

## **Similarities and differences between Web search procedure and searching in the pre-web information retrieval systems**

**[Yazdan Mansourian](#)**

Ph.D. Student of Information Studies, University of Sheffield, UK

*Received June 5, 2004; Accepted August 16, 2004*

---

### **Abstract**

*This paper presents an introductory discussion about the commonalities and dissimilarities between Web searching procedure and the searching process in the previous online information retrieval systems including classic information retrieval systems and database. The paper attempts to explain which factors make these two groups different, why investigating about the search process on the Web environment is important, how much we know about this procedure and what are the main lines of research in front of the researchers in this area of study and practice. After presenting the major involved factor the paper concludes that although information seeking process on the Web is fairly similar to the pre-web systems in some ways, there are notable differences between them as well. These differences may provide Web searcher and Web researchers with some opportunities and challenges.*

### **Keywords**

*Web information retrieval, Web searching, Information retrieval methods*

---

### **Introduction**

Web searching is becoming an indispensable part of the daily life of many groups of people. Web usage is not limited to a few specific purposes and its functionalities are spreading across different aspects of life. Among all web-based applications, Web searching is one of the most common and important ones. However, satisfying the information needs on the Web is not always an easy and straightforward process. Web searching might be very time-consuming and may culminate in unsatisfactory results as well.

Information retrieval through the Web is a multifaceted process and many side elements may affect on this procedure. Nevertheless, our current knowledge about the complexity of this course of actions is limited. In spite of all remarkable progress in developing more sophisticated information retrieval tools on the Web, people may still encounter many difficulties for locating information they need for meeting their information needs. Although developing web-based

search tools has been inspired by pre-web online information retrieval systems, a couple of unique features of the Web environment and the wide range of Web users' characteristics impose new condition for searching and retrieval situation.

## **Pre-web information retrieval systems and their users**

In general, not only resources but also users in the pre-web systems were fairly homogenous and predictable. The electronic resources including online and offline databases, OPACs (Online Public Access Catalogues) contained mainly structured data which has been stored in very well-organised systems. Each document had a specific structure with a number of fields and subfields. This structured environment made the storage and retrieval procedure much easier and more predictable. Accordingly, information resources in the classic pre-web online information systems were well organized and homogenous. On the other hand, the users were limited to some specific groups of the society mostly the academics and researchers, librarians or subject expert people. However, after emergence of the Web everything became different which will be explained in the next section.

## **What happened after the Emergence of the Web?**

The main question is why the emergence of the Web revolutionized the search process considerably. In summary, four main reasons in the recent years are involved in the creating new search environment.

The first one is the surprising and uncontrolled expansion of electronic information in a very uneven, heterogeneous and unstructured environment. Because of the easiness and publicly accessibility of web-based publication facilities every moment new pages appear on the Web without any control on the contents, structure and quality of the information. Consequently, the Web network is a highly dynamic and uncontrolled environment. This situation considerably impacts on the procedure of storage, searching and retrieval of information on the Web.

The second outstanding event is the lack of a logical coincidence between the advance of computer technology in terms of technical sophistication and human computer interaction. Previous investigations about the information seeking and retrieval were mainly focused on the technical aspects and the human and users' side has been neglected in some ways.

The third issue is the rising number of inexperienced users. If previously computer-based equipments were mainly employed by specialist groups, now the majority of the public need to have interaction with different computer-based implements. Previous research indicates that unfamiliarity with search tactics creates difficulties for many users of online information retrieval systems ([Haverkamp & Gauch, 1998](#)). As [White & Iivonen](#) (2001) concluded in their report the majority of Web users are not very sophisticated searchers: preferring known sites, browsing, using only simple searches if they use a search engine, substituting words but not necessarily using other tactics to modify a search, being very trusting about information on the Web, and sometimes being relatively non-discriminating in recognizing relevant information.

The fourth issue is related to the level of understanding about the users' features

and needs while they are searching the Web. In fact, one of the most current important problems goes back to this reality that in spite of the importance of the issue of Web searching, with a user-centred viewpoint, our current knowledge about the information seeking on the Web is still inadequate. Accordingly, search tool designers need to learn more about the real user's information seeking behaviours to develop their products based on their needs. As [Cothey](#) (2002) has mentioned ". . . we need to understand the phenomenon of Web searching more fully to make information provision more effective. "

Regarding all these realities Web searching procedure is not in a satisfactory situation now and there are many issues which should be addressed through future research and then implications of these research findings into the next generation of web-based search tool design.

Until now many interesting investigation have been carried out particularly after year 2000. The literature indicated there is a vital necessity for making reasonable links between the technical-oriented and user-oriented studies about different aspects of the web-based information seeking and information retrieval studies. The importance of this interconnected links goes back to this reality that there is an ever-increasing growth of the diversity in the web-based information resources and Web searchers.

The Web is not a specific media for a small group of people and everybody who may have access to preliminary facilities for connection to the Web could be a potential user. For this reason, after e-mail which is the most popular Internet based application, searching for information on the Web is the second most popular activities in all web-based applications ([Hsieh-Yee](#), 2001). However, as it has been mentioned sometimes looking for an information item on the Web is an easy and pleasant experience. People might become surprised by the efficiency of search engines in finding their desired information so quickly. Nonetheless, sometimes on the contrary they may feel frustrated after a long search procedure. In this situation, it is not easy at all and answering a simple question through the search engines takes a long time or may remain unanswered. In this case people may get confused why search engines are not capable to do such apparently easy tasks. In order to illuminate the current situation in further details it would be useful to review the major information retrieval ways through the Web.

## **Common information retrieval methods on the Web**

There are different methods for information retrieval through the Web environment however; the most common way for information retrieval in this environment is searching through different web-based search facilities including various search engines. [Dennis et al.](#) (2002) have classified web-based searching process into four categories or four paradigms as follow:

1. *Unassisted Keyword Search*, as the most common and straightforward way. When users just entered one or more search term in a general-purpose search engines like Google and search engine retrieves a list of ranked documents.
2. *Assisted Keyword Search*, when search engines provide users with some facilities to expand their initial queries by search engine's suggestion for example in recent version of AltaVista.
3. *Directory-based Search*, when in addition of query based search people can

locate their information needs through browsing hierarchical categories like Yahoo.

4. *Query-by-example* when, users after a search session will be provided with a list of document summaries which is used as the basis of a new query. They also mentioned that most web-based search technology combines more than one of these paradigms. Usually majority of people use search engines to satisfy their information needs.

## Web searching: a holistic view

Search procedure on the Web is a complex course of action and relatively or sometimes entirely different to search procedures in previous online environments. The search process does not begin by typing a search term or a well-defined or unstructured query into search engines and does not come to an end by retrieval of a few relevant documents or avalanche of irrelevant retrieved items.

In addition, the search process for the majority of people does not usually happen just once and then finish particularly for those who are gathering information about a specific topic in a comprehensive approach. Generally, they have to delve into the Web frequently over a period of time. Sometimes it can be a time consuming and possibly frustrating task. Although you may find almost everything on the Web, searching the Web does not go well always and is not all the time successful. There is a potential ability for the Web to answer a given question. However, this is possible that a search process would not be successful. When it happens we can look at this topic from different perspectives. In fact, when somebody can not satisfy his/her information need on the Web there are some possibilities. These possibilities can be summarized as follow:

A. The required information does not exist on the Web. In spite of this reality that the Web is huge and its size is increasing steadily, nobody claims that everything is on the Web. A considerable portion of human knowledge is still in printed format and the other media and there is not the electronic format of them.

B. The required information exists on the Web but because of any reason the user can not find it. Exploring the possible reasons of the unsuccessfulness for locating an information item on the Web is a challenging issue and we can come from many possibilities. The summary is as follow:

- B1. The abilities of current search facilities are not sufficient to locate the information need. This is a very broad area of research and people who are working in information retrieval studies focus on this issue.
- B2. The user's search experience is limited and s/he does not utilise the existing search facilities probably. Therefore, if an experienced or trained user had searched for this given information, it could be found.
- B3. The user's knowledge domain is limited and s/he does not use the correct search term to search.

As [Botluk](#) (2000) mentioned sometimes a few success in quick retrieval of information items through the general search engines might lull people into a false sense of security that these search tools open up the world of information for them. She compared the databases on the Web as the books in the library they contain the relevant information but people need to know where to look.

## Conclusion

This paper concludes the information seeking process on the Web is fairly similar and notably different with classic searching in the classic information retrieval systems or pre-web online systems. It is similar because both are computer based and dealing with electronic information. However, it is considerably different because potential users and information sources in the Web environment are very different with pre-web online information systems.

Information resources on the Web are so heterogeneous, assorted and usually highly unstructured. In contrary, information stored in classic and pre-web online information systems are well organized and homogenous.

This difference is the same about the users of these two groups. Presumably, if not all, the majority of the users of classic and pre-web online information systems have been mainly academics and experienced users while everybody with different computer skills, subject knowledge, and search experience. In contrary, Web users' community includes almost everybody with different search experiences, knowledge, information needs and search patterns.

Regarding all these similarities and differences between Web searching procedure and pre-web search process this paper concludes that in Web searching domain we can utilise the knowledge body, research methodology and research approaches belong to the pre-web time but it is necessary to take all new elements of the Web environment into consideration. In the other word, while Web search research body is being built upon the previous investigations, this new line of enquiry is constructing new independent area of research.

## References

- Botluk, D. (2000, November 15). Mining deeper into the Invisible Web. Retrieved 12 August, 2004, from <http://www.llrx.com/features/mining.htm>
- Cothey, V. (2002). A longitudinal study of World Wide Web users' information-searching behaviour. *Journal of the American Society for Information Science and Technology*, 53 (2), 67-78.
- Dennis, S., Bruza, P., & McArthur, R. (2002). Web searching: a process-oriented experimental study of three interactive search paradigms. *Journal of the American Society for Information Science and Technology*, 53, (2), 120-133.
- Haverkamp, D.S., & Gauch, S. (1998). Intelligent information agents: review and challenges for distributed information sources. *Journal of the American society for information science*, 49(4), 304-311.
- Hsieh-Yee, I. (2001). Research on Web search behaviour. *Library & Information Science Research*, 23 (2), 167-185.
- Wang, P., Hawk, W.B. & Tenopir, C. (2000). Users' interaction with World Wide Web resources: an exploratory study using a holistic approach. *Information Processing and Management*, 36 (2), 229-251.
- White, M.D., & Iivonen, M. (2001). Questions as a factor in Web search strategy. *Information Processing and Management*, 37, 721-740.

Mansourian, Y. (2004). "Similarities and differences between Web search procedure and searching in the pre-web information retrieval systems".

*Webology*, **1**(1), Article 3. Available at:  
<http://www.webology.ir/2004/v1n1/a3.html>

---

**Alert us when:** [New articles cite this article](#)

---

Copyright © 2004, Yazdan Mansourian.