World Wide Web

Information Seeking

Don Turnbull
School of Information
University of Texas at Austin
Austin, TX 78712
donturn@ischool.utexas.edu

A significant amount of the contributions of information seeking theory and models are based on assumptions or smaller-scale studies of user information needs and searching. (N. J. Belkin, N.Oddy, & Brooks, 1982; N. J. Belkin, Oddy, & Brooks, 1982; Kuhlthau, 1991) These models suppose cognitive activity and the feedback given by the information search environment. With the more recent focus on World Wide Web browsing and searching, a new set of data collection and user modeling studies are possible. While these new Web-based models can be primarily based on detectable patterns in either individual interest (content) or situational activity (context), it is hoped that it may be possible to generalize a set of habits from these patterns that can then be either integrated with existing information seeking models or to form a new basis for information seeking studies.

It is possible that by combining these previous studies with new data collection methods and accounting for newer information seeking technologies such as the Web, we may be able to discover, measure and compare information seeking and information retrieval patterns. Over the last few years Choo, Detlor and Turnbull have worked on a hybrid information seeking framework (based on Ellis 1989 and Aguilar 1967) which has potential to be extended or mapped to other information seeking activities such as information search. (Choo, Detlor, & Turnbull, 1998, 1999, 2000a, 2000b)

To extend this model, is it possible to add yet another model dimension that is more quantitative as seen in studies like Huberman, et al.'s "Law of Surfing" study of aggregate Web use or towards the Information Foraging work championed by Pirolli and Card that focuses more on individual users and their general information seeking behaviors? (Huberman, Pirolli, Pitkow, & Lukose, 1998; Pirolli & Card, 1995) Moreover, are these fruitful directions to move in light of the significant previous research into information seeking? Is there anything new to discover in information seeking except to apply existing models to Web environments?

This more behavioral-oriented research shows that habits and routines are dominant in terms of information (seeking) behavior and it may be useful to explore the efficacy of the above models towards more finitely describing user behavior, be it assumptions about cognitive activity or noting the impact of various information technologies. One technique increasingly employed in the recent years is the highly quantitative, Knowledge Discovery in Databases (which includes Data Mining) to analyze large-scale datasets of Web use activity. It would be useful to address the applicability of these techniques to either extend or confirm existing information seeking and use models or devise altogether new models.

One barrier to understanding such large datasets is in visualization methods to make sense of quantitative information seeking data including new metaphors for envisioning and presenting results or altogether new models now more possible due to advances in multi-dimensional or real-time visualization.(Card, 1996; Card, Mackinlay, & Shneiderman, 1999) One application that may have merit is taken from ideas relating to faceted classification, where different views of the same data may be viewed according to context, confirmation of habits, prediction, content identification and classification or by individual or aggregate use. Also, at a simpler level, since much information behavior is

repetitive and iterative, it may be possible to visualize both individual and aggregate information seeking sessions as a spiral model, where the user slowly gets to a centered "target" or solution to the information seeking problem at hand. Additionally, it would be useful to address and formalize a set of standards for data collection, analysis and presentation to make it possible to compare studies of Web information seeking. The current variety of presentation and (path) graphing formats, not to mention the variability of data collection methods and datasets makes any comparison and consensus about Web information seeking problematic. Past information seeking research focuses on two main directions: discernable types of user communities or insights into unique types of user behavior. Why not combine both directions to study Web information seekers and their behaviors as bounded by possible Web-based interactions as defined by the Web browser technology in use? KDD methods may prove to be the primary method to combine these two directions of study.

Recently, there has been an increased focus on modeling information retrieval activity, more specifically Web searching. As Web information retrieval technology improves (and browsing technology stands relatively still), users become habituated to conduct a search as the initial step in information seeking. By focusing first on Web searching, some of the above questions may be answered in an approachable manner and then be built upon for more complex, open-ended information seeking activities. These Web search studies show that users do have particular patterns of search use and yield insight into user mental models of search.(Jansen, Spink, Bateman, & Saracevic, 1998; Silverstein, Henzinger, Marais, & Moricz, 1998) It is possible that a system that can identify and track an individual user's search behavior, in comparison to aggregate patterns, and begin to build a profile of interests and habits can be used to personalize search results and interfaces to augment information seeking sessions. (Pitkow et al., 2002) This type of observation leads to an approach that has potential to both confirm or

leverage current information seeking models or help to unveil altogether new models, that may be either general (e.g. an information seeking episode) or contextual (e.g. eCommerce or content topic specific) to a particular user, a particular situation or as generally applicable to any information seeker. However, the true strength of these new avenues and methods may be to combine with existing information seeking models to provide glimpses into human information seeking combined with the World Wide Web information environment.

References

Aguilar, F.J., Scanning the Business Environment, Macmillan Company, New York, 1967.

Belkin, N. J., N.Oddy, R., & Brooks, H. M. (1982). ASK for Information Retrieval: Part I. Background and history. Journal of Documentation, 38(2), 61-71.

Belkin, N. J., Oddy, R. N., & Brooks, H. M. (1982). ASK for information retrieval: Part II. Results of a design study. Journal of Documentation, 38(3), 145-164.

Card, S. K. (1996). Visualizing Retrieved Information: A Survey. IEEE Computer Graphics and Applications, 16(2).

Card, S. K., Mackinlay, J. D., & Shneiderman, B. (1999). Readings in information visualization : using vision to think. San Francisco, Calif.: Morgan Kaufmann Publishers.

Choo, C. W., Detlor, B., & Turnbull, D. (1998). A Behavioral Model of Information Seeking on the Web - Preliminary Results of a Study of How Managers and IT Specialists Use the Web. Paper presented at the Proceedings of the 61st Annual Meeting of the American Society of Information Science, Pittsburgh, PA.

Choo, C. W., Detlor, B., & Turnbull, D. (1999). Information Seeking on the Web - An Integrated Model of Browsing and Searching. Paper presented at the Proceedings of the 62nd Annual Meeting of the American Society of Information Science, Washington, D.C.

Choo, C. W., Detlor, B., & Turnbull, D. (2000a). Web Work: Information Seeking and Knowledge Work on the World Wide Web. Dordrecht, The Netherlands: Kluwer Academic Publishers.

Choo, C. W., Detlor, B., & Turnbull, D. (2000b, Jan 4-7). Working the Web: An Empirical Model of Web Use. Paper presented at the 33rd Hawaii International Conference on System Science (HICSS), Maui, HI.

Ellis, D. "A Behavioural Model for Information Retrieval System Design", Journal of Information Science, 15(4), 1989, pp. 237-247.

Huberman, B. A., Pirolli, P. L. T., Pitkow, J. E., & Lukose, R. (1998, April 3, 1998). Strong Regularities in World-Wide Web Surfing. SCIENCE, 280, 95-97.

Jansen, B. J., Spink, A., Bateman, J., & Saracevic, T. (1998, Spring, 1998). Real Life Information Retrieval: A Study of User Queries on the Web. SIGIR Forum: A Publication of the Special Interest Group on Information Retrieval, 32, 5-18.

Kuhlthau, C. C. (1991). Inside the Search Process: Information Seeking from the User's Perspective. Journal of the American Society of Information Science, 42(5), 361-371.

Pirolli, P., & Card, S. (1995). Information Foraging in Information Access Environments. Paper presented at the Conference on Human Factors in Computer Systems, CHI-95, Denver, Colorado, USA.

Pitkow, J., Schutze, H., Cass, T., Cooley, R., Turnbull, D., Edmonds, A., et al. (2002). Personalized Search: A Contextual Computing Approach May Prove a Breakthrough in Personalized Search Efficiency. Communications of the ACM, 45(9), 50-55.

Silverstein, C., Henzinger, M., Marais, H., & Moricz, M. (1998). Analysis of a Very Large Web Search Engine Query Log. SIGIR Forum: A Publication of the Special Interest Group on Information Retrieval, 6-12.