

Lightweight Additions to the Web Search Interface Supporting Exploratory Web Search

Orland Hoerber
Department of Computer Science
Memorial University of Newfoundland
St. John's, NL A1C 3X5
Canada
hoerber@cs.mun.ca

ABSTRACT

In this paper, the features of *TheHotMap.com* that support exploratory Web search processes are described. This system grew out of two academic research projects that explored the use of visualization and interaction as a means for supporting users as they conduct Web search tasks. In *TheHotMap.com*, three lightweight interface extensions have been added to the commonly used list-based representation of Web search results. These can be used independently or together to support users as they craft queries and explore search results. A scenario of using the system for exploratory Web search is described in this paper; a live demonstration will be provided at the workshop.

Categories and Subject Descriptors

H.3.3 [Information Storage and Retrieval]: Information Search and Retrieval; H.5.2 [Information Interfaces and Presentation]: User Interfaces

General Terms

interfaces, search, exploration

Keywords

Web search, information visualization, interaction

1. INTRODUCTION

In recent years, Web search has become an important part of the information-seeking and knowledge-generating activities of the general public. A study from 2004 reported that 88% of Web users start their information-seeking tasks with a search engine [13]. Although more recent studies on Web user behaviour do not directly address the frequency of use of search engines, market research has shown monthly increases in their use in the United States [3].

Although searching has become the primary tool for finding information on the Web, the interfaces employed by the

top search engines have changed very little since the early days of Web search. The primary interface features continue to be a query box for capturing the searcher's intent, and a list-based representation of the search results. Although such interfaces are very easy to learn and use, their power for supporting complex or exploratory search tasks is limited.

Our primary motivation for this research has been to explore the use of information visualization and interaction techniques to support Web search activities. Information visualization is a technique for creating interactive graphical representations of abstract data or concepts [15]. Moreover, information visualization promotes a cognitive activity in which users are able to gain understanding or insight into the data being graphically displayed by taking advantage of human visual information processing capabilities [14].

The potential benefits of employing information visualization and interaction techniques to support Web search activities are immense. However, the challenge is to show restraint in the design of such systems, and avoid overly complex visual representations and interaction methods that are difficult to learn and use. Our focus in this paper is on three lightweight extensions to the commonly used list-based representation that support exploratory Web search activities.

2. RELATED WORK

This work is closely related to our previous research activities in the development of visual and interactive interfaces for Web search. In particular, the system is based on a combination of two of our previous research prototypes: HotMap [8] and Wordbars [7]. These prototypes were originally developed with the purpose of exploring visual representations, interaction, and use of various types of information to support Web search activities. Combined together, they allow the searcher to easily switch between their two primary tasks of interactive query refinement and interactive search results exploration [9].

As research tools, these prototypes were useful for validating the potential utility of the proposed techniques [10, 11]. However, they were not designed for public release. *TheHotMap.com* is a complete re-implementation and extension of the methods employed by these previous works.

Others have explored the use of visual interfaces to support the evaluation of Web search results. Heimonen and Jhaveri [6] created an icon-based representation of the locations of specific query terms within individual search results sets. Based on TileBars [5], this system allowed the searcher to see where in the resulting documents their search terms

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, to republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.

HCI 2008 Redmond, WA, USA

Copyright 200X ACM X-XXXXX-XX-X/XX/XX ...\$5.00.

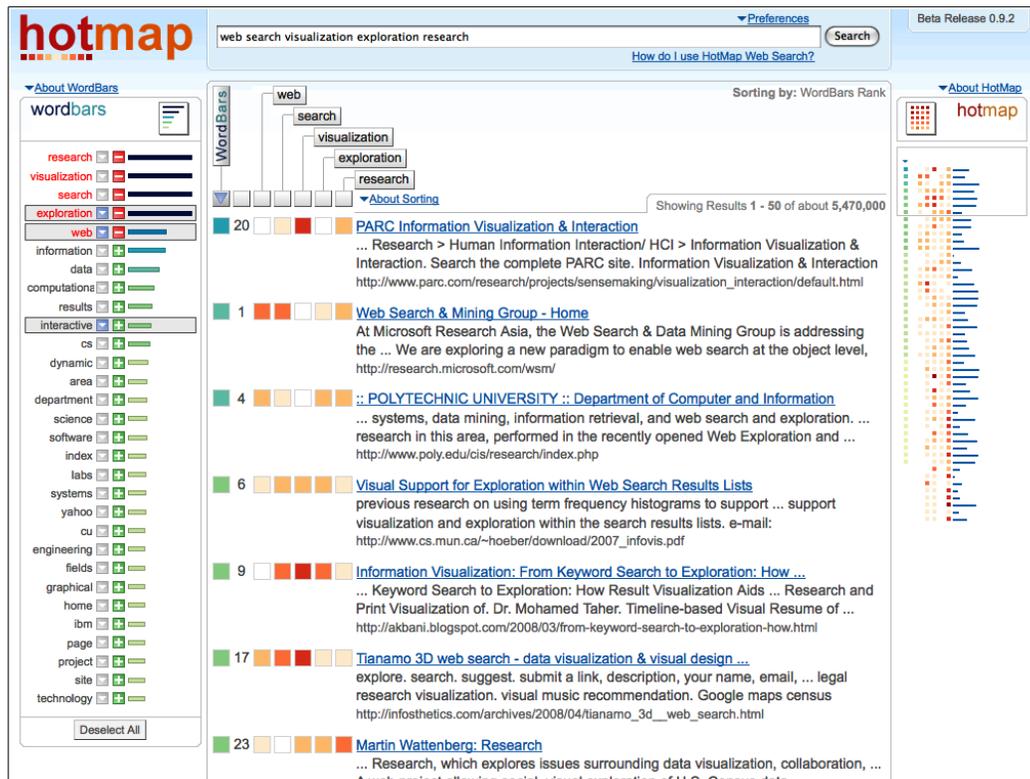


Figure 1: A screenshot of *TheHotMap.com*. Note the lightweight controls representing common terms in the search results set (WordBars histogram, on the left) and the frequency of the query terms in the search results set (HotMap overview, on the right), along with the re-sortable search results list using the HotMap query term headers (centre).

were being used together.

In VIEWER [1], the frequency of all combinations of the query terms were counted within the document surrogates. This information was presented in a histogram representation. Selections within the histogram allowed the searcher to filter the search results set based on specific combinations of the query terms.

Web search clustering systems, such as Clusty [2] and Grokker [4], dynamically identify and label clusters of documents discovered within the search results sets. Normally presented in a tree-based structure, users can expand and select clusters, resulting in a filtering of the search results set. Kules [12] extended the standard paradigm for clustering search engines by providing a consistent naming scheme for the clusters. The result is a system that allows users to learn the names and meanings of the clusters over time.

3. SYSTEM FEATURES

TheHotMap.com is implemented as a Web search interface layer overtop of the search results provided by the Yahoo API [16]. There are three main features that support exploratory Web search activities using the system: the WordBars histogram, the HotMap overview of the full search results set, and the re-sortable search results list using the HotMap query term headers. Each of these features are described in more detail below; specific details on the techniques and their potential benefits are provided in [10, 11].

Figure 1 provides a screenshot of *TheHotMap.com*. Al-

though the number of daily queries is currently limited, the system is available as a publicly accessible demonstration at <http://www.thehotmap.com/>.

3.1 WordBars Histogram

The WordBars histogram provides a visual representation of the most frequently appearing terms within the search results set, allowing the relative frequency of these terms to be easily observed. Users can interactively re-sort the search results set by selecting the arrow icon beside any term of interest. A visual indicator within the search results list (under the vertical WordBars button) depicts the frequency of the selected terms within each search result. Searchers can easily select and un-select terms of interest as they explore the search results. Interactive query refinement is supported by clicking the plus icon beside any term users wish to add to their queries, or the minus icon beside any term users wish to remove from the query.

3.2 HotMap Overview

The HotMap overview provides a compact visual representation of the entire set of search results that are present in the list-based representation. In the current implementation, the system collects 50 search results per page. Colour coding is used to represent the frequency of the query terms within the search results set; bars that are relative to the length of each search result title are included to support the visual mapping between the search results set and the

HotMap overview. The colour coding of term frequencies is also used in the search results list, resulting in the HotMap overview appearing as a “zoomed out” view of the search results set.

The HotMap overview supports the visual exploration of the search results. As users identify documents of interest, they may click on the abstract representation of the search result in the HotMap overview to cause the search results list to scroll to that location. The system temporarily highlights the corresponding search result that was selected in the HotMap overview, allowing users to easily relate their selection in the overview to the scrolled location in the search results list.

3.3 HotMap Re-Sorting

In addition to the re-sorting supported via the WordBars histogram, searchers may also re-sort the search results based on the frequency of use of their specific query terms within the search results. Clicking on any of the query term headers above the search results list will cause the search results to be re-sorted. Although the default sorting method is to perform single-term sorting, an advanced feature is available that supports nested sorting.

4. EXPLORATORY SEARCH SCENARIO

A scenario illustrating the use of *theHotMap.com* when conducting an exploratory Web search based on incomplete knowledge about the task is provided in Figure 2. This scenario shows how a user can start with an initial query (a) and use the features of the WordBars histogram to explore the search results and learn about the topic (b and c). The WordBars histogram also supports the user in making modifications to the query based on what they have learned (d). The HotMap overview allows the searcher to visually inspect areas of interest in the search results set and easily jump to the corresponding location in the search results list (e). The system also supports re-sorting the search results based on the importance the searcher places on their query terms (f).

Although this scenario shows the searcher first using the WordBars histogram features, followed by the HotMap overview and re-sorting features, this order of use is not enforced by the system. Searchers are free to use whichever feature of the system that best supports their current search objective. For example, if the searcher wishes to start with a somewhat vague initial query, and then explore and evaluate the search results seeking relevant terms to add to their query, they may do so easily using the WordBars histogram features. Alternately, if the searcher is already confident in the quality of their query and they wish to explore the search results seeking relevant documents, they may do so visually using the HotMap overview, or via interactive re-sorting of the search results based on their query terms.

5. CONCLUSIONS

TheHotMap.com adds three lightweight additions to the commonly used list-based representation of Web search results. Used together or separately, the features supported by these additions provide flexible methods for conducting exploratory Web search activities, allowing users to interactively refine their queries and interactively explore the search results. Visualization techniques are used to depict information that is relevant to the searchers’ primary tasks and

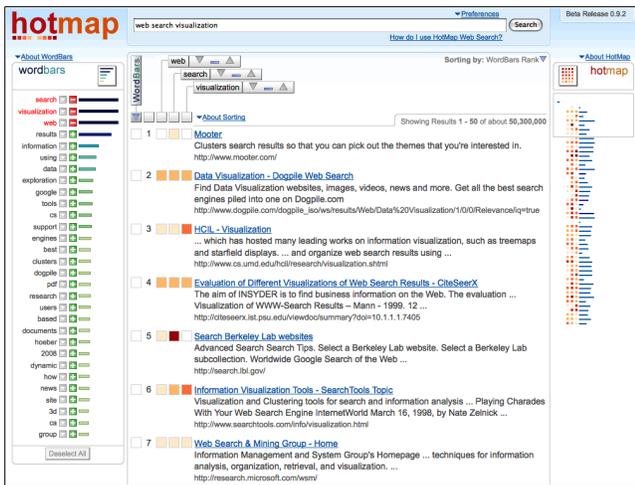
goals. The interactivity of the system allows searchers to take an active role in their Web search activities.

6. ACKNOWLEDGMENTS

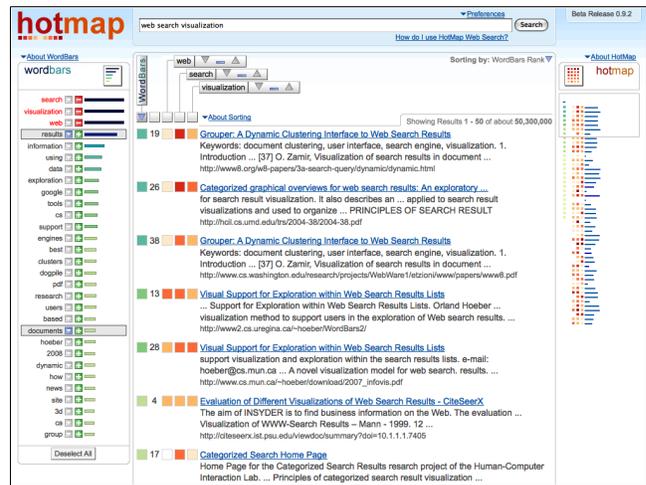
The authors thank the University-Industry Liaison Office at the University of Regina for their support of this work.

7. REFERENCES

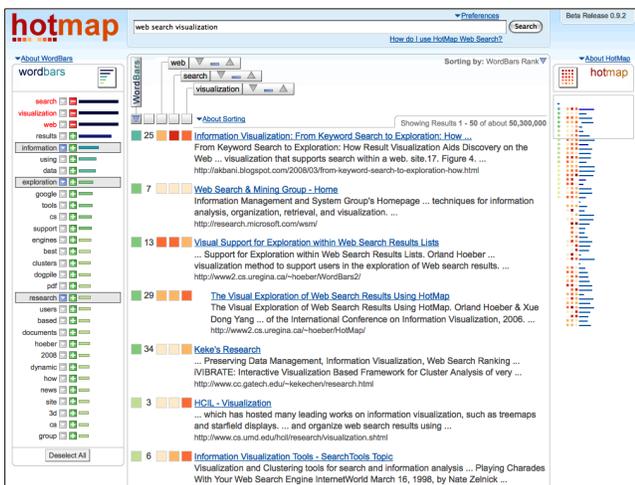
- [1] E. Berenci, C. Carpineto, V. Giannini, and S. Mizzaro. Effectiveness of keyword-based display and selection of retrieval results for interactive searches. *International Journal on Digital Libraries*, 3(3):249–260, 2000.
- [2] Clusty. Clusty: the clustering search engine. <http://www.clusty.com/>, 2008.
- [3] comScore. comScore releases June 2008 U.S. search engine rankings. <http://www.comscore.com/press/release.asp?press=2337>, June 2008.
- [4] Grokker. Grokker - enterprise search management and content integration. <http://www.grokker.com/>, 2008.
- [5] M. Hearst. Tilebars: Visualization of term distribution information in full text information access. In *Proceedings of the ACM Conference on Human Factors in Computing Systems*, 1995.
- [6] T. Heimonen and N. Jhaveri. Visualizing query occurrence in search result lists. In *Proceedings of the International Conference on Information Visualization*, 2005.
- [7] O. Hoerber and X. D. Yang. Interactive Web information retrieval using WordBars. In *Proceedings of the IEEE/WIC/ACM International Conference on Web Intelligence*, 2006.
- [8] O. Hoerber and X. D. Yang. The visual exploration of Web search results using HotMap. In *Proceedings of the International Conference on Information Visualization*, 2006.
- [9] O. Hoerber and X. D. Yang. A unified interface for visual and interactive Web search. In *Proceedings of the IASTED International Conference on Communications, Internet, and Information Technology*, 2007.
- [10] O. Hoerber and X. D. Yang. Evaluating WordBars in exploratory Web search scenarios. *Information Processing and Management*, 44(2):485–510, 2008.
- [11] O. Hoerber and X. D. Yang. HotMap: Supporting visual exploration of Web search results. *Journal of the American Society for Information Science and Technology*, in press.
- [12] B. Kules, J. Kustanowitz, and B. Shneiderman. Categorizing Web search results into meaningful and stable categories using fast-feature techniques. In *Proceedings of the ACM/IEEE-CS Joint Conference on Digital Libraries*, 2006.
- [13] J. Nielsen. When search engines become answer engines. <http://www.useit.com/alertbox/20040816.html>, August 2004.
- [14] R. Spence. *Information Visualization: Design for Interaction*. Pearson Education, 2nd edition, 2007.
- [15] C. Ware. *Information Visualization: Perception for Design*. Morgan Kaufmann, 2004.
- [16] Yahoo. Yahoo search Web services. <http://developer.yahoo.com/search/>, 2008.



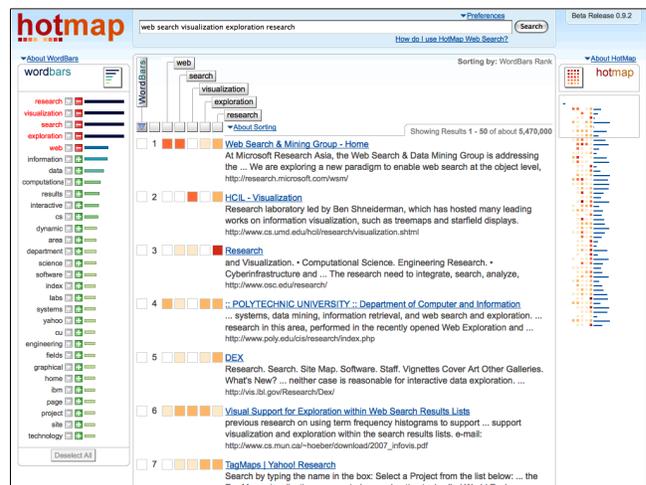
(a) Initial search for “web search visualization”.



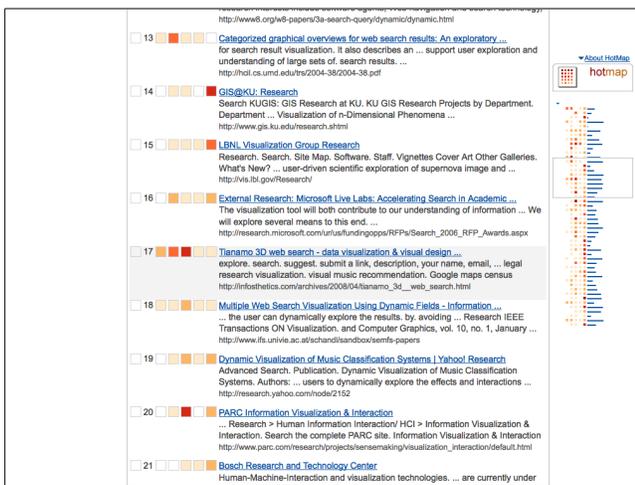
(b) Exploration of the search results by selecting “results” and “documents” from the WordBars histogram. Note the highlighted terms and re-sorted search results.



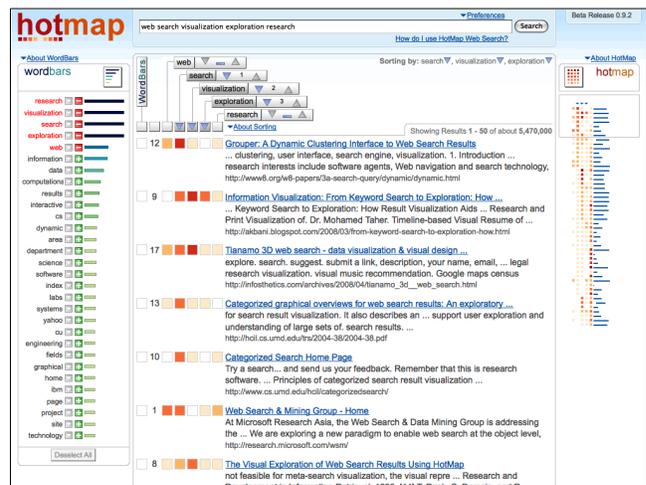
(c) Further exploration of the search results by selecting “information”, “exploration”, and “research” from the WordBars histogram. Note the different order of the search results from the previous step.



(d) The query is refined by adding “exploration” and “research” using the WordBars histogram. Note the new set of search results, the new WordBars histogram, and the new HotMap overview.



(e) Visual inspection of the HotMap overview reveals a potentially interesting document deep in the search results list. Clicking on it scrolls the search results to the appropriate location, and temporarily highlights the document (doc. 17).



(f) Re-sorting the search results based on query terms of specific interest to the searcher allows them to provide supplemental information about the importance of their query terms to the search process.

Figure 2: Screenshots from *TheHotMap.com* illustrating the features that support exploratory Web search.