AWARENESS ABOUT AND USE OF INTERNET SEARCH ENGINES AMONGST SOCIAL SCIENCE RESEARCHERS IN NORTH EAST INDIA

Bikika Laloo   *Narendra Lahkar*

Abstract

Discusses findings of a study conducted on Social Science researchers in two Universities in North East India. Findings show that Social Science researchers in North East India are having the same issues regarding Internet search engines as researchers in other parts of the world. Concludes that better presentation of web sites would help them get indexed by search engines and make them more accessible. Also, improvement in search strategies would help researchers find more relevant information, faster.

Keywords : North East India; Social Science Researchers; Search Engines; Search Strategies

1.   Introduction

The Internet and the World Wide Web made inroads into North East India quite late but once the ball was set rolling there was no stopping it. There is now almost a deluge of websites, both personal and corporate emerging from this part of India. While the National Informatics Centre (NIC) is credited for preparing most of the official websites, individuals and academic institutions too are realizing the enormous potential of the web.

Information on almost all aspects of the North East is now available on the Web. Whether it is well presented is another matter of course. A more important concern is the awareness about this information. And awareness about web resources is brought about by the following means :

- Through verbal communication
- Through print sources
- Through the radio and television
- Through search engines

The concern of this paper is the awareness about and use of Internet search engines amongst the social science research scholars in North East India.

2.   Review

A person venturing into the Internet will find no dearth in the number of services set to help him. The web is crawling with search sites vying to show us the way (Zetter and Mc Cracken, 2000). Conservative estimates put the number at over 8000 including pure search engines, general and

---

special interest directories and metasearchers. Since the Internet revolution began, frequent coverage on search tools has been found in both established computer magazines and new publications. In most articles, Internet-oriented investigators have looked at the utility and/or convenience of various Internet search tools and their ability to locate information on the Internet, then offered their own subjective analysis of how easy such tools are to use. Such articles are updated as more people are introduced to the Internet and as new search tools are introduced and old ones phased out. These investigations have been carried out by writers for the particular magazines, who tested and rated the tools themselves as opposed to investigating the opinions of actual end-users. However, the writers employed by such computer journals can be assumed to have a certain amount of professionalism, with Internet knowledge somewhat advanced beyond the typical end-user, and more on a level of today's professional librarians and library students who have formal training with Internet resources (Vaughan, 1999). These studies vary in terms of their scope and coverage. While most compare features of search tools, some provide more detailed discussions. All these works throw some light on the strengths and weaknesses of various search tools and thus help the user not only to learn about them and make appropriate use of them for retrieving information from the web. (Choudhury, 1999). Articles such as those by Winship (1995), Zorn et.al. (1996) Machovee (1996), Falk (1997), Oppenheim (1997) and Zetter and Cracken (2000) describe the features of different search tools. Others like those by Poynder (1996) and Poulter (1997) are more evaluative and critical comparing search tools on the basis of their generic features such as database content, retrieval software and search interface.

In an effort to better speculate whether a certain set of factors play a role in information professionals' choice of Internet search tools, Vaughan (1999), conducted a survey of MSLS/MSIS graduate students and professional librarians at the University of North Carolina, Chapel Hill. Background discussion on Internet search tool design, usability, field testing and future developments was provided. Two sets of factors were defined for the study – one describing utility functions, the other describing the convenience or ease of use of search tools. The survey revealed a trend in choosing a preferred Internet search tool based on utility factors as opposed to convenient factors. It also suggested a preference for search engines as opposed to subject directories.

Voorbij (1999) conducted a nationwide survey in the Netherlands among students and academics, to explore the use and perceived importance of the Internet for study or work related purposes. More specifically, the study focused on searching information resources on the World Wide Web. The user survey consisted of two parts. First, a rather detailed questionnaire was sent to 1000 members of the academic community. Second, three focus group interviews were held with experienced Internet users. Among other findings the study revealed that searching the World Wide Web is not without difficulty. The Web is being used primarily to search general factual, ephemeral or very specific information. At the moment, full text resources play only a minor role in the academic research process. The Internet may have conquered a place for itself, but it has not pushed aside traditional printed and other information sources.

The Ziff-Davis Publishing Company ranked 14 catalogues and search engines according to usability (interface, design, overall ease of use) and effectiveness (number and precision of returned results) and summarized their opinion of each. Infoseek was ranked slightly better than several close competitors.

Thelwall (2000) conducted a survey to find sites not registered with search engines as well as those that were. The study employed a combinatorial method for finding addresses of Web sites. The survey was conducted in order to test the coverage of search engines and to decide whether their partial coverage is an obstacle to using them to calculate Web impact factors. The results indicated that search engine coverage even of large national domains is extremely uneven and would be likely to lead to misleading calculations.

"How to Stop Searching and Start Finding," is a study by Zetter and McCracken (2000) where they ran a series of queries ranging from broad to specific on twenty search engines, directories and expert sites to see which ones produced the best results. Their objective was to find out which of these search tools provided the most relevant links in the most logical ranking with the least effort. They found that while some search tools did surprisingly well while others forced them to "wade through swamps of irrelevant links," or pointed them to pages that no longer existed. They ranked Google number one among the "search superstars" because it was sleek, simple, fun to use and delivered the web's most relevant results as promised. Their overall conclusion was that search technology while far from perfect, has made great advances in helping find their way effectively through the web.

Killmer and Koppel calculated the retrieval effectiveness of 3 search engines – Google, LookSmart and Metacrawler in 2002. They had students review the first 20 sites returned by each engine. The students noted down the number of relevant sites and the total number of unique sites returned by each search engine. The exercise was given as a homework assignment, therefore the searches were not conducted simultaneously as they may have been in an in-class assignment. The retrieval effectiveness results, with 11 unique relevant sites identified among the three engines were :

- Google : Precision=9/20; Recall = 9/11
- LookSmart : Precision =3/20; Recall = 3/11
- Metacrawler : Precision = 1/3; Recall = 1/11

The general consensus from students was that although many of them used search engines prior to this lesson, they did not realize that all search engines do not return the same results. Furthermore, they believed that search engines searched the entire web as it existed at the moment the search was conducted. Once students learned and witnessed the differences among search engines first hand, they found the concepts of recall and precision useful in their comparison of search engine effectiveness. Students commented that this exercise made them more information consumers. Students also commented that prior to this exercise, they were unaware of the vast number of search engines that exist.

Perfection in retrieval effectiveness is a utopian dream as of now. Visitors to the World Wide Web discover that it is a frustrating task to find specific pieces of information or digital objects despite teasing links that suggest they are only a Website away. (Casey, 1999)

According to Harris (1997), more often than not, a search for specific data can degenerate into an afternoon of dead ends, blind leads and false drops if not outright misinformation.

Too often, the searcher gives up and goes to the library stacks taking two steps at a time. Frequent visitors to the web conclude that accessing digital information is not as convenient as a person using

a book; turning back and forth between two pages for comparison (once one has found the two pages) is as fast as one cares to have it and no command structure and windowing capability stand in the way (Suber, 1992).

Demas et. el. (1995) are of the opinion that just as a library of uncatalogued, haphazardly shelved books is of little use when searching for a specific book, the web will never be a true research tool and resource until a means of directly accessing discrete pieces of information and digital objects is developed. It is clear that the information highway will not magically sort itself out.

3.   Methodology

A set of questionnaires (containing twenty-two questions) was distributed to research scholars working in the departments under the Social Science faculty in Gauhati University – Guwahati and North Eastern Hill University – Shillong. Five questionnaires were distributed to each department. A total of fifty-five questionnaires were distributed in all.

It was difficult to meet all the research scholars in person as most are part time researchers. Only a few were met, with whom discussion was held on the topic in question. The questionnaires were mostly given to the heads of the departments or the office staff to be handed over to the research scholars.

4.   Results

We faced various constraints in getting data. Though every effort was made, a total of twenty researches returned the questionnaires prepared for the purpose. Admittedly, a study with such poor responses cannot present that clear a picture of the situation but it plays its own small role.

Clearly the researchers in the two universities acknowledge the importance of the Internet, many (10) of them agreeing that the Internet is indispensable to their work both now and in the future. Whatever their reservations about it, they were all sure that using the Internet is not a waste of time. And they speak from experience since more than half (12) claim to have been using the Internet for more than one year.

Email comes up tops amongst the ten Internet facilities listed, with six of the respondents claiming to use it daily. Electronic journals follow closely with five respondents using them daily. WWW resources are also not lagging behind with four researches using them daily. Some of the facilities (discussion lists, newsgroups, campus wide information systems, external catalogues, bibliographic databases and documentary delivery services) are either unknown to the researchers or are hardly used.

Following the almost universal trend, Google is crowned number one amongst the eleven search engines listed, by researchers in the North East too. Thirteen respondents admit to using it often. It is closely followed by Yahoo (10 frequent users). Excite, Hot Bot and Web Crawler also find mention with one frequent user of each. Other than the listed search engines, mention is also made of resources like ssrn, sify, inflibnet, rediffmail, metacrawler and dogpile.

Many of the respondents seem to be fairly confident in their use of the search engines judging from their (9) claim of using both simple and advanced search options. Amongst the advanced search options listed, phrase search (5-use often), truncation and case sensitivity (6- use sometimes) come out as the most used, followed closely by field searching (4- use often). Boolean operators don't seem to be popular with the respondents.

The Internet may be 'in' but it is not totally foolproof, a fact substantiated by thirteen respondents who complain that while it does provide "something" to meet their needs, "but not as much as needed." Three respondents felt that it provides "enough" information. One brave soul told it as it is – "the Internet provides more information than is needed." Despite their difference of opinion, the respondents were unanimous in their agreement on the fact that the Internet does have its uses.

Regarding finding information on North East India, ten respondents admitted to it being difficult, six of them attributing this to the lack of relevant web sources on North East India. Three admitted the reason to be their inability to use the right search strategies while two felt that while there are many web sources on North East India, they have not been indexed by the search engines. Six respondents claimed to be able to find information quite easily.

Eleven respondents admitted to not being aware about the sources on North East India, available on the Internet. Those who were aware (5) gave credit to search engines which helped them locate the correct URLs. Two admitted to being directed to the correct URLs by people. Between these two methods, search engines were considered more time and energy saving.

Even where finding information on North East India is concerned, Google (13) and Yahoo (6) were credited for giving the fastest and most relevant results.

The websites on North East India consulted frequently by the researchers include www.kanglaonline.com, www.Epao.nic.in, www.youngamonline.com, www.kukiforum.com, www.ceniseas.org, www.manipuronline.com, www.thesangaiexpress.com, www.neea.com, www.nedfi.com, www.sapt.org, www.ipcsmonitor.com.

5. Discussion

Covering as it did such a small sample, this study is a mere drop in the ocean of studies related to search engines users. But it does manage to throw a little light on the situation in North East India.

From the results, it is apparent that Internet users in North East India are hardly different from those in other parts of the country and the world. While they all admit to the potential of the Internet, not all of them are fully confident in using it. When they do use the Internet, they are willing to try both simple and advanced search options indicating that they are warming up to the Internet.

Where choice of search engines is concerned, there are no surprises. Following the worldwide trend, Google and Yahoo top the list of most used tools.

And as with many users in the rest of the world, a lack of skills in searching becomes a stumbling block for researchers in North East India too.

It is possible that many websites from North East India are not indexed by search engines. This could be because they have not been presented or prepared well as a result of which they are ignored. The loss is the researchers'.

One question that could have been asked but was not is whether the researchers prefer print over Internet sources and why. This could have provided an opportunity for comparison. Perhaps it can be included in other studies.

To conclude, the search engines situation in North East India is hardly different from that found in other parts of the country and the world. If researchers are not finding relevant information, it is because the information has not been indexed owing to poor presentation. Also the lack of searching skills amongst searchers proves a stumbling block. Necessary changes need to be made by the web site designers as well as the searchers.

References

1. Casey, C. (1999), "An Analytical Index to the Internet: Dreams of Utopia," College and Research Libraries,   pp. 586-595.

2. Chowdhury, G.G. (1999), "The Internet and Information Retrieval Research: A Brief Review," Journal of Documentation, Vol.55 No.2, pp. 209-225.

3. Chu, H. and Rosenthal, M. (1996), "Search Engines for the World Wide Web: A Comparative Study and Evaluation Methodology," ASIS Annual Conference Proceedings www.asis.org/annual-96/Electronic Proceedings/Chu.html.

4. Courtois, M.P., Baer, W.M. and Stark, M. (1995), "Cool Tools for Searching the Web: A Performance Evaluation," Online, Vol.19 No.6, pp. 14-32.

5. Galloway, L. et. al. (1999), "Search Engine Retrieval Effectiveness for Medical Information Queries," web.syr.edu/gevander

6. Harris, S.R.(1997) Webliography: "The Process of Building Internet Subject Access" in Acquisition and Collection Development in the Humanities. Irene Owens (Ed.) New York, Haworth Press, p.30.

7. Kimmel, S. (1996), "Robot-generated Databases on the World Wide Web," DATABASE, Vol. 19 No.1, pp. 41-49.

8. Machovec, G.S. (1996), "World wide Search Engines: Open Text, Harvest, 2Ask," Online Libraries and Microcomputers, Vol.14 Nos.6-7, pp. 1-5.

9. Suber, P. (1992), "The Database Paradox: Unlimited Information and the False Blessing of Objectivity," Library Hi Tech Vol.10 No.4, pp. 51-57.

10. Thelwall, M.(2000), "Web Impact Factors and Search Engine Coverage," Journal of Documentation, Vol.56 No.2, pp. 185-189.

11. Tseng, G., Poulter, A. and Hiom, D. (1996), The Library and Information Professional's Guide to the Internet. London, Library Association Publishing.

12. Vaughan, J. (1999), "Considerations in the Choice of Internet Search Tools," Library Hi Tech, Vol. 17 No.1, pp. 89-106.

13. Venditto, G. (1996), "Search Engine Showdown," Internet World, Vol.7 No.5, pp. 79-86.

14. Vorbij, H.R. (1999), "Searching Scientific Information on the Internet: A Dutch Academic User Survey," Journal of the American Society for Information Science, Vol. 50 No.7, pp. 598-615.

15. Winship, I.R. (1995), "World Wide Web Search Tools: An Evaluation," Vine, Vol. 99, pp. 49-54.

16. Zetter, K. and McCracken H. (2000), "How to Stop Searching and Start Finding," PC World, pp. 129-143.

17. Ziff-Davis Publishing Company (1995), "Hide and Go Seek: Rating Internet Search Engines. http://www.zdnet.com/pccomp/features/internet/search/sub1.html.

18. Zorn, P., Panek, M., Marshall, L., and Emanoil, M.(1996), "Advanced Web Searching: Tricks of the Trade," Online, Vol. 20 No.3, pp. 15-28.

19. Killmer, K.A. and Koppel, N.B.(2002), "So Much Information, So Little Time. Evaluating Web Resources With Search Engines,"