CHAPTER- IV

Review of Literature

4.1. Introduction

4.2. Review of Literature
   4.2.1 Use of Internet and Internet Resources
   4.2.2 Search Engines
   4.2.3 Evaluation Criteria for Evaluation of Web based Sources and Websites

4.3. Conclusion
4.1. Introduction

The focus of this study is to examine the use of web resources by the faculty members and research scholars of botany working in universities of Karnataka and Maharashtra. The study also evaluates the botany websites using website testing tools. It is known fact that since the beginning of the present decade quality issues have been discussed far and wide in the information sector. In the latest years we have seen an overwhelming implementation of the World Wide Web (WWW) in most fields including Botany. Many studies of information quality in connection with the WWW include lists of criteria, which may be applied to distinguish between information of “good” and “bad” quality. Designers of Web pages often forget aspects of user-friendliness, especially with regard to navigation between the pages in a complex Web site. It is true that there are numerous guides and handbooks telling how to develop high quality home pages that will attract the attention of potential users or customers.

In this regard, literature survey has been conducted in the area of Internet use, use and evaluation of Web based sources, and use of search engines. In this chapter, an attempt has been made to provide a comprehensive review of related literature on the problem of research. For this purpose the following sources have been used as primary sources of information.
Chapter 4  Review of Literature

- Library and Information Science Abstract (LISA) – both hard and soft copy
- Guide to Indian Periodical Literature: Social Sciences and Humanities
- Current Contents: Social and Behavioral Sciences.
- Online Web articles.

The existing literature has been collected and reviewed with an eye towards evaluation criteria for printed, Web sources, use of Internet and Web based sources as well as evaluation of search engines and evaluation of web resources for their quality. It is observed that quite a good number of core research articles were found in these areas. The literature survey clearly shows that substantial body of literature already exists in the field of Internet use, use of Web based sources and use of search engines. Several significant authors have carried out many studies on these issues and during recent years there has been a continuous and comprehensive discussion about the criteria for evaluation of World Wide Web information. Thus collected articles have been reviewed and classified under the following major headings.

- Use of Internet and Internet Resources
- Web Search Engines
- Evaluation Criteria for Web based Sources and Website Evaluation.
4.2. **Review of Literature**

4.2.1 **Use of Internet and Internet Resources**

User's studies are undertaken in order to assess the use of Internet and Internet resources and secondly, to know, whether these resources are helping the students, research scholars and faculty members to fulfill their information needs. In other words, user studies are focused on users to understand directly or indirectly their information needs, use behavior and use pattern. Innumerable numbers of studies are undertaken to know the use of Internet and its resources, most of them deal with either specific segments of the university faculty or with specific to use of the Internet. However significant studies have been conducted on the use of the Internet by students, research scholars as well as faculty members. Some such published literature has been reviewed as under.

The study of Chu (1994) on the use of e-mail as scientific communication by Faculty was conducted at two U.S universities. The study reports that major percent (92%) of faculty have used e-mail for scientific communication. However, there is a positive relation between e-mail used and variables such as experience with computer. In the similar way Bane and Milheim (1995) conducted a survey on the use of Internet by academicians who subscribed to 231 randomly chosen discussion groups from a list of Scholarly Electronic Conferences. While they reported the results regarding use of number of Internet services in 17 countries, they did not collect or publish data on the use by discipline, just sector
(e.g. commercial, education and government). Among other findings they state that many academicians are still not aware of its resources and possibilities and not all-foreign countries have access.

Another study on Internet use in academia (Ashley, 1995) is a Ph.D. dissertation which examined network Information Retrieval (NIR) among 888 faculty members at the University of Arizona with Internet-accessible computer accounts. Ashley reports that respondents from various colleges at the universities use between 20% and 39% of available NIR technologies, it suggests that NIR is in an early stage of diffusion in all colleges. Another similar Ph.D. dissertation of White (1995) examined a specific Segment of faculty members, but included non-users as well, distributing the survey by mail to faculty members in professional organizations related to the study of mass communication, consumer behavior and advertising and public relations. Unlike previous studies, this study found that the majority (73%) of faculty used computer mediated communication, with younger faculty members and female faculty members showing significantly higher use than the general population.

Further Alexander (1995) as a part of dissertation work conducted study on the use of Internet list servers as post teleconference support to faculty at Community Colleges and institutions. The author observes that 47% of Community college teleconference participants have access to the Internet and 30% know how to use e-mail. An Australian study by Bruce (1997) analyzed data from two samples of academicians from a wide field of disciplines in
13 Australian Universities, to determine how academicians in Australia use the Internet to enhance their teaching. The data revealed that the Internet for Australian academics represents a mechanism for overcoming the disadvantage to academic teaching which may arise from institutional amalgamation, geographic remoteness or the under-representation of certain teaching disciplines in Australian Universities.

In 1995 another similar survey was conducted on faculty use of electronic information technologies and resources by the SUNY University center libraries (Adams and Bonk). The survey included respondents from all academic disciplines and measured the use and frequency of use of electronic information resources, since the questionnaires were distributed by mail, it measured non-use as well. Variations in use among faculty in disciplines are given, but the conclusion state that, in general, most of the common obstacles to the use of electronic information sources is lack of knowledge about what is available. It is also noted that the user training is considered by faculty to be a high priority need. Faculty readiness in terms of necessary equipment and interest to access electronic information sources is almost universal. In addition, the level of use of available electronic resources is quite high and growing.

In case of Clifford Perry (1995) survey on "Travels on Internet: A survey of Internet users" shows that most of the respondents (55%) had been using Internet for less than one year and 25% of that group had used Internet for less than six months. While 17.9% had been using the Internet since lost 13 to 18
months and an equal number of respondents have been using since more than two years. When they were asked to rate the difficulty they found in the use of Internet, majority (57%) of them felt that it was somewhat difficult. This high percentage would indicate that the Internet is not particularly user friendly. The study also shows that almost all the respondents are interested in obtaining textual information (95%). In addition to texts, users also wanted statistics (44%), graphics (13%) and audio (4%), to learn about the Internet and its offerings and respondents indicated that books and periodicals (26.6%) as well as colleagues (25%) were of great help. The most useful popular tools to access information are Gopher (73%) and File Transfer Protocol (63%).

Articles on use of Web based sources have been collected and presented in chronological order. Regarding use of Web based sources; several studies have found that faculty, research scholars and students who are using the Web for research seems satisfied with what they found. Peter Wie He and Jacobson (1996) studied individuals using Internet terminals at the State University of New York at Albony. The study reported that 97% of respondents felt that Web resources are useful, 45% felt that they could find most of the things they need for their research /assignments from the Internet and 34% described that Internet as their most important resource. A survey of students by Lubans (1998) at Duke University’s undergraduate library showed that the students frequently used Web resources for academic purposes. It is also found that they are confident about their ability to navigate it and generally trust its resources. More than 85% of the students rated the accuracy of Web resources as moderate to excellent.
Similarly, Rama Vishwanathan's (1997) study on the "Internet as a medium for online instruction" reveals that 60% (81 out of 135) of the participants intended to use the Internet primarily for research work. About 21% (29 out of 135) of the participants rated their Internet expertise as poor, compared to 29% (39 out of 135) who rated their Internet expertise as adequate. About 72% of the participants are frequent users; with about 52% respondents used it daily whereas 20% of respondents used it only for a few times a week. Study also found that e-mail was most frequently used feature of the Internet, with about 50% (67 out of 135) of the participants said they used it regularly.

Further Lazinger, et.al (1997) conducted a comparative study on "Internet use of faculty members in various disciplines". The primary focus of the research was to examine and compare the use of Internet among various sectors of the faculty at the Hebrew University of Jerusalem. The survey clearly shows that among the respondents, 80.3% (371 out of 462) used the Internet and out of 371 Internet users 362 (97%) used e-mail. This shows that users use e-mail service extensively; primarily for correspondence with colleagues regarding research issues and respondents judge it as most important service. Although e-mail is the most popular Internet service, there is also sentential use of the other services like FTP(48%) and Gopher(45%). The Internet use rate in general is higher among the Science Faculty members than among the Humanities and Social Science faculty members.
A study of students at three Southeastern U.S college by Perry, et.al (1998) found that more than 40% of students in the study used the Internet at least once a week on regular basis and of these; almost one-half used it to find information. Followed by Perry and others, the similar comparative study was conducted at University of Mysore by Pangannaya and Sujith Kumar (2000) on the use of Internet by the academic community. The results of the study fall in the line of Lazinger's (1997) study, which shows that the member of academic community belonging to the Social Science and Humanities disciplines is making low use of Internet, whereas academic community hailing from Science discipline are making high use.

Further Gilette and Viedeon (1998) used a case study to examine the sources cited by forty-seven students in a community college. They found that regardless of a wide range in quality, 50% of the students in their study cited other student papers found online, citations to Web documents were often muddy, with a high degree of faulty links and errors; and students often cited several sections or chapters of one Web site as multiple sources.

In case of the study conducted by Saeed (2000) on Internet use in University libraries of Pakistan, which highlights about the use of Internet tools viz., e-mail, World Wide Web, FTP, telnet and Gopher. The result of the study shows that there are very few university libraries in Pakistan with access to the Internet, even though those who have Internet access suffer from an acute shortage of Infrastructure (e.g. poor telecommunication, limited hardware and software,
library personnel etc). It is because of insufficient funds and these effects on the access of Internet by university library.

In case of study conducted by Schaffner (2001) examines the impact of electronic technology on libraries and scholarship. It focuses on some of the challenges in using electronic resources in research libraries, which include the cost of acquiring electronic formats and effect that such expenditures have on other library services and collection development practices. The article also explores how electronic resources have changed the way of students and research scholars to conduct research. The goal of this Study was not to criticize or condemn electronic formats but, rather to illustrate that electronic technology is simply one tool among others, for the dissemination of information. As such, electronic resources should complement rather than replace other formats.

Another comprehensive study of Zhang (2001) on scholarly use of Internet based electronic resources shows that e-mail was the most frequently used tool. All respondents indicated that they used e-mail at least once in a week, while 93.6% of them used it almost every day. Web browsers are the second most frequently used Internet tool. Nearly 94% of respondents used Web browsers at least once in a week. Next to e-mail and Web browsers, mailing lists and Internet search engines are third and fourth most frequently used Internet tools. In case of Susan Davis Herring (2001) study on “faculty acceptance of world wide web for students research” explores faculty attitude towards the Web as research tool for their students’ research; their use of Web in classroom
instruction; and their policies concerning use of Web by students. In their study
the authors found that although faculty members generally feel positive about the
Web as research tool, they question the accuracy and reliability of Web content
and are concerned about their students’ ability to evaluate the information found
on the Internet.

A study has been conducted on use of Internet resources by midwives of
New Zealand by Stewart Sarah (2001)\textsuperscript{82}. It describes a pilot study, which took the
opportunity to trial the use of the Internet as a research tool. The pilot study also
tested a survey designed to discover how midwives regard and utilize the potential
of the Internet to revolutionize the provision of maternity care. The study was a
sample of 54 midwives working in New Zealand. The midwives were invited by
personal email to access the web site and complete the questionnaire. This pilot
study has shown that there is a population of midwives in New Zealand who
regularly use the Internet for searching information, which they utilize in their
professional practice, and pass onto colleagues and midwifery clients. Major
findings were that respondents were utilizing the Internet to find information for
themselves and women in their care, as well as for communication and
collaboration. Respondents utilized Internet resources such as email, discussion
lists, personal web sites and online newsletters. They were also very enthusiastic
for as many resources as possible to be made available on the Internet such as
electronic claiming, midwifery education programs and online mentoring. The
main barriers to use of the Internet were insufficient time and training.
Another recent study made by Batthini and Madnani (2003)\(^7\) on Web search behavior of LIS professionals of selected libraries of Ahmedabad and Gandinagar clearly shows that 56.25% of libraries are connected with Internet. It also shows the frequency of use of Internet services by LIS professional. Their study found that 31.25% of them use it every day. Another similar study of Varalaxmi (2003)\(^45\) reveals that 76.47% of faculty have used Internet at their home, 42.15% accessed it from private cyber cafe and equally good number i.e.41.17% accessed the Internet from their work place. Further her study also identified the purpose of Internet use and it shows that large number of respondents (78.43%) used it for research and 58.8% and 43.13% of users use it for teaching and publishing purpose respectively.

Kanaujia and Sathyanarayana (2003)\(^45\) study on the use of Web sources shows that most of the users who used the Web resources belong to 18-24 years age group and the least users belong to 45-65 years age group. Pattern of use of World Wide Web in males is found more than the females. So the popularity of World Wide Web uses more in youngsters and males. Regarding time spent on World Wide Web, the study reveals that most of the respondents (49.2%) browse Web for more than 2 to 4 hours. About 37% users used it for 1 to 2 hours. In this study users were also asked to mention the purpose of use of Web and the data shows that large number of users (42%) used it for e-mail and 36.6% of users used it to browse e-journals.
Another study has been conducted by the Carol Tenopir (August 2003) on use and user of Electronic library resources. The study reveals that in terms of information seeking, today's researcher seems to be comfortable with using a wide variety of sources for information. Internet search engines, e-print servers, author Web sites, full-text databases, electronic journals, and print resources are all used to some degree by most users. Both browsing and searching remain important information-seeking behaviors, but there is some evidence that the amount of searching is going up when users have access to multi-title, full-text databases. Browsing through journal issues is done in print issues or in electronic journals for core journal titles. Articles from non-core journals are most often located through searching. Further students are highly responsive to recommendations of specific resources by their teachers, friends, or a librarian. Educating both high school and college students in the best resources, how to evaluate Web resources, and search strategies is important. Convenience remains the single most important factor for information use—all types of users prefer electronic journals only if they make their work easier and give them the information they need. Desktop access, speed of access, and the ability to download, print, and send articles are top advantages of electronic journals for all groups.

The study was conducted on the success of dotcom by Muhammad A. Razi, J. Michael Tarn and Faisal A. Siddiqui (2004) from Haworth College of Business on the success of the web sites. In this study the potential causes for
Dotcom failures and successes have been examined thoroughly. The success and failure of Dotcoms may not be attributed to a single factor, but to a combination of factors. The findings demonstrate that lack of basic business knowledge, poor or non-existent business plans, ineffective promotion, inadequate back-end logistical support, failure to meet customer expectations, poor customer support, misuse of funds, and dwindling investor faith in e-commerce are some of the critical reasons for Dotcom failures. In addition, the mushroom growth of Dotcoms and changes in customers' buying behavior serve as macro-factors. In some cases, investors were simply not comfortable in their continued support of DotComs when DotComs were not perceived as viable business opportunities. On the other hand, many successful online businesses are hybrid, brick-n'-click type with strong back-end logistical support and an existing customer base. Most successful companies have tackled the shortcomings of failed companies and thrived by creating market niches, producing better Web sites, providing superior logistical and customer service and exploiting demand side economies of scale through product/service differentiation.

4.2.2 Search Engines

Since the arrival of the Web in early 1990's, the Web search engines have become an indispensable tool in our everyday life. When we seek information, we often go to our favorite search engine and look at the returned pages. Given the sheer quantity of information available on the Web, the widespread use of search
engines is not surprising. An individual simply cannot read billions of pages available on the Web, so he gets help from search engines to zoom in to a small number of pages worth looking at. Various studies are conducted for the search engines and its effective use in the information retrieval. Some such studies have been reviewed here.

Leighton (1995)\(^{51}\) conducted a study of Web search engines for course work-by employing the evaluation criterion of precision. The findings were not submitted to a journal for publication because of the fast changing nature of the search engines. Leighton evaluated InfoSeek, Lycos, WebCrawler and World Wide Web Worm using 8 Reference questions from a university library as search queries. The author found that Lycos and the free part of Infoseek have the same precision with lycos just a nose ahead while Webcrawler gave surprisingly bad precision. WWW Worm was good enough that usually retrieved at least one or two hits for the given queries with high precision.

In another article written by Courtois, Baer and Stark (1995)\(^{20}\) evaluated the performances of about 10 different Web search aids including CUI, Harvest, Lycos, Open Text, World Wide Web Worm and Yahoo. Using 3 sample search questions along with other information available about the search engines, the authors concluded that among other things, Open Text was the best at the time of their study with its flexible, powerful search interface and quick response. They also concluded, for novices, WebCrawler offers the easiest interface.
In a different study, Scovilee (1996) surveyed a wide range of Web search engines and suggested that Excite, InfoSeek and Lycos should be added to one’s list of favorites because they can retrieve accurate results from easy-to-use interfaces.

Kimmel (1996) examined World Wide Web Worm, Lycos, WebCrawler, Open Text, Jump station II, AliWeb and Harvest based on documentation provided by the search engines along with a couple of single word test searches (e.g., pollution, atlas, computer etc.). The author’s focus was, like many other publications, on describing the features of these various search engines. The author concludes that if the robot-generated databases presented here, Lycos appears to be the strongest system overall. C-NET, a company specialized in evaluating online products and services, conducted a comparative study of 19 Web search engine on its Web Site (Leonared, 1997). The search engines were tested on their accuracy of results, ease of use and provision of advanced options using 15 queries specifically composed for the evaluation. Most of the queries resemble reference questions asked in public libraries. According to the two feature tables generate by the evaluation, Alta Vista seems to be the best choice among individual search engines, while all-in-one search page and the Internet Sleuth Achieved the highest ranking for meta or unified search engines. The reported findings obviously do not appear to agree with one another. The methodologies and evaluation criteria used by those studies differed as well. Can a
Feasible methodology be developed to help Web users to select search engine, out of that number of choices that is most appropriate to their specific search needs? The Authors of this study are trying to do this by first evaluating the searching capabilities and performance of selected Web search engines currently available.

Chu and Rosenthal (1996) like many other researchers, used sample search Queries that were based upon real reference questions. These queries were structured to evaluate the search engines' ability to deal with variety of query syntax, for example different Boolean logic searches. The authors tested the first ten results from three search engines for precision and concluded that Altavista outperformed Excite and Lycos in both search facilities and recall. Although Lycos had the largest claimed coverage of Web resources, another factor that was taken into consideration was the response time of the search engines, which perhaps surprisingly did not vary between search engines. Other researchers have conducted similar experiments using different queried and analyzing varying numbers of the results for relevance.

Ding and Marchionini (1996) Considered the first twenty hits for precision using five queries. All the search engines searched within twenty minutes. Gauch and Wang (1996) used twelve queries on a number of major search engines and reported precision figures for the first twenty hits.

Tomaiuolo and Packer (1996) studied the precision of the first ten hits from 200 queries but did not list the query topics or the exact expressions or
operators used. Links may not have been visited to check for broken links. Like many other researchers, the criteria for relevance were not given in this study.

In an important paper Clarke and Willett (1997)\(^{15}\) compared the effectiveness of altavista, Excite and Lycos using thirty different searches. The paper is important because of the critical evaluation of earlier research that it provides and also because it offers a realistic and achievable methodology for evaluating search engines. The authors developed a method for comparing the recall of the three sets of searches, despite the fact that they are carried out upon non-identical sets of Web pages, because each search engine has indexed a different set of documents. They developed an algorithm for calculating relative recall by checking how many of the so called relevant found by one search engine were present at all in the universe of documents covered by the other search engine. The study clearly shows that Alta Vista is significantly better at retrieval performance than either Lycos or Excite-this is in accordance with the other studies.

Leighton and Srivastava (1997)\(^{52}\) evaluated five robot search engines, during 31st January and 12th March 1997. These search engines are compared for precision on the first twenty results returned for fifteen queries. The authors rated the services of each search engine based on the percentage of results within the first twenty returned, that were deemed relevant or useful. Their analysis shows that overall, Alta Vista, Excite and infoseek performed on short, unstructured queries, while Hotbot performed better on structured queries.
Daniel Dreilinger (1997) has conducted an experiment on “Experience in selecting the search engines using meta search”. The experimental findings suggest that a metaindex approach can be effective in making search engine selection decisions. However, the potentially large amount of knowledge required to make these decisions raises some question about the overall efficiency of the system. Initially, the categorical version is far superior to the metaindex version. With a small amount of term experience (only four or five uses), the metaindex surpasses the performance of the categorical version on the no Results; because temporary poor performance of a search engine is accommodated through the ranking penalty rather than the query score, the metaindex will reflect long-term poor performance on a query term. However, considerable knowledge (100–200 uses of a term) is required for metaindex to surpass categorical on the Visits measure. These findings suggest that the metaindex approach is better at predicting where not to send a particular query with relatively little word knowledge. Given continued increases in how many words have been repeated, one would expect that, at some point, metaindex visits performance would exceed categorical performance. However, considering the substantial proportion of new words still arriving by the end (9.1%), the crossover point of selecting search engines using metasearch 100–200 previous uses is inordinately large. Major robot search engines report dictionary sizes in the millions of terms; the SavvySearch metaindex contains less than 50,000 unique stemmed tokens. The metaindex approach requires a large user base with frequent updating; assuming a word acquisition rate of 2000 terms per day (the current approximate rate), one million
search terms will be accumulated in just over a year. Consequently, adequately supporting the metaindex scheme requires developing techniques for accumulating more words faster.

Meghaghah and others (1998) examined the effectiveness of five World Wide Web search engines: Yahoo, WebCrawler, Infoseek, Excite and Lycos. The Study involved five queries that were checked against each of the search engines in original and refined formats, a total of fifty searches. The queries vary in terms of specificity and level of difficulty in finding Internet resources. The results of the study revealed that Yahoo had the highest precision ratio for both original and refined queries. Infoseek maintained second place with respect to refined queries and Lycos with respect to original queries. Query refinement resulted in a higher precision ratio. The addition of promising pages, selected by the researchers as likely to be relevant, increased precision in all queries and across all engines.

Schwartz (1998) reviewed the way Web search engines operate and noted the growing body of evaluation literature, much of it is not systematic in nature. She concluded that performance effectiveness is difficult to assess in the context of the Web. She also concluded that significant improvements in general content search engine retrieval and ranking performance may not be possible and she also argued that they are probable not worth the effort. However, search engine providers have introduced some rudimentary attempts at personalization, summarization and query expansion. She suggested there might be a trend towards smaller resource collections with rich metadata and navigation tools.
The literature of the evaluation of Internet search engines is reviewed by Oppenheim et al (2000)\(^6\), although there have been many studies, there has been little consistency in the way such studies have been carried out. This problem is exacerbated by the fact that recall is virtually impossible to calculate in the fast changing Internet environment, and therefore the traditional Cranfield type of evaluation is not usually possible. A variety of alternative evaluation methods have been suggested to overcome this difficulty. The authors recommend that a standardized set of tools is developed for the evaluation of web search engines so that, in future, comparisons can be made between search engines more effectively, and that variations in performance of any given search engine over time can be tracked. The paper itself does not provide such a standard set of tools, but it investigates the issues and makes preliminary recommendations of the types of tools needed.

In middle of second half year 2002, Biradar and Sampath Kumar\(^1\) have made a comprehensive study on evaluation of major search engines. For the evaluation purpose six major search engines viz., Yahoo, Google, Altavista, Excite, Lycos, Hotbot were considered. Search was conducted for a query with using search strategy, which is given by respective search engine and without using search strategy. It is found that when search techniques were used for each search engines, number of hits reduced and majority of them are relevant to the query.
The paper by Thorsten Joachims (2002) on optimizing search engines using clickthrough data presents an approach to mining log files of WWW search engines with the goal of improving their retrieval performance automatically. The key insight is that such clickthrough data can provide training data in the form of relative preferences. Based on a new formulation of the learning problem in information retrieval, this paper derives an algorithm for learning a ranking function. Taking a Support Vector approach, the resulting training problem is tractable even for large numbers of queries and large numbers of features. Experimental results show that the algorithm performs well in practice, successfully adapting the retrieval function of a meta-search engine to the preferences of a group of users.

Batthini and Madnani (2003) conducted a study on use of search engine by LIS professionals. About two-third of the respondents in this study are dependent on search engines to retrieve information. Of these 60% are depending on them frequently. Study also shows that the professional used Boolean operators (50%) for searching information. Regarding the level of satisfaction, it was found that 70% of the users felt that they satisfied with information retrieved through search engine. In the similar way another study by Kanaujia and Sathyanarayana (2003) reported that Google(97.4%) and Yahoo(93.4%) were most popularly used search engines followed Alta Vista (71.6%). Regarding the use of search strategy the study shows that truncation, phrase searching and field searching are used most of the times as search strategy on World Wide Web. Boolean searching has not been
used extensively as search strategy on Web. Wild card searching is also used
sometimes among all other types of strategies. Another study by Panda and Sahu
(2003) also shows that Google (85%) and Yahoo (75%) are popular search
engines for respondents.

In a study by Daniel E Rose (2004) on understanding user goals in web
search indicates that of user goals has already yielded two unexpected patterns in
web search. First, so-called “navigational” queries appear to be much less
prevalent than generally believed. Second, many queries appear to be motivated
by a previously unexplored goal involving the need to obtain online and offline
resources. More importantly, an understanding of search goals provides a
foundation for tackling the larger problems of conveying user goals to a search
engine (either explicitly or by inference), and modifying the engines’ algorithms
and interfaces to exploit this knowledge.

A paper by Junghoo Cho (2004) focuses on how the popularity of a Web
date evolves over time and how search engines affect the popularity evolution.
Through an experimental study conducted over 7 months showed that popular
pages are indeed getting more popular while unpopular pages are getting relatively
less popular. We then analyzed two reasonable Web models and tried to estimate
the potential impact of search engines on the popularity evolution of Web pages.
The result from the analysis is immensely worrisome. It shows that when search
engines rank pages based on their popularity, it takes several orders of magnitude
more time for a new page to become popular even if the page is of high quality.
Given that Page Rank and its variations are being used by major search engines, the result strongly indicates that many high-quality pages are ignored by Web users, simply because no one has discovered them yet. The author believes that their study demonstrates an urgent need to develop a new ranking mechanism that can potentially identify high-quality pages early on and promote them, so that the problem can be alleviated.

Search engines usually offer a date-restricted search on their advanced search pages. But determining the actual update of a web page is not without problems. Dirk Lewandowski Et al (2004) conducted a study of testing date-restricted queries on the search engines Google, Teoma and Yahoo! Finds that these searches fail to work properly in the engines examined. Finally, discusses implications of this for further research and search engine development.

The study by Hong Iris Xie (2004) evaluates two different types of online information retrieval (IR) systems: online databases and Web search engines, in terms of user generated criteria. It also compares four types of Web search engines: directories, search engines, meta-search engines, and specialized search engines. The results show that three elements are essential to users in the evaluation of online IR systems: interface design, system performance and collection coverage. While participants preferred the ease of use and intuitive interfaces of Web search engines, they also liked the credible and useful information offered by online databases. Based on the discussion of advantages
and problems of online databases and Web search engines, implications of the design of IR systems are further suggested.

Alireza Isfandyari Moghaddam (2006) conducts a study on overlapping of search results in metasearch engines and their common underlying search engines. The findings indicate that MSEs (Metasearch engines) are more likely to find the same documents which are common in their underlying search engines. Furthermore, MSEs which have a successful recall ratio are identified, which a finding of great practical relevance to library and information practitioners is helping users exploit the Internet to best effect. This study provides clear descriptive evidence for the underlying retrieval patterns of important search tools which are commonly used by today’s Internet users.

Jin Zhang, Suyu Lin (2007) has conducted a study on “Multiple language support in search engines”. The study aims to investigate the multiple language support features in Internet search engines. The diversity of the Internet is reflected not only in its users, information formats and information content, but also in the languages used. As more and more information becomes available in different languages, multiple language support in a search engine becomes more important. The study was conducted through survey about existing search engines and to identify search engines with multiple language support feature. The study is further analyzed, compared and characterize the multiple language support features in the selected search engines against the proposed five basic evaluation
criteria. Finally, the strength and weakness of the multiple language support features in the selected search engines are discussed in detail. The results of the study reveals that Google, EZ2Find and onlinelink respectively are the search engines with the best multiple language support features in their categories. Although many search engines are equipped with multiple language support features, an indispensable translation feature is implemented in only a few search engines. Multiple language support features in search engines remain at the lexical level. The findings of the study will facilitate understanding of the current status of multiple language support in search engines, help users to effectively utilize multiple language support features in a search engine, and provide useful advice and suggestions for search engine researchers, designers and developers. Dirk Lewandowski (2008) also conducts the study on problems in use of web search engines to find results in foreign languages. The results of the study show while none of the search engines faced problems providing results in the language of the interface that was used, both Google and MSN faced difficulties when the results were restricted to a foreign language. Searchers should not use the language restriction in Google and MSN when searching for foreign-language documents. Instead, searchers should use Yahoo or Ask. If searching for foreign language documents in Google or MSN, the interface in the target language or country should be used. This study demonstrates a problem with search engines that has not been previously investigated.
A study by Peter Jacso (2007) on clustering search results had a purpose of examining clustering search results. Traditionally, search results from professional online information services presented the results in reverse chronological order. Later, relevance ranking was introduced for ordering the display of the hits on the result list to separate the wheat from the chaff. The results of the examination shows despite its proven efficiency, clustering is not available, except for Ask, in the primary Web-wide search engines (Windows Live, Yahoo and Google).

David D. Oberhelman (2007) study on “Search platform revisited” aims to look at the user-friendliness of search interfaces of different databases. The results of the study shows libraries are increasingly offering multiple means of accessing the data contained in their products, and the era of a single search platform for all library resources is not that far off, and in some cases is already here (with varying degrees of success). Thus the burden on vendors to create ever better search experiences for their users has become all the greater, and the competition both for superior content and superior search interface designing has become a little more intense. This paper shows search capabilities that are available now, and what may be available in the near future.

In case of Dirk Lewandowski (2008) discussions is on retrieval of effectiveness of search engines. The purpose of this paper is to compare five major web search engines (Google, Yahoo, MSN, Ask.com, and Seekport) for their retrieval effectiveness, taking into account not only the results, but also the results...
descriptions. The result shows that the two major search engines, Google and Yahoo, perform best, and there are no significant differences between them. Google delivers significantly more relevant result descriptions than any other search engine. This could be one reason for users perceiving this engine as superior. The paper implies that search engines should focus on relevant descriptions. Searchers are advised to use other search engines in addition to Google.

Myriam Arrue and others (2008) conducted a study on web accessibility awareness in search engines. The goal of this study is to propose the integration of web accessibility measurement into information retrieval processes. Firstly, quantitative accessibility metrics are defined in order to accurately measure the accessibility level of web pages. Secondly, a model to integrate these metrics within IR processes is proposed. Finally, a prototype search engine which re-ranks results according to their accessibility level based on the proposed model is described.

4.2.3 Evaluation Criteria for Evaluation of Web based Sources and Websites

The literature on evaluation of Web resources began to appear in late 1990’s. The World Wide Web information resources have created a new challenge and competition for information professionals. These resources provide greater opportunities for transfer of information from one location to another very easily. The information can be automatically searched in order to locate items of interest and facilitate better and faster access to information through networking. Many
studies have been conducted for evaluating the web resources. Here in this study, an attempt has been made to review the various studies conducted to evaluate the web resources.

Starr (1994)\textsuperscript{81} provides the best Criteria for the evaluation of Internet sources in his research paper “Evaluating Physical Reference sources on the Internet”. He classified and reviewed representative samples of Physical science reference sources and suggested few important criteria viz., purpose, authority, scope, audience and format for evaluating the Physical Science Reference Sources. Although the focus of the study is on reference sources in Physical Science, the evaluation criteria can also apply to other disciplines as well.

Further Retting (1996)\textsuperscript{39} pointed out that criteria could not have been transferred from the print medium to the electronic medium without modification, of course. Even though there is overlapping among them, many of these criteria can be applied to evaluate Web resource. Retting’ criteria for evaluation of Web resources mainly addresses the issue like information accuracy, authority or its creator, appropriateness of links to other resources, quality of organization and ease of use, graphic design, timeliness, level of treatment, indexing and search ability, appropriateness for Internet medium, comparability to other resources on the Internet or in other media and uniqueness.

Pratt and others (1996)\textsuperscript{71} pointed out that with the rapid growth of the Internet, today library users can rapidly access resources from many parts of the globe and from many types of information providers. Meanwhile the Internet
sources frequently publish without filtration and without considering the needs of
the users is poorly maintained and may be only transiently or intermittently
available. Thus user or library authority must be aware of criteria for evaluation of
web resources. For libraries, the basic evaluation criteria of quality, credibility,
accessibility, scope and cost are still issues but once that must be viewed in new
ways.

In this regard Smith (1997) made a comprehensive work on evaluation of
Internet resources, where he reviewed and proposed a toolbox of criteria for
evaluation of Internet resources. The author also made an attempt to evaluate some
Internet resources on the basis of some significant criteria.

Knowlton (1997), writing in the New York Times, noted that many
educators feel all students do not realize how much of vast amount of information
is readily available over the Web is unreliable. He quoted one professor who
noticed a decline in the quality of his students’ papers after they had begun using
the web resources for their research work. Bruce and Leander (1997) writing on
the use of digital libraries and other information technologies in education,
identified major issues in Internet use as perceptions of the value of information,
questions concerning authority and computer anxiety.

McMardo (1998) in his paper “Evaluating Web information and design”
has give comprehensive criteria for evaluation of web resources, which are almost
similar to Smiths’ Internet information evaluation criteria. The paper mainly focused on the methods for critically evaluating the content of web resources and also focuses on the designing Internet documents. The author also aims to identify and provide a manual of good practices from such sources.

Another comprehensive work by Skov (1998) regarding Internet quality states that given the nature of the Internet, where quality resources mix with vanity publishing advertisements and sheer junk, evaluating quality has become essential. Exercising evaluative skills with regard to printed material and traditional databases has always been part of the information professional’s responsibilities. Thus, the author has come to the conclusion that many of the criteria employed at the time of scrutinizing traditional sources are still valid.

While Kapoun (1998) in his article “Teaching undergraduates Web evaluation: A guide for library instruction” has clearly stated that over the last few years, the faculty members are demanding more web usage from their students. In fact, some faculty members may exclude most printed resources in favor of web resources. Thus the students or faculty members must be aware of evaluation criteria for web resources and in this regard he provides best criteria which are most helpful to students and faculty members for the evaluation of web pages. Thus the suggested criteria of Kapoun include Accuracy, authority, objectivity, currency and coverage of web document. As web resources are revolutionizing the academic librarians, many librarians believe that these resources have changed the principles of selection radically; some believe that they will virtually eliminate
selection. Although it is true that the art of selection is undergoing profound change, the selection of materials is still crucial for libraries. In this context, Holleman (2000) suggests four basic criteria for selection—quality, library relevancy, aesthetic and technical aspects and cost—remain the same in electronic era of information. What they mean and how they are used have changed. But, even quality and cost, the two most controversial criteria, carry great importance for the responsible selection of electronic resources. The criteria that are suggested by the foregoing authors, some of them are also suggested by, Metz (2000), Case (2000), Lord and Ragon (2001).

Similarly other researchers Regan (1998), Connel (1999) have also focused on the questions of the accuracy and reliability of web resources. While Regan (1998) noted that poor quality exists even in sites created and maintained by commercial news organizations. A study by Connel and Tipple (1999) tested the accuracy of information on the web using Altavista search engine. A sample of sixty-reference questions were chosen for the study and after retrieving information from the search engine, authors found that only 27% of the pages provided correct answers or mostly correct answers. However, only about 9% of web pages provided wrong answers and 64% of the pages provided no answers to the questions at all.

Although benchmarking has touched many areas of an organization, including information systems, very few examples are available on how this powerful methodology can be used to specifically address one of the fastest growing elements within information systems – the World Wide Web. The study
by Johnson L. Kelsey & Mark M. Misic, (1999) presents a case study on how benchmarking was used to determine how one organization's Web site compared to Web sites of related schools and professional organizations. The results of the benchmarking study provided a measure of how Web site compares to the sites of related organizations, ideas on how we may further enhance our site, and also a way to regularly evaluate our site.

Further Stoker and Cooke (2000) suggest the techniques and evaluation criteria for information sources available on the Internet. The paper highlights on need for qualitative judgments about library materials and acceptable techniques and evaluation criteria for printed reference works are outlined. The emergence of non-book materials generally and electronic information formats in particular, has created the need for new techniques and evaluation criteria relating to Online services and CD-ROM products. Suggested criteria for the evaluation of network information sources are as follows: Authority, genealogy, scope and treatment, purpose, coverage, currency and method of revision, accuracy, objectivity, audience, format, arrangement, technical considerations, price and user support.

Similarly Cottrell (2001) has identified some important evaluation criteria to teach students to evaluate web resources more critically. His research paper mainly focuses on some criteria viz., identifiable-accreditation of source, quality of the source, links to other sites, accessible on older versions of browsers, hyperlinks and navigation, clear directions in the site, credentials of sites, links are updated, information organization and aesthetic features of the site.
Further Grimes and Boening (2001) discuss the evaluation of Web resources in their paper "Worries with the Web: A look at students' use of Web resources". Authors reviewed numerous online guides regarding evaluation criteria and developed a checklist and the authors found a great deal of similarity among the guides, highly reflective of traditional selection criteria used in evaluating print resources in academic libraries.

Similarly Biradar and Sampath Kumar (2001) have made an attempt to bring out a comprehensive evaluation criteria for Web and printed resources and authors also made an attempt to compare possible common criteria applicable for Web and printed sources, will also apply to the Web resources with some modification, Study also suggests some new criteria, particularly for evaluation of web resources.

In the similar way another study made by Biradar and others (2001) on "use of Internet based information resources" at Indian Institute of Science, Bangalore where respondents were asked to rate the World Wide Web information resources based on Yin Zhang study's features. It is observed that 74.07% of respondents rated the features of e-resources as good, 25.37% of respondents rated the features of e-resources as excellent and only few (0.55%) are rated as fair.

In 1999 the University of the South Florida libraries decided to embark on a usability study to coincide with the rollout of a new interface design. Because this type of study had been conducted by Allen Maryellen (2002) with the initial
interface, its implementation, and completion were paramount in the development of new design. Details of preliminary activities, testing methodologies and results of usability testing by the USF virtual library projects usability study group.

The ADA mandates that the library programs and services be accessible to people with disabilities. Suzane L Byerley and Mary Beth Chambers (2002)* study examines the accessibility of two popular web-based abstracting and indexing services, periodical abstracts offered by OCLC FirstSearch, and Gale Group’s expanded Academic ASAP. When accessed by blind users using screen-reading programs. The study measured accessibility based guidelines from the amended section 508 of the Rehabilitation Act of 1973 and on the web content Accessibility guidelines issued by the WWW consortium. The finding indicates that, while each database has a high degree of accessibility using testing of commercial databases with people who rely on the screen readers for access to the web. Librarians must cognizant of accessibility issue and demand assurance from the database vendors that their products are accessible.

Joy Tillotson (2002)** has conducted a study on Web site evaluation. Using a questionnaire distributed to participants in the library instruction programmes at two Canadian universities, the author studied students understanding of the need for Web site evaluation and their ability to articulate criteria for evaluation. The results show that students view web sources somewhat critically and are aware of standard Web site evaluation criteria.
Chapter 4 — Review of Literature

The corporate web site is an indispensable part of any e-commerce venture. If users find it difficult to do electronic transactions, they will visit another online source, or go to a physical source instead. Evaluating the performance of e-commerce Web sites has received some attention in the past few years, but there is still a pressing need for a more comprehensive approach. A study by Rian Van der Merwe and James Bekker, (2003) attempts to address this need by providing an e-commerce Web site evaluation framework and method built on solid business principles, and using multidimensional scaling to analyze evaluation data and present the results in a graphic manner for easy interpretation.

Tsai-Young Hung (2004) has conducted a study on “Undergraduate student’s evaluation criteria when using web resources for class papers”. The purpose of this study is to investigate what core quality criteria undergraduate students use to evaluate web resources for their class paper and to what extent they evaluate the web resources. This study on five web page evaluations and a questionnaire survey of thirty five undergraduate students in information technology and informatics program at Rutgers University. Results show that undergraduate students have become increasingly sophisticated about using web resources but not yet sophisticated about searching them. Undergraduate students only used one or two surface quality criteria to evaluate web resources. They made immediate judgments about the surface features of web pages and ignored the content of the documents themselves. This research suggest that undergraduate
instructors should take the responsibility for instructing students on basic web use knowledge or work with librarians to develop undergraduate students information literacy skills.

A study by Kshirod Das and Jeevan V K J (2006)\(^2\) aims to present the various evaluation methods, but this study utilizes some of the useful checklists/criteria for evaluating information found on web. The results of this study can be useful for information scientists studying and/or working with portal-based retrieval of e-books, and also for the designers and developers of future online interactive information retrieval systems. The evaluation framework should be taken as starting point for further development and refinement by including functionalities not covered. By eliminating the disadvantages identified, these systems could be made more easy to use. Also features need to be critically studied with user interfaces of each portal to find which resources have more user-friendly features and best suited to satisfy the information needs of target users.

Adel M Aladwani (2006)\(^1\) has conducted a study to understand the relationship between web site quality and consumers’ web attitudes and purchases. Only the technical dimension of web site quality influences consumers’ purchasing behavior both directly and indirectly through their attitudes towards the web site. Specific content quality and appearance quality have relatively stronger association with consumers’ attitudes towards the web site than technical quality and general content quality do. Understanding the link between multiple
dimensions of web site quality and purchasing behavior of web consumers should help organizations know how to improve forward integration with their customers. Previous research, although helpful, failed to examine this important relationship. This paper tries to fill this void in the literature.

The purpose of study conducted by Vaughan C. Judd et al (2006)\(^\text{90}\) is to discuss the attempt to find an evaluation instrument for undergraduate students to use to evaluate public web sites, the analysis of the variety of instruments discovered, the subsequent development of an appropriate instrument, and the application of the instrument in workshops with students. The results of the study shows although a number of diverse evaluation instruments from the literature and from web-based sources were examined, none was deemed suitable for students to use, so the authors developed their own. The authors concurred that, based on their assessment of the learning environment, the focus of an instrument should be on evaluation as a process.

Chang S. Nam et al (2007)\(^\text{63}\) conduct a study on web based learning environment. The study addresses two weaknesses i.e. (1) integration of the user interface design with instructional design and (2) development of the evaluation framework to improve the overall quality of web-based learning support environments. Web-based learning support environment for Global Positioning System (GPS) education: Web-based distance and distributed learning (WD2L) environment. The research goals of the study focused on the improvement of the
design process and usability of the WD2L environment based on a theory-based Integrated Design Process (IDP) proposed in the study. Results indicated that the proposed IDP was effective in that the study showed (1) the WD2L environment’s equivalence to traditional supplemental learning, especially as a web-based supplemental learning program and (2) users’ positive perceptions of WD2L environment resources. The study also confirmed that for an e-learning environment to be successful, various aspects of the learning environment should be considered such as application domain knowledge, conceptual learning theory, instructional design, user interface design, and evaluation about the overall quality of the learning environment.

The research by J. Michael Pearson and others (2007) aims to investigate the relative importance of five key criteria in assessing web usability. These criteria are navigation, download speed, personalization, ease of use, and accessibility. It is believed that these factors play a significant role in assessing web usability. The research indicated that, as expected, ease of use was the criteria that the respondents considered most important in assessing web usability. Interestingly, personalization and customization was consistently reported as the least important criteria. It was also found that males and females view these web usability criteria differently. The results indicate that less effort/resources should be devoted to personalization and customization, and more in making sure those web sites are easy to use and have clear navigation capabilities.
Gnaneswaran V and others (2008) have conducted a study on Evaluating Website Quality Using E-Service Quality Dimensions with objective to determine the key dimensions that affect user experience on the website and evaluate a tool for measuring website quality. Two e-commerce websites, both from same business domain, were tested across the service quality dimensions identified in this study using a questionnaire. 15 subjects from a company group and 10 subjects from a student group participated the user testing. The results of user testing were then compared with website heuristic evaluation and indicated that the e-service quality dimensions could be used to capture user experience and evaluate quality of a website from a user satisfaction point of view. The results also have ramifications for designers of business websites.

Further Lan Anh Tran, (2009) conducted a study on "Evaluation of community web sites" with a purpose to explore an evaluation approach and to develop a model of web site evaluation that includes the specification of evaluation criteria, key issues to discuss and recommendations for improving the web site – in this case, that of the Community Social Planning Council of Toronto (CSPC-T). The evaluation of the CSPC-T web site found mostly positive results in terms of the content of the site and its ability to deliver information to the residents and local communities of Toronto. However, the design of the web site had some negative points requiring improvement, including site structure, page layout and interface design. This paper contributes to the knowledge base in the fields of community networks and web site evaluation, and to the improvement of the quality of information and service delivery to local communities via the web.
Chapter 4  Review of Literature

The study by Candice Dahl, (2009)21 seeks to suggest that criteria commonly used to teach undergraduates to evaluate online resources are inadequate when dealing with non-academic items in the public domain. It aims to argue that these resources should not be ignored by librarians or undergraduates, but that they must still be evaluated. An alternative method of evaluation, based on the concepts of comparison, corroboration, motivation and purpose is to be proposed.

Yasemin Gülbahar (2010)93 has conducted a study to develop a web-based interactive system, Web Macerasi, for teaching-learning and evaluation purposes, and to find out the possible effects of this system. The study has two stages. In the first stage, a WebQuest site was designed as an interactive system in which various Internet and web technologies were used for infusion of technology into teaching and learning process. The Web Macerasi site was used for project work by 92 prospective students who attended different courses in different years. For collecting the students' perceptions about the implementations of the system, a questionnaire of WebQuest effectiveness and a focus group interview guide were developed. Next, the first phase of the study was concluded, and the WebQuest system was updated based on the data gathered from students. In the second phase, 27 students from a different course used the system, and their perceptions were collected through the questionnaire and analyzed. It was found that the students favored the technology-supported media, were more willing to collaborate, found
the feedback very useful, and agreed on the positive contribution of planned works. Consequently, the Web Macerasi site was found to be successful and to have been used effectively in terms of its aims. Further studies should be carried out for diffusion of this technology into the teaching-learning processes.

J.A. Martinez Usero et al conducted a research based on the recently adopted European Interoperability Framework and its Spanish equivalent in order to design an evaluation model to identify key aspects related to information and knowledge interoperability in public organizations. The main objective of the project is oriented to diffuse the information interoperability standards in public organizations and to foster the usage of new techniques and procedures for information integration and management. The method used to carry out the research consists of three main phases: the analysis of information interoperability related technologies, the design of a model to evaluate information interoperability in public web sites, and the empirical analysis of key factors affecting information interoperability. The project main result is an electronic tutorial to facilitate the information and knowledge evaluation process in public organizations and to plan the conversion from non-interoperable technologies and formats to interoperable ones.

A paper by Li, Peng (2010) proposes an automated web site evaluation approach using machine learning to extract evaluation criteria from the existing evaluation data. Evaluating web sites is a significant task because evaluated web
sites provide useful information for users to estimate sites validation and popularity. Although many practical approaches have been taken to present a measuring stick for web sites, their evaluation criteria are set up manually. Thus, authors develop a method to obtain evaluation criteria automatically and rank web sites with the learned classifier. Evaluation criteria are discriminate functions learned from a set of ranking information and evaluation features collected automatically by web robots. Conducted experiments, confirmed the effectiveness of approach and its potential in performing high quality web site evaluation.

Mohammad Hassanzadeh and Fatemeh Navidi, (2010)\textsuperscript{33} has conducted to study provide knowledge about methods used to evaluate the accessibility of Iranian web sites through a real research environment. Findings and evaluative studies show that each of the methods has deficits and it is necessary to employ a combination of the methods in order to conduct a reliable accessibility assessment. The different rankings of ministerial web sites in various evaluations have confirmed this view.
4.3. Conclusion

In this chapter an attempt has also been made to review the articles on use of Internet and web resources by the botany professionals. It is observed from the review of literature that large numbers of literature are from abroad. From late 1990's considerable number of articles are contributed from Indian researchers also.
References


46. Kanaujia (Shiva) and Sathyaranayana (N.R.) Op.Cit. PP587-593.


Chapter 4

Review of Literature


