Leading Article

Tracing medical information over the Internet

Sajed M. Al-Mutairi, *MBBS, MSc.*

ABSTRACT

The Internet became with do doubt a huge and valuable source of information for researchers. The wealth of information on the Internet is second to none and medical information is no exception. Yet with the vast expansion of the Internet and the World Wide Web in specie, to find the kind of information one is looking for, he/she needs to browse thousands of web sites and the experience would be like digging into a stack of hay looking for a needle. That's why search engines and subject indexes, as means to overcome this problem, were introduced and grew so rapidly. In general, there are three approaches to retrieve data from the World Wide Web; the subject directories, search engines and detailed subject indexes. However, there is no single comprehensive search engine or directory and it is recommended to use more than one with different keywords and synonymous.

Keywords: Internet, World Wide Web, search engine, subject directory, medical informatics.

Saudi Medical Journal 2000; Vol. 21 (5): 421-423

The Internet has become a part of our society with breath taking rapidity. Even for those who are not yet Internet literate, references to it appear daily and with increasing frequency because this powerful network carries a vast body of information that is transmitted quickly and efficiently.

Medical information makes an actively increasing large part of this network. Some authors believe that the Internet will be the sole guide to creating one worldwide scientific community.[1]

The speed by which the Internet is expanding and developing led LaPorte and his colleagues to proclaim the death of paper medical journals. They argued that, the Internet most certainly is faster and research findings will be presented to scientists in days rather than years. It is also less expensive and potentially more equitable, at least for the type of information transfer entailed in a short research communication. Moreover, they added that the Internet is spreading rapidly to almost all countries; thus researchers in developing countries will have access to communication that they do not have through print.[2]

An advantage of the Internet over the professional databases is that it provides a lot of supplementary information; on authors, their works and research projects and on the supporting foundations behind them.

Locating information on the WWW. Tracing down relevant information on the WWW showed to be a difficult and tedious task. As a result, during the last few years, the WWW witnessed the development of significant and substantial databases and search engines of its continuously increasing resources. However, these databases and search engines still have a long way to perfection[3] and it is relatively rare to find the required information as a result of a first search using these search engines and databases.[4] Trials indicate that the process of finding information via search engines was simply too complicated for Internet naive people to use without very high levels of support.[4] On the other hand, users, after gaining the necessary experience, may find searching the Internet for information more efficient than locating the same information through a library.[5]

The continually changing nature of the WWW and

From the Occupational Health Department, Ministry of Health, State of Kiwait.

Address correspondence and reprint request to: Dr. Sajed Al-Mutairi, PO Box 957, Al-Surra, Kuwait. Fax. +965 326 2045 E-mail. Sajed@kuwaiticorner.com

its overall lack of organization are the reasons that make it more difficult to search than commercial electronic databases. It requires different search strategies as the systematic approach of finding one relevant citation in a controlled vocabulary index and then finding more based on its subject headings does not work well in WWW databases which lack such controlled vocabulary systems.[3]

There are 3 main approaches to locate information on the WWW: either by using Subject Directories, by using Search Engines or by using Detailed Subject Guides. All are electronic cataloguing and registering systems that break the information on the WWW into manageable bits and return an online list of sites, documents, pictures, and sound or video clips.[6] There are more than 20 Subject Directories, Search Engines, and Detailed Subject Guides available for the WWW.

Subject Directories. Subject Directories provide a hierarchical index of the WWW. Such an index is divided into logical categories such as arts, business, education, government, health, etc. and from there into sub-categories,[7,8] (for example; the logical category "health" is divided into 43 sub-categories in Yahoo.[9] By selecting from the logical categories, and then from the series of lists of increasingly more specific sub-categories, the user is guided to a list of WWW sites that offer information relevant to the topic of interest.

Subject Directories can also be searched by keywords, however one study showed that users prefer searching for information from hierarchical categories rather than keywords.[4]

Examples of Subject Directories. Yahoo (http://www.yahoo.com) is the best known and most popular of the Subject Directories.[10,11] It has links to approximately 370,000 web sites.[12] A survey conducted by the market research company, Relevant Knowledge, identified Yahoo as the most "highly trafficked site on the WWW", with 26 million users accessing it in January 1998.[7] Yahoo is still one of the few Subject Directories that did not incorporate an active search engine.

Excite (http://www.excite.com) is another Subject Directory which has a special section that is called the Excite Reviews Database, that holds more than 50,000 WWW edited sites reviews.[12,13] In contrary to Yahoo, Excite incorporated an active search engine component in addition to its Subject Directory.

Strengths of Subject Directories. They allow the user to identify Internet resources from a broad subject base, thereby negating the need to identify highly specific search terms.[6,7] Their hierarchical and browseable structures provide a logical and accessible route to the Internet and thus are an ideal starting point for the new user.[4,7,10]

Weaknesses of Subject Directories. As the creation of subject heading is human-based, the directories tend to be smaller and less up to date than the Search Engines[6,7] and they only become aware of new sites when notified of them by a site's author or administrator.[12]

Relevant Internet resources maybe be misplaced.[7,8] They might point to WWW pages that no longer exist.

Search Engines. These are active databases that provide WWW pages' addresses by using a Robot (also called a spider, wanderer, or crawler).[8] Robots are computer programs that scour the Internet automatically and systematically, server by server, looking for new resources to index and check that previously identified sites are still available.[7] In addition to the search robots, most Search Engines have a Subject Directory component added to them.[13] Search Engines are searched by keywords and phrases, and most of them provide Boolean logical operators functionality.[14]

Examples of Search Engines. AltaVista (http://www.altavista.com) which is created by Digital Equipment Corporation has the largest coverage of the WWW (110 million distinct pages in March 1998).[2,8] AltaVista is the search tool of choice as it can be searched in seconds, whilst its "refine" features allow users to focus a search to a more manageable number of relevant hits.[7]

Northern Light (http://www.northernlight.com) is ranked as one of the 3 largest Search Engines available.[14] This Search Engine provides a different approach to web search, as it includes an index of around 50 million WWW resources with a special collection of full-text journal and newspaper articles. Northern Light provides health professionals with a rich and, at times, a unique source of information.[7]

Hotbot (http://www.hotbot.com), Infoseek (http://www.infoseek.com) and Lycos (http://www.lycos.com) are other Search Engines available.

Strengths of Search Engines. Users can undertake a very specific search and find WWW pages that precisely match the requirement.[7] Due to the automated comprehensive indexing of these engines, only few searches fail: no matter how obscure the search, the engine will return a resource location on the WWW.[7]

Weaknesses of Search Engines. The non-discriminatory method of document retrieval inevitably returns a number of irrelevant WWW pages.[13]

The resources added to these databases are not evaluated by any means.[7] They might point to WWW pages that no longer exist.[13]

Some of them have no duplicate detection, the thing that will generate large lists of repeated WWW links.[6]

The Boolean processing is not always accurate in most of them.[14]

Detailed Subject Guides. A number of organizations (and individuals) maintain Detailed WWW Subject guides and starting points. Some of

these are focused on specific topics while others offer a systematic overview of the WWW's resources. Some of the Detailed Subject Guides contain specialized search engines that focus on the topic at hand, and which have access to information unavailable to more general search engines.[8]

Examples of Medical Detailed Subject Guides. OMNI (Organizing Medical Networked Information) is one of the most developed specialized Medical Detailed Subject Guides. It is funded by the Joint Information Systems Committee (JISC) and operated by a consortium including the British Medical Association Library, Medical Research Council, Wellcome Centre for Medical Science and others. OMNI contains a searchable and browseable database of UK and global medical education and research resources, providing a brief description of each resource.[7,15] OMNI can be reached at the URL (http://omni.ac.uk).

Medical Matrix (http://www.medmatrix.org) which is produced by Healthtel Corporation with support of the Internet Working Group of the American Medical Informatics Association is considered as one of the more established medical search services.[7,15] It is well organized and has a user-friendly layout, however some of the inclusions could be heavily augmented (e.g. the disease category).[6]

Health on the Net (http://www.hon.ch) is produced by the Health on the Net Foundation which is a non-profit organization dedicated to "realizing the benefits of the Internet and related technologies in the fields of medicine and healthcare". This Foundation has created a Code of Conduct for the health sites on the Internet. Sites complying with this Code meet a defined quality standard and accordingly included in the database. However, the Health on the Net database is not, surprisingly, limited to sites that comply with code. Moreover, it would be a mistake to assume that sites not complying with code are not worth using.[7]

Strengths of Detailed Subject Guides. Usually, professionals evaluate the resources before being included in these guides.[7] The resource entries are accompanied by short descriptions that enable the user to have a fair idea in advance.[7]

Weaknesses of Detailed Subject Guides. Most of them are still in their beginnings and accordingly their databases are relatively small.[7] Being compiled by individuals, potentially useful resources might be overlooked.[7]

Optimizing the WWW search experience. Few guidelines should be followed to optimize the experience of conducting a search of the WWW. First, search terms should be as specific as possible. Due to the literal nature of the registration process in most Search Engines, keywords and phrases related directly to the topic of interest must be used in order to avoid getting unrelated resources.[7]

Second, synonyms and different wording for the topic of interest should be used as most of the Search Engines and Subject Directories lack the presence of synonyms detection systems.

Third, most Search Engines and Subject Directories have advanced search instructions, reading the instructions of each is very important to gain the most out of the search experience. Advanced search techniques are not standard among these Search Engines and Subject Directories.

Finally, there is no single comprehensive Search Engine or directory. It is wise to use more than one, as each will return many unique resources.[3,7,12,14]

References

1. Onganer E, Cinar Y. Online with physicians accessing the Internet [letter; comment]. JAMA 1996; 276: 1218.
2. LaPorte RE, Marler E, Akazawa S, Sauer F, Gamboa C, Shenton C, et al. The death of biomedical journals. BMJ 1995; 310: 1387-1390.
3. Notess GR. On The Net. Internet Search Techniques and Strategies. Database. Updated in (Unknown). Accessed in 29/5/1999. Available online from: http://www.onlineinc.com/onlinemag/JulOL97/net7.html.
4. Pollock A, Hockley A. What's Wrong with Internet Searching. In: Designing for the Web: Empirical Studies; 1996 30/10/1996; Microsoft Corporate Headquarters, Redmond, Washington: Microsoft Usability Group; 1996.
5. Kipp DE, Radel JD, Hogue JA. The Internet and the nutritional scientist. Am J Clin Nutr 1996; 64: 659-662.
6. Gibbs SR, Sullivan-Fowler M, Rowe NW. Mosby's medical surfari: A guide to exploring the Internet and discovering the top health care resources. St. Louis: Mosby; 1996.
7. Kiley R. Medical information on the Internet: A guide for health professionals. 2nd ed. Edinburgh; New York: Churchill Livingstone; 1999.
8. Barrett DJ. Netresearch: Finding information online. 1st ed. Sebastopol, CA: Songline Studios: O'Reilly & Associates; 1997.
9. Yahoo! Health. *Yahoo*. Updated in (Unknown). Accessed in 28/5/1999. Available online from: http://dir.yahoo.com/Health/.
10. Notess GR. On The Nets, Comparing Net Directories. Database. Updated in (Unknown). Accessed in 29/5/1999. Available online from: http://www.onlineinc.com/database/FebDB97/nets2.html.
11. Bride M. Teach yourself the Internet. London: Hodder & Stoughton Ltd; 1998.
12. Venditto G. Search engine showdown. Internet World 1996; May 1996: 79-86.
13. Missingham R. Will You Still Need Me When I'm 64? Libraries and the Internet. In: The Eighth Australasian Information Online & On Disc Conference and Exhibition; 1997 21/1/1997; Sydney Convention and Exhibition Centre, Sydney, Australia; 1997.
14. Notess GR. Comparing Internet search engines. In: The Ninth Australasian Information Online & On Disc Conference and Exhibition; 1999 19/1/1999; Sydney Convention and Exhibition Centre, Sydney, Australia; 1999.
15. McKenzie BC. Medicine and the Internet: Introducing online resources and terminology. 2nd ed. Oxford; New York: Oxford University Press; 1997.