

# Exploratory Search Interfaces: Categorization, Clustering and Beyond

## Report on the XSI 2005 Workshop at the Human-Computer Interaction Laboratory, University of Maryland

Ryen W. White<sup>1</sup>, Bill Kules<sup>2</sup> and Ben Bederson<sup>1,2</sup>

Human-Computer Interaction Laboratory

<sup>1</sup> Institute for Advanced Computer Studies

<sup>2</sup> Department of Computer Science

University of Maryland

College Park, MD 20742 USA

### Abstract

The development and testing of systems to support users engaged in exploratory search activities (i.e., searches where the target may be undefined) is an challenge for the online search community. In this article we report on a workshop on exploratory search issues organized in conjunction with the University of Maryland Human-Computer Interaction Laboratory's Annual Symposium and Open House in June 2005. This workshop brought together researchers from the fields of Information Seeking (IS), Information Retrieval (IR), Human-Computer Interaction (HCI) and Information Visualization (IV) for a cross-disciplinary exploration of these and other issues. Although originally intended to focus on interfaces to support exploratory search the workshop blossomed into a rich discussion of not only interface issues, but also evaluation, the cognitive processes that underlie information exploration and research methods.

### Introduction

Although the Semantic Web is making new richer and more flexible kinds of Web-based querying possible a burning question remains "Are current search technologies enough to meet all user demands?". The answer of course is yes and no. Search engines, bibliographic databases and digital libraries provide adequate support for users whose information needs are well-defined. However, they do not work well in situations where users lack the knowledge or contextual awareness to formulate queries or navigate complex information spaces [2]. For example, what if you want to find something from a domain where you have a general interest but not specific knowledge? How would you find classical music you might enjoy if you do not know what Beethoven or Berlioz sounds like? What a Sonata is? The difference between Baroque or Romantic? What do you type into Google?<sup>1</sup>

In situations where information needs are vague or evolving users generally have developed coping strategies to compensate [1,2,5]. For example, they submit a tentative query that retrieves documents and browse from there, dependent solely on their ability to

---

<sup>1</sup> Example taken from <http://www.mspace.fm>, a site describing the mSpace exploratory search browser.

interpret contextual cues and navigate between documents. In short, they *explore* available information, selectively seeking and passively obtaining cues about where next steps lie. The goal of information exploration is the refinement of a vague information need that leads, through interaction with information objects and information resources, to a more thorough understanding of the problem. Searchers may benefit from interfaces that provide additional support, for example by enabling grouping of results and/or guided discovery processes [3,4,6]. Researchers from communities such as information retrieval, user interface design, visualization and library sciences have been working for some time on techniques to support such *exploratory searches*.

Exploratory searches are typified by uncertainty about the space being searched and the nature of the problem that motivates the search. People engage in exploratory searching both intentionally and incidentally to other activities. It includes situations where: (i) the target of the search is unknown or partially unknown, (ii) the search begins with some certainty about what is known but changes into one unknown and unfamiliar on exposure to new information and (iii) users recognize useful information objects by scanning through information resources, evaluating their usefulness and determining the content/structure of a set of information objects. Relevance feedback (RF), and in particular implicit relevance feedback, is a technique that complements users exploration of the search space and the reduction in uncertainty about the target of the search that can result from this activity.

Domains such as the World Wide Web create tantalizing opportunities for learning and research, not only for students and researchers, but also for journalists, attorneys, and practitioners in many fields. The growth in interest in technologies to search these domains has meant that the problems associated with search have occupied center stage in the computing community for some time. The broad nature and scope of these problems meant that those within this community may benefit from the sharing of expertise that an inter-disciplinary workshop of this nature encourages.

## **Participation**

The workshop brought together around 40 researchers from academia, government and industry for a cross-disciplinary exploration of issues about exploratory search. Everyone contributed to a lively and engaging discussion that was well received by participants. A strength of the workshop was the diversity of participant backgrounds and the variety of perspectives on similar problems in a mixture of domains. The workshop demonstrated a high level of interest in this topic, so much so that we had to close registration early.

Participation was generally by invitation only. A number of presenters were selected prior to the workshop based on the estimated contribution of their research to the diversity of the proceedings. Those presenting were asked to provide a one paragraph abstract for their presentation and their one paragraph definition of exploratory search. The abstracts were used to arrange the speakers into topically-coherent groups. Presentations varied in length from brief one minute introductions to twenty minute keynotes. We invited four special guests to give keynote presentations:

- *Can Exploratory Search be Distinguished from Other Kinds of Information Seeking Behaviors, and Can a Search Interface Support a Variety of Such Behaviors?* (Nicholas J. Belkin, Rutgers University)
- *Are there optimal ways to do exploratory searching?* (Marcia Bates, University of Southern California)
- *Search: Lessons learned from browsing movies and pictures* (Steven Drucker, Microsoft Research)
- *Exploring Information Needs, Results Analysis, and Information Use* (Susan Dumais, Microsoft Research)

The definitions of exploratory search provided by participants prior to the workshop were used as part of an introductory “ice breaker”. As part of the introductory activity we placed everyone’s definition of exploratory search in large font on a separate sheet of paper and pinned two sets of these definitions to the wall on opposite sides of the room. We then divided the participants arbitrarily into two groups and asked them to rearrange the definitions in a way that made that made sense to them. They were given a large sheet of paper and asked to take notes on the reasons behind their decisions and any emergent themes. Subjects clustered definitions in interesting ways and provided interesting rationale for their decisions; a discussion of this is included in future work.

A workshop web site was created at <http://umiacs.umd.edu/~ryen/xsi> containing details of the program and the slides from all presentations.

The remainder of the day was divided into three oral presentation sessions broadly classified as “Seeking”, “Retrieval” and “Visualization”. The afternoon was devoted to a discussion of themes defined by participants in a pre-discussion brainstorming session. The themes that emerged were:

- *users needs* (i.e., what support user wants; what feels good, natural or correct)
- *search strategies* (i.e., how users search)
- *tasks* (i.e., what they are trying to accomplish)
- *evaluation measures* (i.e., how to evaluate support offered)
- *workspaces to support exploratory search* (i.e., what help should systems offer)

We had originally planned for breakout sessions in the afternoon covering each of these issues. However feedback from participants indicated a preference for a single discussion involving all participants devoting time to each of these issues.

### **Some implications for interface design**

The unique nature of searches where the target is unknown or undefined calls for new interface designs that help users. Participants agreed that interfaces to support user exploration are a vital (and missing) part of most current mainstream search technology. Some emergent themes from the workshop were that exploratory search systems:

- should use the context of user search activity and contextual information available the target documents to aid user understanding and reduce uncertainty about the nature of the problem and the information being searched;
- need to be able to support a diverse range of user search strategies such as recall and recognition, facet-based search and domain selection;
- are problematic to evaluate and longitudinal, ethnographic and scenario-based methodologies may be appropriate to test them, and;
- should contain “workspaces” to support a spectrum of activities, from unstructured note-taking to integrated authoring environments.

The discussion notes from the workshop are available on the workshop website. A more detailed analysis of the implications will be presented in future work.

### **Concluding remarks**

Those attending the workshop reinforced the understanding that exploratory search is a multi-disciplinary problem that is troublesome to support at the interface and difficult to evaluate interfaces that do attempt to support it. Participants felt that not only was addressing the problem timely it was also crucial to help users and a fascinating opportunity to extend current levels of understanding and apply well-established theory to new contexts. Across the variety of disciplines where techniques to support exploratory are studied there is a need to improve the contextual awareness of the user, bring information that supports exploration to the fore and establish how the effectiveness of these systems can be determined.

Further activities that will help define the field are planned. Perhaps the most exciting is that as a result of the workshop, two of the authors of this article (Ryen White and Bill Kules), alongside Steven Drucker (Microsoft Research) and m.c. schraefel (University of Southampton), are editing a special issue of the Communications of the ACM in April 2006 entitled “Supporting Exploratory Search”. The special issue will:

- bring exploratory search into the spotlight and reflect the cutting edge research being conducted into this important problem
- describe how exploratory search is handled in a variety of disciplines:
  - combine theory and practice in each of the fields of interest
  - describe how these disciplines overlap and interact
  - demonstrate the value of interdisciplinary collaboration
- propose evaluation methodologies and metrics for exploratory search interfaces

A further activity is the organization of a workshop on exploratory search at future venues such as SIGCHI, SIGIR and ASIST. At these forums it is hoped that we can attract large numbers of participants to continue the discussion of the issues raised in this workshop.

The scope of search technologies must expand beyond known search targets to support information exploration. It gave all of the organizers great pleasure to arrange this workshop and focus the minds of those interested in online search on this problem.

## Acknowledgments

We are grateful to AOL for their financial support, to Doug Oard and Ben Shneiderman for their invaluable comments and guidance, and Kiki Schneider for her administrative assistance.

## References

- [1] Baldonado, M. Q. W. and Winograd, T. (1997). SenseMaker: An information-exploration interface supporting the contextual evolution of a user's interests, In *Proceedings of the Conference on User Interface and Software Technology*, pp. 11-18.
- [2] Bates, M. J. (1989). The design of browsing and berrypicking techniques for the online search interface. *Online Review* 13 (5), pp. 407-424.
- [3] Bederson, B. B. (2001). PhotoMesa: A zoomable image browser using quantum treemaps and bubblemaps. In *Proceedings of the Conference on User Interface and Software Technology*, pp. 71-80.
- [4] Kules, B., Shneiderman, B. and Plaisant, C. (2003). Data exploration with paired hierarchical visualizations: Initial designs of PairTrees. In *Proceedings of the National Conference on Digital Government Research*.
- [5] Pirolli, P. and Card, S. (1995). Information foraging in information access environments. In *Proceedings of the Conference on Human Factors in Computing Systems*, pp. 51-58.
- [6] White, R. W., Jose, J. M. and Ruthven, I. (2005). Using top-ranking sentences to facilitate effective information access. *Journal of the American Society for Information Science and Technology*, 56 (10), pp. 1113-1125.