Beyond Logs and Surveys:
In-Depth Measures of People's Web Use Skills [1]

Eszter Hargittai
research@eszter.com
Sociology Department
Princeton University

Abstract

Finding information on the Web can be a much more complex search process than previously experienced on many pre-Web information retrieval systems given that finding content online does not have to happen via a search algorithm typed into a search field. Rather, the Web allows for a myriad of search strategies. Although there are numerous studies of Web search techniques, these studies often limit their focus to just one part of the search process and are not based on the behavior of the general user population nor do they include information about the users. To remedy these shortcomings, this project looks at how people find information online in the context of their other media use, their general Internet use patterns in addition to using information about their demographic background and social support networks. This paper describes the methodology in detail and suggests that a mix of survey instruments and in-person observations can yield the type of rich data set that is necessary in order to understand in depth the differences in people's information retrieval behavior online.

INTRODUCTION

There are many studies that look at how people use information retrieval systems and, in particular, how people search for information on the Web (for a review of this literature see Jansen and Pooch 2001). The Web Use Project at Princeton University adds to the literature on information retrieval in the following three important ways: 1. it recognizes that with the Web, searching for information is no longer limited to entering search queries in a search engine, rather, there are numerous ways in which one can go about finding information and these ways may lead to different results and differences in the efficacy of the particular information retrieval technique used; 2. it considers the search patterns of users drawn from the general population instead of solely relying on people in the academic and information science communities for data; 3. it collects data not only on users' search activities but also on their use of other media for information retrieval, their demographics and their social support networks allowing for a study of their online skills in the context of their social attributes.

This paper describes a method that allows us to measure differences in people's skills with respect to Web use. First, I briefly discuss the current state of research on Web use for information retrieval and why the approaches in existing studies are not always suitable to gaining a refined understanding of the differences in how people locate content online and what explains these differences in people's information retrieval skills. Next, I discuss how this project remedies some of the limitations of existing studies. Then, I describe the project methodology including a discussion of the sampling technique, technical specifications of the project, the two survey instruments, and a detailed description of the in-person observation session of people's online search techniques. Finally, I discuss how the various parts of the data can be pulled together for analysis.

EXISTING RESEARCH ON WEB USE FOR INFORMATION RETRIEVAL

Scholars from many fields have explored how people use the World Wide Web for information retrieval from fields as diverse as computer science, economics and library sciences. Advertising and marketing specialists often refer to users as "consumers" emphasizing their particular interest in people's online actions, namely their shopping behavior (Jarvenpaa, Sirrka and Todd 1996, Bell and Tang 1998). These studies often analyze users' behavior on only one particular site as opposed to exploring users' overall online behavior. Moreover, their sole interest is in how people decide to make online purchases, what influences these decisions, and how much shopping people engage in.

Much work conducted in the human-computer interaction field also tends to concentrate on particulars. Researchers in this area analyze people's use of specific design features and distinct Web site layout (see, for example, the Special Issue on World Wide Web Usability of the International Journal of Human-Computer Studies (1997)). Furthermore, they also look at features of software programs to assess important usability issues (see, for example, Greenberg and Cockburn (1999) for a detailed discussion of the "Back" button on browsers).

Alternatively, computer scientists draw on large-scale aggregate logs about people's Web use by analyzing all Web activity over a specified period (Catledge and Pitkow 1995, Huberman et al 1998.). An important limitation of many such studies is that they concentrate on the behavior of a small segment of the population by limiting participants to university faculty and students (e.g. Catledge and Pitkow 1995) or long-term users from the information technology profession (e.g. Choo, Detlor and Turnbull 1999). Although concentrating on such groups may be informative for particular research questions, such sampling techniques limit the extent to which findings can be generalized to a larger segment of the Web user population. In cases where data are derived from larger segments of the online population (e.g. Huberman et al. 1998, Hoelscher 1998, Silverstein et al. 1999, Jansen, Spink and Saracevic 2000), no information is available about specific users and thus it is impossible to make any claims about how attributes of users may be related to their online behavior.

Private research corporations collect data on what sites people visit and how much time they spend on each page (e.g. MediaMetrix and Nielsen//NetRatings collect Web behavior information this way). However, such information is proprietary and does not include information on what users are actually looking for (if anything) and whether they are satisfied by the options presented to them on the screen. Although some have tried to develop more general models from these types of data (Goldfarb 2001, Sinai and Waldfogel 2001), these studies are based on assumptions about users' behaviors that cannot be verified. These data sets do not contain information on how users perceive what they see and how they make the particular choices in their linking behavior and search strategies.

Researchers in the library and information science community have also conducted numerous studies on people's use of library resources that are often increasingly run on Web-based applications. Abramson (1998) looked at how people used the Web at public access computers by recording logs of use via a computer connected to the machines she was observing. However, she only collected information about visited sites and time of day and week without any information about users. Numerous case studies exist on the implementation of specific search programs in libraries (e.g. Payette and Rieger 1997) but these also limit their scope to a distinct user base and Web search protocol or library interface. There are also many studies (Hsieh-Yee 1993, Koenemann and Belkin 1996, Siegfried, Bates and Wilde 1993) that look at searches performed on various information retrieval systems (pre-Web applications as well), however, they focus on the details of query specifics (e.g. number of queries in the data set, session length, query length, use of advanced search functions) without considering information about user demographics or other information retrieval practices of the users.

Closest to the methods presented in this paper are some of the in-person user studies that have been conducted by library and information science researchers. Wang, Hawk and Tenopir (2000) collected data by observing how respondents search for information specified by the research team. Their project generated synchronized video-audio data, which were then analyzed for detailed information about respondents' search techniques. However, as often is the case in such studies, the participants for the study were graduate students and faculty in an information science program. In order to gain a better understanding of how the general population is using the Internet, it is important to include people from beyond the academic community in such studies.

The methods used in the studies cited above provide important information for a baseline understanding of how certain people navigate particular parts of the Web. However, existing studies either limit their scope to specific user populations (e.g. IT professionals or people who go to libraries), do not collect background information about users, or look at use patterns on an aggregate level without collecting data about the specific goals of a Web session. The Web Use Project remedies these shortcomings by collecting information about all these attributes of users and their online actions concurrently in one study. The next section describes the methodology in detail.

COLLECTING IN-DEPTH DATA: STRUCTURED OBSERVATIONS AND INTERVIEWS

Sampling

In order to be able to generalize from the findings, it is important to conduct the study on a random sample of users. The Web Use Project looks at the online use patterns and skills of a randomly selected Internet user population. A random sample of residential addresses is obtained for Mercer County, New Jersey from Survey Sampling, Inc. and is checked against the National Change of Address Database maintained by the U.S. Postal Service. [2] Potential respondents are first contacted via postal mail. They are sent a letter explaining the project and requesting participation with a brochure that presents more details about the study. People are also pointed to http://www.webuse.org on the Web for more information and are given the option of calling/writing to the research center to schedule an appointment. A few days after the letters have been sent, the households are contacted by telephone. The eligible adult (i.e. Internet user adult over 18) with the next nearest birthday is selected in order to randomly sample from within the household [3]. Web users are defined as people who go online at least once every month for more than using email. Although this is a low threshold for including people in the study, it is used to ensure that low frequency users who are nonetheless familiar with the Internet are also included.

People are offered $40 for their participation, which they receive after the observation session. Respondents are asked to come to the research site on the university campus. [4] The respondents' email address is recorded and a time for the session is scheduled. Respondents are informed that they will receive a follow-up letter in the mail or an email message (based on their preference) for confirmation and with directions to the research site. The day before the study a reminder phone call is placed to the respondent.

Technical Specifications

Both a PC and a Mac are used for the study to allow for variation in people's computer experiences. The three most popular browsing software applications are all available on both machines; Internet Explorer, Netscape Communicator, and America Online. [5] The computers connect to the Internet on a T1/T3 university network line. The HyperCam (Hyperionics 2001) software program is used to record the observation sessions on the PCs. This program creates audio-visual files (.avi) of the activity on the screen accompanied by the respondents' comments. A similar program, SnapZPro

(Ambrosia Software 2001), is used on the iMac. Additionally, a program called Don't Panic (Panicware 2001) is used to erase the browser and URL history on each browser program so that each respondent starts out with a clean slate and is not influenced by previous users' actions. Files are stored in a password-protected directory of the university network that is backed-up nightly offering safeguards against data loss.

Survey Instruments

The in-person sessions start with a twenty-minute interview about basic Web use questions. This interview draws on the Internet Module of the General Social Survey (GSS) 2000 and the HomeNet project at Carnegie Mellon University (Kraut et al 1998). [6] The questionnaire presented to respondents in this study replicates sections of the GSS Internet module in order to allow for comparisons with a larger population of users. The questionnaire is administered verbally to establish a rapport between the researcher and the respondent. Administering the questionnaire right before the observation session proves to be very useful. Because the questions explore many facets of Web use, respondents are prompted to think about numerous details of their Web experiences before sitting down at the computer and embarking on the tasks presented by the researcher. After the observation session (described below), respondents fill out another questionnaire online. [7]

The types of data collected in the two surveys include information about the frequency and location of respondents' regular Internet use, the types of sites they visit, the types of activities they perform online, their use of other media, their time spent on various social activities and their social support networks. Details about the types of sites users visit are important to collect in order to have baseline information about the types of sites with which they are familiar. Someone who often visits political sites is likely to exhibit different browsing strategies while searching for such Web sources, not necessarily because of a general higher level of skill in searching

for information online, but because of prior experience with this particular type of task.

Finally, a long list of computer and Internet related terms are presented on the survey and respondents are asked to rank their understanding of these terms. A list of multiple-choice questions measuring people's actual knowledge of these computer and Internet related terms was added later in the study. The goal is to see whether the level of skill measured by analyzing people's actions online correlates with people's scores on these knowledge variables. Because the methods used in this study are time- and labor intensive and costly, a longer-term goal of this project is to suggest ways in which people's skills can be assessed via survey questionnaires instead of always relying on such elaborate data collection to assess users' online search skills.

Observation Sessions

Conducting the observations at a university research site has both advantages and shortcomings. Requesting users to come to a location affects response rates. It also places people in a location with which they are not familiar and requires them to use a computer that is configured differently from the machine they usually use for browsing. This may influence the results, as certain settings (e.g. the default homepage and bookmarks) are not equivalent to their own. However, this approach controls for quality of Internet connection, and hardware/software differences. It also allows us to concentrate on Web use knowledge in a setting that is equally different and new for all. Moreover, using one computer allows the setup of particular software applications that are required for data recording as described earlier. No default page is set on browsers in order not to influence respondents' initial actions once online. The session is started off by the researcher asking the respondent to recall – if possible – the default homepage on the computer she uses the most. The respondent is also asked to comment on how much the browser used in the study in front of her resembles the one she uses most

frequently. The respondent is also asked whether she has personalized anything on the browser and whether she has any bookmarks/favorites set.

Users are given a list of seventeen tasks to perform on the Web to see how they would find certain information online. These resemble the question on the GSS 2000 Internet Module (Q19) that asks how users would go about finding information about a political candidate. However, instead of the hypothetical question asked on that survey, the researcher is able to watch users go through the process of finding a page and take detailed notes on what they do. Respondents are encouraged to continue searching without giving up too easily. A minimum of five minutes is given for each task unless the respondent exhibits frustrations and expresses a need for moving on in which case the researcher reads the next task.

Although a recording of users' visits gives us information about what pages users see (as per the type of data collected by commercial marketing corporations and analyzed in some studies [Goldfarb 2001, Sinai and Waldfogel 2001]), it gives us no information on what type of information users are actually looking for and how satisfied they are with what they find. Moreover, most logs of uses do not record information about moves that concern the local cache of the machine. Consequently, these large-scale data sets miss information about details such as use of the "Back" button on browsers, which is a considerable problem given that its use comprises up to 30 percent of people's browsing activities (Tauscher and Greenberg 1997) and may be considered a part of one's level of search sophistication.

The task-oriented method is repeated for several types of sites such as arts, current events, volunteer organizations, shopping, music, health-related and job search services. Some tasks are fairly general (e.g. finding information about a political candidate or the contact information of a long lost friend) whereas others ask for very specific types of content (e.g. a Web site with a recipe explicitly stated as acceptable for someone with lactose intolerance). Presenting different types of tasks is important in order to gather

information on universal versus topic-specific search strategies. Someone who is universally skilled in finding information may have highly sophisticated skills in locating any type of information whereas topic-specific skills imply that the user has considerably different search skill levels depending on the topic being sought. An example of topic-specific search skills is someone who possesses sophisticated methods for finding Web sites on online music, but has little knowledge of how to arrive at Web pages with reliable information about a health concern. Some of the tasks were familiar activities to some respondents but not to others. However, some of the tasks were new to all respondents which allows for comparison across cases with respect to a formerly uncharted territory. [8]

During the session, the respondent is encouraged to make comments about her actions. She is asked if the actions she has been asked to perform are ones she has performed before. If the respondent gets enthusiastic about showing off a skill or search that she has recently engaged in, the researcher encourages her to do so even if this action is not directly related to the specific task at hand. The one constraint to keep in mind here is that all such additional information needs to be collected in between tasks. This is important so actions not directly related to the search do not contaminate the time-to-completion measure calculated for each task.

By talking to people, we learn more about their actions and motivations than if we were simply observing recordings of the pages they visit. In other words, this project is not only about studying people's sequence of use but also their search strategies, the underlying motivations of their actions, and their levels of satisfaction with their Web experiences.

CODING AND ANALYZING THE DATA

The audio-visual (.avi) files generated by the screen capture application are coded while being viewed with a multi-media program such as QuickTime or Windows Media Player. Hargittai (2000) developed an exhaustive – albeit not mutually exclusive – list

of ways in which one can arrive at a Web site. This list is used as the basis for coding people's online actions and includes information about whether the user: a. accessed a URL directly (e.g. by typing it into the location bar); b. performed a search in a search engine; c. followed up on the results of a search; d. clicked on a category directory or Web guide; e. clicked on an advertisement; f. simply clicked on a link on a page. Additional information is collected on whether the user took advantage of various browser features such as the use of auto-complete forms, particular uses of the Back button or History file, and the use of URL truncation in the case of missing pages. Regarding open searches, all search terms are tallied as are the particular types of search results users pursue (whether they are "recommended" sites, "matching" sites or "sponsored" links). Additionally, the URL of each visited Web page is recorded , as is the exact time of every action. The sessions are also coded with particular emphasis on whether a user was successful in completing a task and if yes, how long she took to do so.

The audio component of the file is transcribed and the content is coded for additional information about users' online experiences. Since exact replication of the users' regular Web use environment is not possible in this study, users' comments are used as a source of additional information about their everyday strategies. The interview component reveals the following types of information about users: whether users depend on social support networks for assistance (e.g. do they ask for help when looking for particular content and from whom, do they get site recommendations from others); where users learn about sites (e.g. advertisements in other media or how-to articles); how users feel about their online experience (e.g. frustrations with particular sites or tasks, confusion about various Web features); what exactly users look at when browsing the results of a search (e.g. do they know how to read a URL and do they look at that for clues about the search results); and whether users know who put up a site and whether they can assess its reliability. We know such information by having prompted the respondents to talk

through their actions and by asking specific questions after search sessions.

Finally, information about search strategies is matched with how easily and quickly the respondent was able to find the desired information. Again, it is important to note that while someone can be extremely skilled at locating one type of information, they may have fewer skills in locating a different type of information. For this reason, both topic specific and universal search strategies are assessed. Finally, the data obtained from analyzing the search sessions are merged with data from the surveys and can be analyzed together.

FINDINGS

Preliminary results suggest that there is large variance in the amount of time people take to complete all seventeen tasks ranging from twenty minutes to over 100 minutes. Most people are eventually successful in locating most of the requested content although some fail in succeeding with as many as half of the tasks.

An interesting finding of this study is the extent to which members of the general user population lack the basics of surfing the Web. A few people barely know what a Back button is and thus have an incredibly hard time moving from screen to screen. Many people rarely use search engines and solely rely on functions of their browsers or Internet service providers. Some respondents also have a hard time entering valid search terms including the common occurrence of spelling mistakes. One recurring mistake is entering multiple term queries without any spaces. When asked about this practice, several users have replied that you are not supposed to use spaces on the Web thus the exclusion of spaces in between search terms. [9] Others' exhibit the exact opposite behavior by typing search terms in the location bar itself. However, given that most browsers now automatically redirect those terms into a search engine, this seems to cause less confusion and fewer dead-end sessions.

More knowledgeable users vary their strategies depending on the type of task. They alter the use of open-ended searches with browsing in category directories compiled by large sites. Knowing some of the intricacies of how to use a search engine can be extremely valuable as well (e.g. use of Boolean operators). People who recognize the value of typing in more than one search term especially in the case of a complex search have a much easier time finding sites that address their queries. Moreover, understanding how search engines rank pages, and being able to read search results (including the URLs of the results) can be quite valuable.

CONCLUSION

Overall, there is great variance in how long people take to find content online and whether they are successful in the first place. The methods presented in this paper allow us to focus on the Web as a complex set of information retrieval services instead of only exploring a single aspect of its search features. Moreover, given that many users from the general population do not know about or do not use search engines, it would be a mistake to restrict all studies on searching to that one search technique. A mix of different methods – survey instruments coupled with in-person observations – leads to the type of rich data set that allows us not only to understand people's very diverse set of search strategies but also explore what social factors explain the differences in their actions. Focusing on the general population instead of particular specialized groups allows us to generalize our findings to the larger user population. Due to these features of the research design, the findings from such a study can be used more broadly to inform training programs to educate users about efficient Web use techniques, and to suggest ways in which online content organization can be improved to facilitate users' access to information.

NOTES

1. I would like to thank Paul DiMaggio for his insightful comments throughout this project and Stan Katz for his ongoing support. Barbara Wildemuth and the anonymous reviewers from the ASIST SIG USE Research Symposium offered very helpful suggestions. I am also grateful to Edward Freeland, James Chu and Jeremy Davis-Turak for their help with the survey components of the project, to Carolyn Mordas for help with recruiting and to Inna Barmash for her help with interviews and coding the data. Generous support from the Markle Foundation is kindly acknowledged. The project has also been supported in part by NSF grant #SES9819907, a grant from the Russell Sage Foundation, and through a grant from the Pew Charitable Trusts to the Center for Arts and Cultural Policy Studies, Princeton University. I am also grateful to the Fellowship of Woodrow Wilson Scholars at Princeton University.
   A similar version of this paper was presented at the Telecommunications Policy Research Conference on October 29, 2001 in Alexandria, VA.
2. This study is also being conducted in other locations where the same sampling methodology is used.
3. If this randomly selected person from the household is not willing to participate then the household is coded as a refusal even if another member of the household would have been willing to take part in the study. Such strict measures of random selection assure that the participants of the study represent a truly random sample of the Internet user population. Households without any Internet users are not eligible for the study. This study only includes adult English-speaking users. Two follow-up studies are already being planned for Spanish-speaking users and high school students.
4. Respondents are offered transportation if they cannot provide their own.

5. The research project has its own AOL account in order not to burden respondents' accounts with the time spent at the session.

6. The GSS is conducted every few years on a random sample of the American population with a response rate (70-80 percent) rarely achieved by other surveys. The GSS interviews are conducted face-to-face with people in their homes. For decades, the core section of the questionnaire has been replicated on every survey allowing for time-series analyses about people's political beliefs and social attitudes in addition to a myriad of other attributes.   The GSS also contains topical modules that differ from year to year. In Spring 2000, a twelve minute Internet Module was added to ask people about their Internet use at home, work, school and other locations (e.g. libraries). Questions were asked about what online services people use, what types of sites they visit, and how they use the Web for political and cultural activities. The mean response time of the GSS Internet module was 12:26 and ranged from zero to 45 minutes. The HomeNet Project has been administered since 1995 and contains numerous questions about people's social well-being.

7. An online questionnaire allows for convenient automatic coding of responses. This online survey uses the Princeton University Survey Facility, which is an application available to members of the Princeton University community for administering Web surveys (http://www.princeton.edu/~jkchu/Survey/).

8. Two tasks that seemed new to most people are: 1. searching for a site that compares presidential candidates views on abortion; and 2. looking for a page that displays art by children.

9. Respondents were asked about some of their online actions after the full search session had been completed to make sure that the rest of their search behavior would not be influenced by the researcher's question.

REFERENCES

Abramson, A.D. (1998). Monitoring and Evaluating Use of the World Wide Web in an Academic Library: An Exploratory Study. *Proceedings of the 61st Annual Meeting of the American Society for Information Science, Vol. 35*, 315-326.

Ambrosia Software. (2001). SnapZPro. http://www.ambrosiasw.com/

Bell, H., & Tang, N.K.H. (1998). The effectiveness of commercial Internet Web sites: a user's perspective. *Internet Research, 8*(3): 219-228.

Catledge, L.D, & Pitkow, J.E. (1995). Characterizing browsing strategies in the World-Wide Web. Paper presented at the Third International World Wide Web Conference (Darmstadt, Germany, April 10-14).

Choo, W.C, Detlor, B., & Turnbull, D. (2000, February 7). Information seeking on the Web: an itegrated model of browsing and searching. *First Monday, 5*(2). http://www.firstmonday.dk/issues/issue5_2/choo/index.html.

Goldfarb, A. 2001. Analyzing Website Choice Using Clickstream Data. Paper presented at the 29th annual Telecommunications Policy Research Conference, Alexandria, VA.

Greenberg, S., & Cockburn, A. (1999). Getting back to back: alternate behaviors for a Web browser's back button. Proceedings of the 5th Annual Human Factors and the Web Conference (NIST, Gaithersburg, Maryland, USA, June 3, 1999). http://www.cpsc.ucalgary.ca/grouplab/papers/

Hargittai, E. (2000). Open portals or closed gates? Channeling content on the World Wide Web. *Poetics, 27*: 233-253.

Hoelscher, C. (1998). How internet experts search for information on the Web. Paper presented at the World Conference of the World Wide Web, Internet, and Intranet, Orlando, FL.

Hsieh-Yee, I. (1993). Effects of search experience and subject knowledge on the search tactics of novice and experienced

searchers. *Journal of the American Society for Information Science, 44*(3): 161-174.

Huberman, B.A., Pirolli, P.L., Pitkow, J.E., & Lukose, R.M. (1998). Strong regularities in World Wide Web surfing. *Science, 280*(5360): 94-97.

Hyperionics. 2000. HyperCam. http://www.hyperionics.com/

International Journal of Human-Computer Studies. (1997). Special Issue on World Wide Web Usability. http://www.hbuk.co.uk/ap/ijhcs/webusability

Jansen, B.J., & Pooch, U. (2001). A review of Web searching studies and a framework for future research. *Journal of the American Society for Information Science and Technology, 52*(3): 235-246.

Jansen, B., Spink, A., & Saracevic, T. (2000). Real life, real users, and real needs: a study and analysis of user queries on the Web. *Information Processing and Management, 36*(2): 207-227.

Jarvenpaa, S.L., & Todd, P.A. (1996). Consumer reactions to electronic shopping on the World Wide Web. *International Journal of Electronic Commerce, 1*(2): 59-88.

Koenemann, J., & Belkin, N. (1996). A case for interaction: a study of interactive information retrieval behavior and effectiveness. Paper presented at Conference on Human Factors in Computing Systems. Vancouver, Canada.

Kraut, R., Patterson, M., Lundmark, V., Kiesler, S., Mukopadhyay, T., & Scherlis, W. (1998). Internet Paradox. A Social Technology that Reduces Social Involvement and Psychological Well-being? American Psychologist, 53 (9), 1017-1031.

Panicware, Inc. (2001). Don't Panic 4.0. http://www.panicware.com.

Payette, S.D., & Rieger, O.Y. (1997, April). Z39.50: the user's perspective. *D-Lib Magazine.* http://www.dlib.org/dlib/april97/cornell/04payette.html.

Siegfried, S., Bates, M., & Wilde, D. (1993). A profile of end-user searching behavior by humanities scholars: the Getty Online Searching Project Report No.2. *Journal of the American Society for Information Science, 44*(5): 273-291.

Silverstein, C., Henzinger, M., Marais, H., & Moricz, M. (1999). Analysis of a very large Web Search engine query log. *SIGIR Forum, 33*(1): 6-12.

Sinai, T & Waldfogel, J. 2001. Geography and the Internet: Is the Internet a Substitute or a Complement for Cities? Paper presented at the 29th annual Telecommunications Policy Research Conference. Alexandria, VA.

Tauscher, L., & Greenberg, S. (1997). How people revisit Web pages: empirical findings and implications for the design of history systems. *International Journal of Human Computer Studies, 47*(1): 97-138.

Wang, P., Hawk, W.B., & Tenopir, C. (2000). Users' interactions with World Wide Web resources: an exploratory study using a holistic approach. *Information Processing & Management, 36*(2): 229-251.