Visual Interfaces for Stimulating Exploratory Search

Ralf Krestel
L3S Research Center
Hannover, Germany
krestel@L3S.de

Gianluca Demartini
L3S Research Center
Hannover, Germany
demartini@L3S.de

Eelco Herder
L3S Research Center
Hannover, Germany
herder@L3S.de

ABSTRACT

Exploration is an activity that people undertake to broaden their knowledge on a certain topic. In contrast to regular search, which is typically aimed at obtaining a specific answer to a specific question, exploratory search should give a more complete overview of a topic. Further it should enable the discovery of related aspects, such as people, places, times and locations. Exploration demands more time, effort and creativity from the user, but rewards the user with deeper knowledge. Therefore, users need to be stimulated to bring exploration in regular goal-directed search activities. In this paper we present a user study in which we investigate different kinds of exploratory behavior and goals, as well as different kinds of visualizations to support exploration.

Categories and Subject Descriptors

H.5.2 [**Information Interfaces and Presentation**]: User Interfaces—*Evaluation/methodology*; H.5.4 [**Information Interfaces and Presentation**]: Hypertext/Hypermedia—*Navigation*; H.3.3 [**Information Storage and Retrieval**]: Information Search and Retrieval—*Clustering*

General Terms

Experimentation, Human Factors, Measurement

Keywords

Search Interface, Visualization, Exploratory Search

1. INTRODUCTION

Searching information is among the top activities when it comes to time spent online. But search is not only important on the Web: digital libraries provide users with a colossal reservoir of information. Specialized archives store years of years of accumulated data. News archives contain manifold data about persons, events, and background knowledge. Without a suitable user interface and powerful search algorithms, this information is not accessible for non-professional users.

Because different users have different goals when entering a keyword in a search engine field, searching comes in a variety of flavors. Besides the goal of a search, users also differ in the way they try to achieve this goal. As a first taxonomy, Broder [1] classified the search activities of Web users into three query types: navigational, transactional, and informational to account for different objectives. Commercial search engines try to help users as much as possible to quickly find what they are looking for. They do a good job for navigational and transactional queries, but for informational queries – where users are often interested in finding answers to a specific questions or to acquire knowledge about broader topics - standard search engines do not achieve optimal results. In contrast to Web search tasks, exploratory search can be characterized as learning-oriented or investigative [3]. Especially for exploratory search, where users have a more fuzzy information need and want to explore a topic, concept, or event the result presentation is a decisive element together with the ranking algorithm to offer the best support for the information seeker.

Discovery of unknown facts and identifying relations within a topic of interest demands for result presentations beyond a pure list of documents. Current search engines do not sufficiently support exploration and discovery, as they do not provide an overview of a topic or assist the user by finding related information. A step in the direction of exploratory search is query suggestion where the search engine recommends related queries. There are also approaches that cluster search results[1] which can help users dive into a topic.

2. VIZIO: A SYSTEM FOR VISUALLY EXPLORING NEWS COLLECTIONS

In order to study how users perform exploratory search we developed a search system on top of the New York Times (NYT) collection [4]. This search engine is taking as input keyword queries from the user allowing her to select one of the five possible result visualization type: world map, timeline, related words, related persons and places, or the classic list of results.

The developed system, named Vizio, combines state-of-the art tools to provide a visual search interface. A user can submit a query and the system runs it against an inverted index of news articles. More than just the full-text we consider also additional metadata of the news articles. We include information about the countries related to the news article, the publishing date, and we extract entities (i.e., persons, locations, and organizations) from the text. The screenshot of Vizio in Figure 1 shows a generated timeline visualization[2].

[1] http://search.yippy.com/

[2] For implementation we used Lucene, LingPipe, and Protoviz libraries.

JCDL'11, June 13–17, 2011, Ottawa, Ontario, Canada.
ACM 978-1-4503-0744-4/11/06.

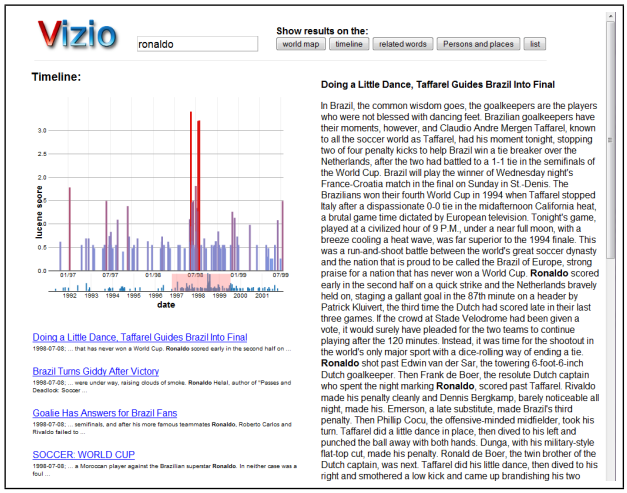

Figure 1: A Screenshot of Vizio: Our System for Exploratory Search on Top of a News Corpus

3. PRELIMINARY EVALUATION

With the Vizio System we conducted a task-oriented user study on exploratory search on a range of topics that are known to be of public interest and that are covered by news articles in the NYT corpus. We recruited ten participants who were invited to explore these topics in individual user sessions, followed by a structured interview. The sessions were recorded making use of screen and audio capturing equipment.

3.1 Dataset and Tasks

For the experiments we took a corpus of news articles from the New York Times (NYT) [4]. From this collection we picked a period of ten years from 1991 to 2001. This dataset consists of all articles published by the NYT within these ten years.

In order to ensure sufficient variety in the tasks we chose 3 topics for the user to explore from three different categories, "politics", "culture", and "sports". The topics were: "Terror Attacks", "Nobel Prize", and "Olympic Games". The basic task associated with each topic was to conduct leisure browsing in the set of news articles between 1991 and 2001 related to these topics. For each topic we prepared a short introduction in which we mentioned some major events that the participants were likely to remember. For example, when introducing the topic 'terror attacks' we mentioned the Twin Towers, Osama Bin Laden and the IRA; for the topic 'Nobel Prize' we mentioned peace, medicine and physics as major fields. After the short introduction, the participants could freely search and browse using the Vizio system. We anticipated that some participants would at some point lose interest and therefore prepared suggestions for subtopics, to be given sparingly.

3.2 Results and Discussion

The timeline and the world map were the most popular visualizations - they were used more often than the list, related words or persons and places. At the same time, the average usage duration of both visualizations was higher as well, see Figure 2. Further, user interaction with the timeline and world map was higher, as signified by the use of the zoom

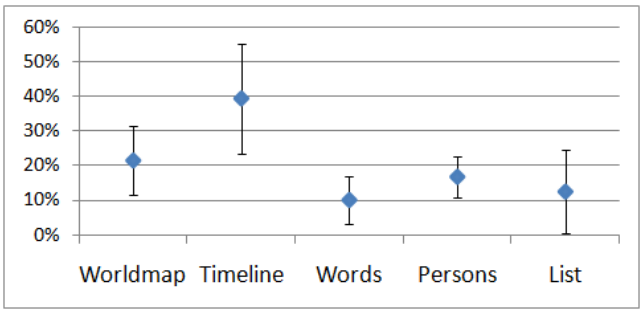

Figure 2: Average Relative Time Spent by Users on Each Visualization Together with Standard Deviation

function. These statistics confirm that events and places are considered important dimensions of news items [2].

A more qualitative look at the video recordings revealed that the different visualizations lead to fundamentally different kinds of exploratory behavior. The timeline and world map provide "topic-neutral" cues in the form of event peaks and highlighted countries. These peaks and highlights typically resulted in the user trying to find what caused it. For example, virtually all participants clicked on the highlighted country Myanmar (Burma) on the worldmap for "nobel prize" and then found the famous 1991 laureate of the Nobel Peace Prize. Similarly, the 1993 peak in the timeline for "terror attacks" led most participants to learn about the 1993 bombing of the Twin Towers, an event that most participants could only vaguely remember, as it was overshadowed by the 9/11 event.

4. CONCLUSIONS

Exploratory search is a means for acquiring knowledge in digital information spaces. Standard search interfaces do not support exploration of data sufficiently. We presented a system incorporating different result visualizations to stimulate exploratory search. A user study was conducted to analyze user behavior and exploratory search strategies. The results reveal the need for advanced visualizations to assist users doing exploration. To this end we encourage search interface designers to include more visual tools to provide an overview of the search results and to allow for easier narrowing down towards specific aspects. In particular we found a correlation of different search strategies and used visualization as well as a connection between the topic and the chosen visual aid.

5. REFERENCES

[1] A. Broder. A Taxonomy of Web Search. *SIGIR Forum*, 36 (2), 2002.

[2] M. Dörk, S. Carpendale, C. Collins, and C. Williamson. VisGets: Coordinated Visualizations for Web-based Information Exploration and Discovery. *IEEE Trans. Visualization and Computer Graphics*, 14 (6):1205–1212, 2008.

[3] B. Kules, R. Capra, M. Banta, and T. Sierra. What Do Exploratory Searchers Look at in a Faceted Search Interface? In *Proc. JCDL 2009*, 2009.

[4] E. Sandhaus. The new york times annotated corpus. *Linguistic Data Consortium, Philadelphia*, 2008.