **Cuil** (<http://www.cuil.com/>) : Search engine with results shown with images and a drill-down menu. General feature, webmaster, and investor information.
4. **Entireweb** (<http://www.entireweb.com/>) : The English language version of a Swedish search engine with a clean interface and large database.
5. **FactBites.com** (<http://www.factbites.com/>) : A cross between a search engine and an encyclopedia. The results return complete, informative sentences about the search topic. Related topics are suggested.
6. **Hakia** (<http://www.hakia.com/>) : Ontological semantic and natural language processing (NLP) based search engine.
7. **Lexxe** (<http://www.lexxe.com/>) : Processes natural language queries and delivers results in clusters by topic. Queries can be keywords, phrases or short questions.
8. **Live Search** (<http://www.live.com/>) : Offers general web search and various verticals, including images, news, classifieds, maps and local by Microsoft.
9. **Mahalo** (<http://www.mahalo.com/>) : Human-compiled and edited search engine.
10. **Mojeek** (<http://www.mojeek.com/>) : Search engine with a simple interface and no advertising. Offers the ability to create a customized topical or personal search index.

11. **Mozdex** (<http://www.mozdex.com/>) : Uses open-source search technologies and an index seeded from the Open Directory.
12. **Ms Dewey** (<http://www.msdevey.com/>) : Search engine with a touch of humour that plays prerecorded movie clips of a female actress to entertain you while searches are being performed.
13. **Objects Search** (<http://www.objectssearch.com/>) : An uncluttered interface offers options to search the web, news and blogs, including cached pages and 'anchors', which are titles and descriptions from the Open Directory.
14. **Scrub The Web** (<http://www.scrubtheweb.com/>) : Search engine with sponsored links at the top of results. They are also offering a meta tag builder and analyzer.
15. **Search Hippo** (<http://www.searchhippo.com/>) : Indexes sites listed in the major internet directories. Offers free web services, including a search toolbar, and code for webmasters to provide search boxes on their sites.
16. **Searchles** (<http://www.searchles.com/>) : Offers social search features (allowing others' opinions to influence your results) and allows users to tag, group and save links.
17. **Ulysseek** (<http://www.ulyseek.com/>) : A straightforward search engine with a simple interface.
18. **Web Gaps** (<http://www.webgaps.com/>) : User-rated results, provides filters for web, news, feeds, people and includes shared website suggestions.
19. **Wotbox** (<http://www.wotbox.com/>) : Search engine with country-specific searching. Options to display country flags, and include preview and translation features. Sponsored links appear in panels separate from the main results.

20. **About** (www.about.com) : Advertising itself as “ The network of sites led by an expert guide,” about.com is an online neighbourhood of hundreds of helpful experts, eager to share their wealth of knowledge with visitors on the subject of Health, Humanities etc.
21. **AltaVista** (www.altavista.com) : One of the largest databases with some 30 million web pages. It is very fast and takes a few seconds to perform a simple search. Most importantly it indexes, full-text of all retrieved documents.
22. **All the web** (www.alltheweb.com) : Claiming to be the largest of the search engines, capable of tracking more than 200 million URLs, its product (commercially known as Fast Search) boasts over 300 million web pages nicely indexed. Searchers can ask for exact search-phrase returns, partial results, or all search terms in no particular order.
23. **Dogpile** (www.dogpile.com) : It offers access to and very accessible information on, such everyday and diverse topics as automobiles, furniture, travel, online trading, pharmaceuticals, free music, news, self-improvement and computer.
24. **Google** (www.google.com) : This is the most popular search tool, Huge database available on this site. Very authentic and systematically arranges the search result.
25. **Hot Bot** (www.hotbot.com) : HotBot is based on search technology developed by the University of California. Berkley. HotBot claims to have the largest index of web pages, at the same time it is one of the most accurate, using relevance scoring technology. It is most useful in a comprehensive search.
26. **Grokker** (www.Grokker.com) : a team of experienced executives and distinguished directors and advisors to address

the challenges associated with the emerging trends in information management and communication. Together they created Grokker, a web-based enterprise search management platform that leverages the power of federated content access and visualization to maximize the value of information assets for enterprises, content publishers, libraries and other research-intensive organizations.

27. **Mamma (www.mamma.com)** : Mamma a Meta search engine also known as the Mother of all search engines. When users enter a query, Mamma simultaneously queries 10 of the major search engines and properly formats the word and syntax for each source, organizes the results into a uniform format, and presents them in order of anticipated relevance.
28. **Lycos (www.lycos.com)** : Like AltaVista, Lycos has a very large database. It claims to have indexed nearly 130 million web pages. It only indexes the title and then the first 20 lines of the text and headings. Lots of good sources are provided on the Net.
29. **MetaCrawler (www.metacrawler.com)** : Its slogan is "Search the search engines". Simultaneously search into 10 search engines, and rank the results from each site in 20 categories in three languages.
30. **Yahoo (www.yahoo.com)** : The oldest and largest web directory of sites and the most popular website, which has indexed about 25 million web pages. This engine is useful for beginners on the Net.

Conclusion:

Most of the libraries and librarians are offers ICT based services in the library to its patrons. Search engine is an effective and additional tool to search pin-pointer information to save the time and energy of the

user. This is very much essential and needed that library staff must have adequate knowledge to use various Boolean search facilities as well the various functions available on different search engines which will helps to provide better services and educate to users.

References:

Barlow, L. (2004). A Helpful Guide to Web Search Engines—How Search Engines Work (page 4). Retrieved December 13, 2022, from <http://www.monash.com/spidap4.html>

Editors of the American Heritage Dictionaries. (2022, May 12). Search Engine [Online Dictionary]. Retrieved May 12, 2022, from American Heritage Dictionary of the English Language website: <https://www.ahdictionary.com/word/search.html?q=search+engine>

Kale, R. S. (2009). E-RESOURCES IN MANAGEMENT WITH SPECIAL REFERENCE TO OPEN ACCESS:AN EVALUATION (p. 119) [Unpublished Mphil Dissertation]. Waghodia, Vadodara,: Sumandeep Vidhyapeeth University.

Rigdon, J. C. ed (Ed.). (2016). Dictionary of Computer and Internet terms (1st ed.). Cartersville, GA: Eastern Digital Resources.

Shailendra, K., & Anas, S. (2006). Effective tools to navigate for digital information on the world wide web. In C. Anandanm ed & M. Gangatharan ed (Eds.), Digital Libraries ; From Technology to Culture (1st ed., pp. 230-235.). New Delhi: Kanishka Publishers and Distributors.

Wolff, A. (Ed.). (2006). BRITANNICA CONCISE ENCYCLOPEDIA: Revised and Expanded editon (Revised and Expanded editon). Chicago: Encyclopaedia Britannica, Inc.