Innovation and Evaluation of Information

A CHI98 Workshop

Gene Golovchinsky and Nicholas J. Belkin

Abstract

This report summarizes a workshop held at CHI 98 that focused on several aspects of information exploration, including user interfaces, theory, and evaluation. Information exploration is a common activity that spans a variety of media and is an integral component of many information seeking behaviors that people engage in. The complexity of this activity, and the need to support it appropriately, led us to propose this workshop. Over the course of two days, we examined several aspects of this problem, struggled with a few definitions, and came away with a better understanding of the design space. Here we summarize those efforts.

Introduction

Traditional Information Retrieval is concerned with improving effectiveness of indexing and retrieval mechanisms, and with supporting one information seeking behavior: specified searching through query formulation. This has been predicated on support for one kind of user population, with one kind of information need. But the networked information environment has resulted in a shift in the user population of information retrieval systems. This change has introduced new classes of users, in the sense of levels of expertise, and has also made clear that there are different kinds of information needs and different kinds of information seeking behaviors than those supported by traditional IR systems and techniques. This workshop focused on developing understanding of one such information seeking behavior, Information Exploration, on interface design for supporting this behavior, and on evaluation methods and measures for assessing such interfaces.

Information Exploration addresses the goal of refining a vague concept into a more thorough understanding of the problem which led to the information interaction. We believe that information exploration research falls squarely in the domain of human-computer interaction with some emphasis on information retrieval, rather than vice versa. Thus one of the thrusts of this workshop was to attempt to characterize the activities users engage in, to design for those activities, and to identify evaluation techniques and measures that provide appropriate insights into users' behavior and performance.

Organization

About 20 people participated in the workshop. They were chosen on the basis of initial brief submitted position papers, and represented a broad spectrum of industry and academia. Participants came from France, Canada, Germany, and the U.S. After acceptance, participants were asked to submit longer (4-5 page) position statements that described relevant research and perspectives a few weeks prior to the workshop. These papers were made available through the workshop web site, and participants were encouraged to review and comment on them.

Submissions were organized into three categories: Interface, Evaluation and Theory. Each category was further subdivided into themes that suggested themselves. Thus a number of interface submissions concerned information visualization; three of five evaluation-related submissions focused on expertise, and the theory section split evenly between frameworks and representation of information.

On the morning of the first day, workshop activities were organized based on the three topics we had initially defined. After the morning introductory session, we split the workshop into three new working groups, based on the results of that discussion.

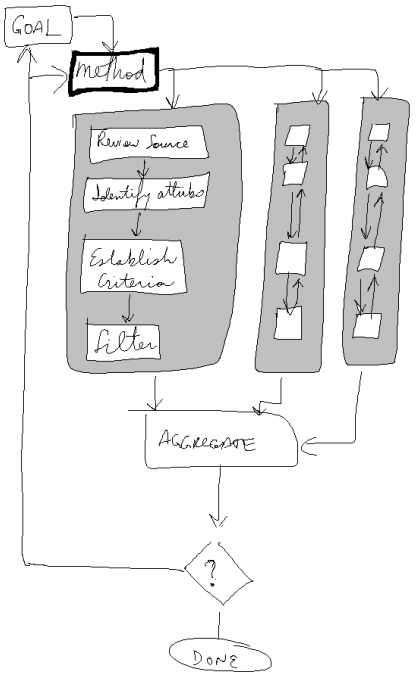

Figure 1. Information exploration (gray box) situated in the broader task. The black "method" box may involve a recursive information exploration step to identify information sources.

Discussion Highlights

It seems obligatory for a workshop to debate the definition of the concept that brought people together; we embraced this orthodoxy with a vengeance. One of the recurring themes of

the workshop was whether information exploration is a process that a person engages in, or rather a lower-level sub-task (akin to searching, reading, writing, etc.) that is part of a larger information seeking activity.

At the end, we decided that it was not useful to characterize information exploration as the larger activity because this definition did not help us to elaborate on information seeking in general. Instead, we adopted the sub-task alternative. There was agreement, however, that uncertainty regarding information sources, incomplete or inadequate mental models of the sources and/or the domain, and learning were all important aspects that characterized an activity as information exploration, and distinguished it from other kinds of interactions with information. A number of such aspects of information exploration identified. These included:

- Uncertainty of information
- Incomplete information
- Incomplete or inappropriate mental model
- Learning
- Use of metadata

From this we can synthesize an approximate definition:

Information exploration is the process of examining metadata regarding information sources in the presence of an incomplete or inadequate mental model for the purpose of addressing some information need.

We also described the process graphically, as shown in Figure 1 and illustrated with a relevant example in Table 1. The information seeking process starts with a goal, often vague or ill-defined, regarding some information need (e.g., "What are some good restaurants in Los Angeles where we could have dinner this evening?"). Next, the method of finding information is identified. The information exploration step consists of reviewing the source, identifying its attributes, establishing criteria on which to select information, and filtering the results. Several such passes may occur simultaneously, and the results are then aggregated to produce a set of candidate answers. At this point, one of three possibilities exists: the information need is satisfied, and the process ends; another source is selected, and the process repeats; or the goal is modified and the process repeats.

It is interesting to note that the method step itself may be recursive: the task of identifying sources of information may rely on an information exploration step. For example, we could have tried to find out if the local night-life paper (LA Weekly, for example), contained listings of the kinds of restaurants in which we were interested. Selecting Zagat's, on the other hand, would not constitute an exploration step because we were already familiar with its purpose and reputation.

Step	Example
Goal	Find a good restaurant in Los Angeles where we can have dinner this evening
Method	look in conference materials and Zagat's
Review source	See how lists are organized
Identify attributes	distance, quality, cuisine, price, atmosphere
Establish criteria	good food, quiet, nearby
Filter	identify several candidate restaurants
Aggregate	compare Zagat's and conference lists results
Go back to method	
Method	ask concierge for recommendations
(review - filter)	
Goal	Go to bar instead

Table 1: Information exploration process

Groups

We split the workshop into three groups to allow more focused discussion of topics that recurred during the morning presentations. One group focused on the use of multiple views to support exploration, one focused on tools to support information exploration tasks, and the third handled evaluation.

Multiple Views

The multiple views group addressed issues of visualization in support of information exploration. During the first day, they identified several issues that should be addressed during the design of visualizations, and in the second day they created design guidelines based on these dimensions. They identified two goals of visualization:

1. Show data at different levels of granularity
2. Show data at different structural/ syntactic/semantic/lexical levels

The issues included

- When and how to use multiple visualizations / views:
  Information-theoretic perspective (support for tasks)
  Provide a holistic view
  Accommodate different users' abilities (different devices, media, contexts)
  Facilitate chance encounters
  Concern: Cognitive cost of switching
- When and how to couple multiple views:
  CARE taxonomy – Complementarity, Assignment, Redundancy, Equivalence
  Time dimension
  User control
  Concern: Loss of information
- How to provide smooth transitions among views:
  Animation
  Coordinated navigation
  Perceptual highlighting
  Linked views at different levels of granularity
  User-driven manipulation of transitions
  Perceptual cues to accentuate changes at periphery (e.g., movement)

Tasks and Tools

The tasks and tools group took a two-pronged approach: they enumerated some typical activities that users engage in that motivate information

exploration activity, and then described some low-level activities that comprise the information exploration step. High level tasks included conducting literature reviews, doing information analysis (knowledge building), collecting nuggets and growing pearls (aspectual searching), looking for anomalies in data, and support for making predictions.

One approach to decomposing information exploration activity was to identify several atoms that would combine to produce various exploratory behaviors. Categorization, relevance judgment, knowledge building and classification were proposed as candidates. This group also characterized some tools with respect to the information exploration tasks they support, as shown in Table 2.

Activity	Tool
Discovering structure, cataloging, metadata	Forms, examples
Understanding vocabulary	Thesauri, 2-level hypertext
Understanding organization of collection	Hierarchy browsers, clustering
Understanding content of collection	Clustering
Finding patterns	Visualizations, clustering
Finding outliers	Datamining

Table 2: Tasks and tools

Evaluation

Traditionally, evaluation of information exploration activities has focused on system effectiveness or efficiency. Increasingly, however, users' behavior has become recognized as an object worthy of study. Dimensions along which users and their activities can be classified and measured are shown in Table 3.

Measures can also be classified as task dependent and task-independent. The popularity of traditional measures such as recall and precision in information retrieval work has been based chiefly on the relative ease with which data could be collected and compared between systems and experiments. Unfortunately, these comparisons are not always meaningful as they ignore differences in users' goals, in experimental tasks, and in user interfaces. Thus a range of task-dependent measures that are based on user goals, task goals and system goals can be more appropriate for evaluating interactive systems. In many cases, it is difficult or impossible to achieve between-system control even with traditional measures, which strengthens the argument in favor of measures derived from the specific circumstances (system, user, task) being evaluated.

Category	Measures	Measurement techniques
Performance	Effectiveness, "goodness", workload	System logs, performance on secondary tasks
Demographics	Age, sex, experience, education, etc.	Self-reporting
Psychographics	Spatial ability, Myers-Briggs, etc.	Questionnaires, test questions
Mental models	Learning	Self-reporting, comprehension questions (e.g., "where would you find...?", "Draw a map of…", etc.)
Efficiency	Time, number of steps	System logs
Strategy	Patterns of actions	System logs, self-reporting, verbal protocols

Table 3: Dimensions and measures of users' performance and behavior

Information exploration is a complex activity. Experiments involving information exploration tasks are typically conducted to test user interface or retrieval system designs. To characterize information exploration activity with adequate fidelity, attention must be paid not only to system and user interface design, but also to the users and their tasks and goals. Thus the data collection space may also be characterized by the three dimensions of users, tasks/goals, and environment, in which each dimension varies from "real" to "synthetic", as shown in Figure 2. This space can be used to classify different data collection situations, ranging from laboratory experiments to field studies. Point A in the figure represents the situation when a group of subjects perform an assigned task in a laboratory setting. The people are not chosen specifically for any expertise related to the task, they perform activities that are not part of their real work, and they carry out these activities in a laboratory or other artificial setting using an interface created solely for the experiment. The other extreme is represented by point B: here, the users, their tasks, and the system they use are "real" in the sense that they were not brought together for the experimenter's convenience. The experiment may be a field study or a controlled experiment; what is important here is that the complexity of actual work is captured, and thus the results retain a high degree of ecological validity. It is possible, of course, that generality of findings may be compromised. Finally, point C represents a situation where users bring their actual information needs to an experimental system.

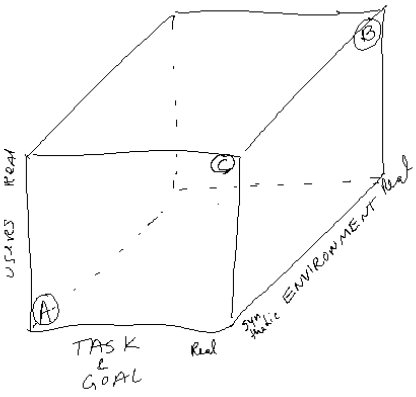

Figure 2. Design space for experiments that represents users, tasks and the environment. Point A represents a highly-artificial situation in which randomly-

selected people perform an assigned task. Point B represents, for example, a field study of real use of a real system by real people. Point C. Represents a deployment of a test system (i.e., a synthetic environment) to which real people bring their real information needs.

The evaluation of information exploration systems involves not only the design of an interface or system, but also requires the experimenter to pay careful attention to the goals and motivations subjects bring to the experiment, and the collections they search as the assigned task. Only careful handling of all three dimensions – user, task and system – will insure that the data collected in the experiment will provide truly useful information that can inform subsequent design.

Conclusions

This area of research brings together many research disciplines, including user interface design, human-computer interaction, information retrieval, visualization, cognitive and experimental psychology, ethnography, computer sciences, and others. The problems posed by the activity of making sense of information require solutions that draw on these multiple skills. The participants of this workshop brought many different perspectives on a variety of issues related to information exploration to the discussion.

The first step to solving a problem is recognizing that you have one. This workshop provided an excellent opportunity to define some of the problems characteristic of information exploration tasks. We are trying to solve systems problems: not merely computer systems, but systems that involve human actors engaging in difficult cognitive activities. Thus it is important for practitioners in all aspects of this field to understand the complexity of the issues. We believe this workshop was a useful step toward a broader and more comprehensive understanding of the behaviors that comprise information exploration and of their implications for the design of systems.

Acknowledgments

We would like to thank the participants of this workshop: they made it the success that it was. We would also like to thank ACM and SIGCHI, and the reviewers of our proposal who made many helpful suggestions.

Participants

Nicholas J. Belkin, Rutgers University, USA (Organizer)
Peter Brusilovsky, Carnegie Mellon University, USA
Mark H. Chignell, University of Toronto, Canada
Gene Golovchinsky, FX Palo Alto Laboratory, USA (Organizer)
Beth Hetzler, Pacific Northwest National Laboratory, USA
Jürgen Koenemann, German National Research Center for Information Technology (GMD FIT), Germany
Rick Kopak, University of Toronto, Canada
Allan Kuchinsky, Hewlett Packard Laboratories, USA
Laurance Nigay, Laboratoire CLIPS_IMAG, France
Allison L. Powell, University of Virginia, USA
Morgan N. Price, FX Palo Alto Laboratory, USA
Frank M. Shipman III, Texas A&M University, USA
Diane H. Sonnenwald, University of North Carolina at Chapel Hill, USA
Russell Swan, University of Massachusetts, USA
Elaine G. Toms, Dalhousie University, Canada
Frederic Vernier, Laboratoire CLIPS_IMAG, France
Michelle Q Wang Baldonado, Xerox PARC, USA
Karen Wilson, Lexis-Nexis, USA
Allison Woodruff, UC Berkeley, USA

Authors' Addresses

Gene Golovchinsky
FX Palo Alto Laboratory, Inc.
3400 Hillview Ave., Bldg. 4
Palo Alto, CA 94304
USA
e-mail: gene@pal.xerox.com
Tel: +1 650 813 7361

Nicholas J. Belkin
Rutgers University
School of Communication, Information and Library Studies
SCILS 311, 4 Huntington Street
New Brunswick, NJ 08901-1071
USA
e-mail: nick@belkin.rutgers.edu
Tel: +1 732 932 8585