

Methods for evaluating changes in search tactics induced by exploratory search systems

Bill Kules

Takoma Software, Inc.
Takoma Park, MD 20912

wmk@takomasoftware.com

ABSTRACT

Mixed research methods provide a way to understand how exploratory search systems change search tactics and strategies. Longitudinal studies may help understand how searchers adapt search strategies and tactics in support of creative, challenging, individualized activities.

1. INTRODUCTION

Evaluation of exploratory search systems is an exciting research challenge. The situated nature of exploratory search tasks can lead to many different task outcomes for different searchers. This can make it difficult to specify objective performance measures like time to completion, error rates, precision, or recall. Completing an exploratory task often involves developing and refining an information need that is specific to the individual. Mistakes, dead-ends, and back-tracking are part of the process as searchers learn concepts and vocabulary. Documents that have great utility or novelty to one person may have little value to another, because of variations in domain knowledge, interests, and previously encountered information, so establishing ground truth for a measure of relevance is problematic.

The strategies and tactics that searchers use are affected by the capabilities provided by the search system [1, 2]. Designers build interfaces to support specific strategies, based on intuition or analysis. The effect of new capabilities on search tactics may not be what designers anticipate. Unexpected problems may negate expected benefits. Serendipitous possibilities may present to searchers. In response, searchers may adapt their tactics and strategies as they become familiar with the capabilities. Our research seeks to understand how exploratory search systems with rich user interfaces change the way that searchers think about and pursue their searches. What strategies and tactics do exploratory search interfaces enable? And, ultimately, do they enable searchers to achieve their higher-level objectives?

Task-based evaluation of exploratory search systems using controlled experiments has been effective for showing subjective satisfaction differences between systems, but less effective at showing objective differences in task performance, particularly in task outcomes. [4, 13]. Evaluations have assessed and rated the quality of a task outcome to generate quantitative measures on a

lesson plan creation task [4] or measured incidental learning that occurred during a search session [7]. Exploratory tasks have been decomposed or narrowed to constrain the task [3]. A combination of quantitative and qualitative evaluation methods have also been used [10, 13].

Controlled experiments and in-depth case studies are two approaches to evaluation of exploratory search systems. This position paper describes two approaches to evaluation of exploratory search systems. A mixed method approach was used to evaluate categorized overviews of search results. The longitudinal approach is proposed to extend the mixed methods.

2. MIXED METHODS

My dissertation research investigated the use of categorized overviews of web search results based on meaningful and stable categories to support user exploration and understanding of large sets of search result [6]. Web search engines are effective at generating extensive lists of results that are highly relevant to user-provided query terms. For known-item queries, users often find the site they are looking for in the first page of results. However, a list may not suffice for more sophisticated exploratory tasks, such as learning about a new topic or surveying the literature of an unfamiliar field of research, or when information needs are imprecise or evolving [12]. When searchers need to gather information from multiple perspectives or sources, categorized overviews can organize results from web or digital library searches. Categorized overviews can help searchers explore alternative sources, assess utility of results, and decide on next steps. When searchers' information needs are evolving or imprecise, categorized overviews help by stimulating relevant ideas, provoking illuminating questions, and guiding searchers to useful information they might not otherwise find.

Research prototypes and commercial search engines have incorporated categorized overviews, but there have been few user studies of categorized overviews for exploratory web search, and there is little research explaining whether they are effective, why, and under what circumstances. Research is needed to understand how categorized overviews change the way users conduct web searches, to guide the design of search engine interfaces, and to justify the entry and maintenance of category metadata.

To study this, we adopted a mixed methods approach, using an experimental design that counterbalanced two interface conditions and collecting qualitative data for analysis. The task was described in the context of a journalistic scenario, and the 24 subjects were recruited primarily from journalism students. During the two-hour session, subjects were provided training and practice, and then they conducted four searches using a think-aloud protocol, ending with a 30 minute semi-structured interview. Screen video and audio were recorded, and interactions (queries, clicks, scrolling, mouse movement, etc.) were logged

using a custom JavaScript-based tool. Based on previous research (ours and other studies), we expected to observe quantifiable and significant differences relative to several behavioral measures, including how deep in the search results subjects explored. A qualitative approach extended the hypothesis tests by looking for phenomena not modeled by the research variables. For example, we expected that the categorized overview interface would prompt tactical and cognitive changes, but there was no *a priori* list; that was developed from the data.

The study identified seven tactics that searchers began to adopt when the categorized overview was available. It highlighted two important considerations for future evaluation of exploratory search systems. First, it takes time for searchers to reflect on their searches and refine their tactics. The semi-structured interview fostered in-depth reflection by subjects. Analysis of their responses complemented the analysis of two questionnaires [5], in which users of a clustering search tool answered two questionnaires administered 6 weeks apart, reporting differences in their search tactics. Second, the strategies and tactics employed by exploratory searchers are individualized and varied.

3. LONGITUDINAL STUDIES

Longitudinal studies may be useful for addressing these two considerations. Longitudinal studies have been used, for example, to examine changes in tactics and query terms in relation to changes in searchers' information problem stage while developing a research proposal [11]. In-depth, longitudinal case studies have been used to evaluate information visualization interfaces and creativity support tools [8, 9]. These techniques integrate ethnographic and quantitative methods, using participant observation, surveys, interviews, and usage logs to study users performing complex tasks with individually defined goals. These techniques may be beneficial for investigating how searchers adapt their tactics when rich web search interfaces like interactive categorized overviews are available. They present the opportunity to observe changes as searchers become familiar with an exploratory search system and tactics mature. They also present challenges, because search is often a means to an end, and individual searches may be initiated to satisfy a higher level task. The search sessions may not be readily organized into blocks of time that can be scheduled with a researcher. We are tackling these challenges as we undertake a study using this methodology.

4. CONCLUSION

Understanding how exploratory search systems change searcher tactics and strategies is a necessary step toward designing better systems. Mixed research methods provide a way to understand how exploratory search systems change search tactics and strategies. Longitudinal studies may help understand how searchers adapt search strategies and tactics in support of creative, challenging, individualized activities.

5. ACKNOWLEDGEMENTS

This research was partially supported by an AOL Fellowship in Human-Computer Interaction and National Science Foundation Digital Government Initiative grant (EIA 0129978).

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] Golovchinsky, G. Queries? Links? Is there a difference? in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, Atlanta, GA*, ACM Press, New York, 1997, 407-414.
- [3] Janecek, P. and Pu, P. (2005). An evaluation of semantic fisheye views for opportunistic search in an annotated image collection. *Journal of Digital Libraries*, 5 (1). 42-56.
- [4] Kabel, S., Hoog, R.d., Wielinga, B.J. and Anjewierden, A. (2004). The added value of task and ontology-based markup for information retrieval. *Journal of the American Society for Information Science and Technology*, 55 (4). 348-362.
- [5] Käki, M. Findex: search result categories help users when document ranking fails. in *Proceeding of the SIGCHI Conference on Human Factors in Computing Systems, Portland, OR*, ACM Press, New York, 2005, 131-140.
- [6] Kules, B. Supporting Exploratory Web Search with Meaningful and Stable Categorized Overviews (Unpublished doctoral dissertation), University of Maryland, College Park, 2006. Retrieved June 29, 2006, from <http://hcil.cs.umd.edu/trs/2006-14/2006-14.pdf>.
- [7] Pirolli, P., Schank, P., Hearst, M. and Diehl, C. Scatter/gather browsing communicates the topic structure of a very large text collection. in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems: Common Ground, Vancouver, British Columbia, Canada*, ACM Press, New York, 1996, 213-220.
- [8] Shneiderman, B., Fischer, G., Czerwinski, M., Resnick, M., Myers, B., Candy, L., Edmonds, E., Eisenberg, M., Giaccardi, E., Hewett, T., Jennings, P., Kules, B., Nakakoji, K., Nunamaker, J., Pausch, R., Selker, T., Sylvan, E. and Terry, M. (2006). Creativity support tools: Report from a U.S. National Science Foundation sponsored workshop. *International Journal of Human-Computer Interaction*, 20 (2). 61-77.
- [9] Shneiderman, B. and Plaisant, C. Strategies for evaluating information visualization tools: Multi-dimensional in-depth long-term case studies *Beyond Time and Errors: Novel Evaluation Methods for Information Visualization (BELIV '06): A Workshop of the AVI 2006 International Working Conference*, Venezia, Italy, 2006. Retrieved June 29, 2006, from <http://hcil.cs.umd.edu/trs/2006-12/2006-12.pdf>.
- [10] Toms, E.G., Freund, L., Kopak, R. and Bartlett, J.C. (2003). The effect of task domain on search. *Proceedings of the 2003 Conference of the Centre for Advanced Studies on Collaborative Research*. 303-312.
- [11] Vakkari, P. eCognition and changes of search terms and tactics during task performance: A longitudinal case study. in *Proceedings of the RIAO 2000 Conference*, 2000.
- [12] White, R.W., Kules, B., Drucker, S.M. and schraefel, m.c. (2006). Supporting exploratory search. *Communications of the ACM*, 49 (4). 36-39.
- [13] Yee, K.-P., Swearingen, K., Li, K. and Hearst, M. Faceted metadata for image search and browsing. in *Proceedings of the SIGCHI Conference on Human factors in Computing Systems, Ft. Lauderdale, FL*, ACM Press, New York, 2003, 401-408.