Speaking the Same Language About Exploratory Information Seeking

Bill Kules
School of Library and Information Science, The Catholic University of America
Washington, DC USA
kules@cua.edu

1. INTRODUCTION

"Exploratory search" – little did I realize when I started my dissertation six plus years ago that a term that I naïvely thought my advisor and I had coined (we hadn't) would become a central element of discussion for an NSF-funded workshop. As a computer science Ph.D. student just beginning to learn about the field of library and information science, the notion of exploratory search seemed to capture the kinds of tasks that I was investigating. Of course, this was hardly the only term in use – it wasn't even the most common. As I read more, I realized that there was a broad and deep literature on the topic, but one which contained myriad terms, definitions, and models. The lack of common language presented a challenge as I tried to understand and apply that rich body of knowledge to my own work in faceted search interfaces, and it remains a challenge today.

This paper advocates the adoption of a common term (I suggest "exploratory information seeking"), as well as a common definition and reference model. I make this proposal at the risk of sounding naïve – perhaps it reflects too little knowledge of the rich history of the field; maybe it suggests oversimplifying or "dumbing down" the field to fit a least-common denominator approach; or it may simply echo a consensus already evidenced by the existence of this workshop. But I believe that the field is at a point where these three things are necessary in order to make the research and knowledge accessible to a larger audience.

We need to communicate with multiple audiences: researchers working directly in this field (many of whom are at this workshop); researchers outside of the LIS field (in areas such as computer science, human-computer interaction, psychology, management, and the humanities) who can benefit from and contribute to this work; students who are learning about the field; and the funding agencies, journalists, and general public that need to understand how our individual efforts fit into and advance a larger research agenda. A common framework will also support meaningful dialog about systems and evaluation. This need has been acknowledged at the workshops I have previously attended. In at least two instances, promising discussions were slowed by the lack of a common framework and the need to (re)establish a common understanding among participants.

2. TERMS AND DEFINITIONS

To start, we need to select a term and define it with care. The term should provide an immediate indication of the topic. In addition, we need a short definition that captures the essence (an "elevator speech"), and a long definition that more fully delineates the contexts, situations, tasks and techniques.

The term "exploratory search" is problematic, in part, because of its use of the word "search." Search has a variety of meanings and connotations, only some of which are relevant to the notion of exploratory search as I understand it. Past discussions at workshops have often spent a fair bit of start-up time trying to establish a working definition. They often stumble because there are many behaviors included under exploratory search that are not strictly search – browsing being a common example.

As a starting point for conversation, the term "exploratory information seeking" begins to capture the kinds of work that we are trying to support.

3. TASK MODEL

The second challenge is to develop a common reference model for tasks. This is important for the same reasons that a common definition is needed – we need to communicate effectively within and beyond our community. Of course, no single model will effectively capture the breadth of elements, concepts, variables, and relationships that have been identified. Rather, it should provide a rough "map" of the field that enables researchers to situate and elaborate on their work. As with the short definition, the model should be simple enough to be quickly comprehended.

The TREC conferences provide historical motivation for a common framework. Their system-oriented and quantitative focus has been criticized, but the common framework enabled dialog among researchers, development of tasks for evaluations, and comparison of systems. Over the years, they have carefully defined, refined, and examined specific information retrieval tasks, contributing to the robust development of the IR community. Our challenge is to adapt that idea to exploratory information seeking research without foreclosing on the correspondingly broader set of research methods that are required.

A particular challenge is to develop a taxonomy of exploratory information seeking tasks that span multiple levels of context. Because context and high-level goals are important, the tasks will span a range from individual tactics and actions within a specific interface to high-level scenarios such as writing a research paper or conducting competitive intelligence analysis. Inherent in this challenge is the need to accommodate both quantitative and qualitative methods. It may not be possible to directly compare data between a controlled laboratory study and an ethnographic investigation in the field, but a common understanding of the tasks will help us to triangulate between a varying study designs and compare evaluations in a systematic manner. This in turn will enable meta-evaluations that integrate the results of individual research projects into a larger understanding.

4. A SAMPLE REFERENCE MODEL

This section presents a three-level model of the exploratory information seeking process as a partial illustration of the proposals in this paper (Figure 1). This specific model does not capture the diversity of research interests represented by

workshop attendees, but it illustrates a simple framework that can be used for situating discussions and comparing research.

This model combines Marchionini's electronic browsing model (Marchionini, 1995) with the three-level Byström & Hansen model. This model defines five activities: recognize an information need (to satisfy a work task), define an information-seeking problem (to satisfy the information need), formulate query, examine results, and view documents. It places activities in the context of the three levels of information-seeking and work tasks. It shows how search activities are sequenced within the iterative search process. Each higher-level activity can involve multiple subsidiary activities.

Search is a necessary step within a larger information seeking process, the objective of which is to satisfy a perceived information need or problem (Marchionini, 1995). In turn, the perceived information need is motivated and initiated by a higher-level work task (Byström and Hansen, 2002; Ingwersen & Järvelin, 2005) or personally motivated goal (Kari, 2006).

Work tasks are situated in the work organization and reflect organizational culture and social norms, as well as organizational resources and constraints. The work task is similar to Marchionini's recognition and acceptance of an information problem, but the work task specifically situates these in an organizational context. In the context of the work task, a second level of context is defined, in which information-seeking tasks are identified. These tasks vary as the work task progresses. The third level of context is the information retrieval context, wherein searchers identify sources, issue queries, and examine results. Reflection is inherent in each activity, and each activity can return to a previous or higher level activity.

The model is limited in a number of ways. For example, it is focused on work tasks; it considers a limited set of search behavior; and it does not account for non-IR tasks that are part of information seeking. Our challenge is to integrate the many models and theories already available (including Bates, 1972, 1989, 1990; Belkin, 1980; Choo, Detlor & Turnbull, 2000; Dervin & Nilan, 1986; Ellis, 1989; Fidel, 1985; Furnas & Rauch, 1998; Pirolli & Card, 1999, Saracevic, 1996) and summarize them in a manner that is meaningful to multiple audiences.

5. CONCLUSION

This paper proposes the adoption of a common term, definition and reference model for exploratory information seeking. It provides an example of a simple reference model as a straw man for discussion. One inherent challenge is to capture the breadth and depth of the field without oversimplifying it. This will benefit the field by providing a common frame of reference for discussion of systems and evaluation. It will also make the research and findings more accessible to researchers, practitioners, students, and funders – anyone who can benefit from and contribute to the field.

6. REFERENCES

[1] Bates, M. (1990). Where should the person stop and the information search interface start. *Information Processing and Management, 26*(5), 575-591.

[2] Bates, M. J. (1979). Information search tactics. *Journal of the American Society for Information Science, 30*, 205-214.

[3] Bates, M. J. (1989). The design of browsing and berrypicking techniques for the on-line search interface. *Online Review, 13*(5), 407-431.

[4] Belkin, N. J. (1980). Anomalous states of knowledge as a basis for information retrieval. *Canadian Journal of Information Science, 5*, 133-143.

[5] Byström, K., & Hansen, P. (2002). Work tasks as units for analysis in information seeking and retrieval studies. In H. Bruce, R. Fidel, P. Ingwersen & P. Vakkari (Eds.), *Emerging Frameworks and Methods* (pp. 239-251). Greenwood Village, CO: Libraries Unlimited.

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

[7] Dervin, B., & Nilan, M. (1986). Information needs and uses. In M. Williams (Ed.), *Annual Review of Information Science and Technology* (Vol. 21, pp. 3-33). White Plains, New York: Knowledge Industries.

[8] Ellis, D. (1989). A behavioral model for information retrieval system design. *Journal of Information Science, 15*(4-5), 237-247.

[9] Fidel, R. (1985). Moves in online searching. *Online Review, 9*(1), 61-74.

[10] Furnas, G. W., & Rauch, S. J. (1998). Considerations for information environments and the NaviQue workspace. In *Proceedings of the Third ACM Conference on Digital libraries, Pittsburgh, PA* (pp. 79-88). New York: ACM Press.

[11] Ingwersen, P. & Järvelin, K. (2005). *The Turn: Integration of Information Seeking and Retrieval in Context*. Dordrecht, The Netherlands: Springer.

[12] Kari, J. (2005). Evolutionary information seeking: A case study of personal development and Internet searching. *First Monday, 11*(1). http://www.firstmonday.org/issues/issue11_1/kari/index.html

[13] Kules, B., Wilson, M., schraefel, m.c., Shneiderman, B. (in process) From Keyword Search to Exploration: How Result Visualization Aids Discovery on the Web.

[14] Marchionini, G. (1995). *Information Seeking in Electronic Environments*: Cambridge University Press.

[15] Pirolli, P. L., & Card, S. K. (1999). Information foraging. *Psychological Review*, 106(4), 643-675.

[16] Saracevic, T. (1996). Relevance reconsidered. In *Information science: Integration in perspectives. Proceedings of the Second Conference on Conceptions of Library and Information Science (CoLIS 2)*. Copenhagen (Denmark), 14-17 Oct.1996. pp. 201-218.
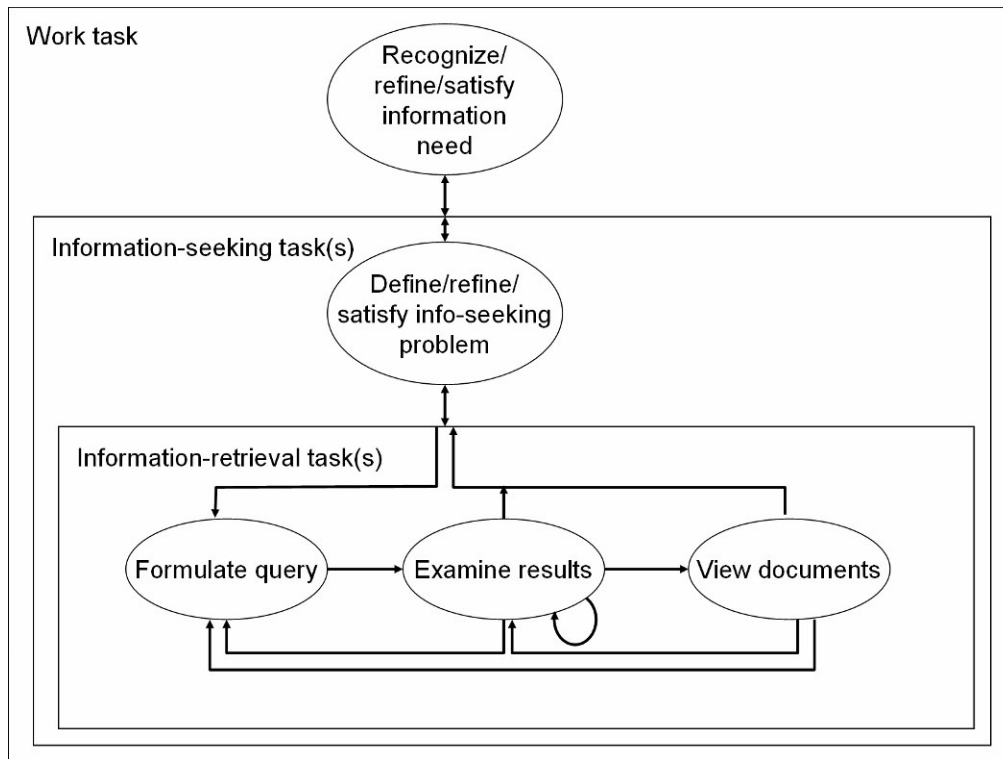
Figure 1. A three-layer model of exploratory information seeking, showing the context of work and information-seeking tasks.