TECHNICAL RESEARCH REPORT

Categorized Graphical Overviews for Web Search Results: An Exploratory Study using U.S. Government Agencies as a Meaningful and Stable Structure (2004)

by Bill Kules, Ben Shneiderman

TR 2005-71

Categorized graphical overviews for web search results: An exploratory study using U. S. government agencies as a meaningful and stable structure

Bill Kules and Ben Shneiderman
Department of Computer Science, Human-Computer Interaction Laboratory, and Institute for Advanced Computer Studies University of Maryland at College Park, College Park, MD 20742 {wmk,ben}@cs.umd.edu

ABSTRACT

Search engines are very effective at generating long lists of results that are highly relevant to user-provided query terms . But the lack of effective overviews presents challenges to users who seek to understand these results, especially for a complex task such as learning about a topic area, which requires gaining overviews of and exploring large sets of search results, identifying unusual documents, and understanding their context. Categorizing the results into comprehensible visual displays using meaningful and stable classifications can support user exploration and understanding of large sets of search results. This extended abstract presents a set of principles that we are developing for search result visualization. It also describes an exploratory study that investigated categorized overviews of search results for complex search tasks within the domain of U. S. government web sites, using a hierarchy based on the federal government organization.

Keywords

Categorized search results; search result visualization; information seeking; information retrieval; graphical user interfaces; human-computer interaction.

INTRODUCTION

The World Wide Web creates tantalizing opportunities for learning and research. Every day, information seekers attempt to find, organize, understand, and ultimately learn from information on the web. A teacher seeks current material to make lessons relevant to increasingly information-glutted students. A journalist uses the web to gather background material for a story about a newly patented device. These users struggle with information overload, coping with an overabundance of information that lacks a comprehensible organization.

Examining and understanding search results is a necessary step in a higher-level information seeking task (Marchionini, 1995). Based on their examination of search results, searchers take one of several possible actions, such as examining specific documents, refining/reformulating the query or revising their conceptualization of the information seeking problem. Search engines are very effective at returning large quantities of results that are highly relevant to user-provided query terms. But the lack of effective overviews presents challenges to users who seek to understand these results, especially for a complex task such as learning about a topic area, which requires gaining overviews of and exploring large sets of search results, identifying unusual documents, and understanding their context. The prevalence of short queries (Spink, Wolfram, Jansen, & Saracevic, 2001) produces large sets of results with high relevance to the provided query terms but limited situational relevance and begs for a more effective presentation.

Information visualization techniques help overcome limitations of typical text interfaces such as Google. Graphical presentations have been shown useful for providing overviews of large sets of categorical data. These techniques encourage exploration and understanding of data sets, helping users discover information that they would not otherwise find. The success of search result visualization, however, has been mixed.

Classifications and other knowledge structures can be applied to search result visualizations and used to organize the presentations of search results. Soergel (1999) observes that classifications, taxonomies and ontologies pro vide semantic roadmaps to fields of knowledge, improve communication and learning, and support information retrieval, among other benefits. They help searchers understand what concepts, ideas and relationships are relevant in a domain. Recent work on the presentation of textual result lists has shown that grouping by a subject classification speeds user retrieval of documents (Dumais, Cutrell & Chen, 2001). Other classifications, such as organization charts, and geographic and temporal hierarchies, can also be used to map search results. Clustering, in which documents are grouped by similarity measures, has also been shown to improve on ranked lists, although generating meaningful clusters and effective labels are recognized problems.

Cognitive and perceptual theories provide insight into how users scan, comprehend and assimilate visual information, but have not been adequately integrated with information-seeking theory to guide the display of and interaction with categorized search results. To that end we are developing a set of principles for search result visualization. We believe that categorizing search results into comprehensible visual displays using meaningful and stable classifications can support user exploration and understanding of large sets of search results. Displaying search results in the context of a known classification can help users understand the results better and lead to interesting findings that would not otherwise be observed.

This exploratory study investigates categorized overviews of search results for complex search tasks in U. S. government web sites. We are particularly interested in searchers' use of hierarchies to understand search results in a learning-oriented search task.

RELATED WORK

The most common presentation format for web search results is the textual list, typically showing document titles and a few other pieces of information such as URL and a snippet of text (possibly with matching query terms highlighted), but rarely with a category indicator. The results are typically ordered by a computed relevance rank. Drori & Alon (2003) compared four textual lists based on permutations of two variables (document category and lines from the document). The interface with categories and query-relevant lines from each document produced the fastest performance and was preferred by subjects. Dumais, Cutrell & Chen (2001) found that grouping web search results by a two -level subject classification expedited document retrieval.

Traditional online public access catalogs allow users to browse and search bibliographic databases usin g subject classifications. Allen (1995) describes two digital library interfaces based on two hierarchical classifications, the Dewey Decimal System and the ACM Computer Reviews classification. These interfaces show search results against the classification hierarchy and integrate several other features. HIBROWSE, an online public access catalog (OPAC) (Pollitt, 1997) exploits faceted hierarchies to provide visual query specification and to organize results. Flamenco (Hearst et al., 2002) provides interfaces to specialized collections (art, architecture and tobacco documents), using faceted hierarchies to produce menus of choices for navigational searching. The Technical Report Visualizer prototype (Ginsburg, 2004) allows users to browse a digital library by one of two user-selectable hierarchical classifications, which are displayed as hyperbolic trees (Lamping & Rao, 1996) coordinated with a detailed document list.

Clustering, in which documents are grouped by similarity measures rather than an explicit categorical attribute, has also been shown to improve on ranked lists, although generating effective labels is a recognized problem, as noted by Rivadeneira & Bederson (2003). Vivisimo

(www.vivisimo.com) uses this technique, producing an expandable menu of labeled clusters.

As noted above, the success of search result visualization has been mixed. Several web search (or metasearch) engines, including Grokker (www.grokker.com), Kartoo (www.kartoo.com), and FirstStop WebSearch (www.firststopwebsearch.com) incorporate visualization. Grokker has been found to compare poorly with textual alternatives (Rivadeneira & Bederson, 2003). The authors found that the textual interfaces were significantly preferred.

The GRiDL (Shneiderman, Feldman, Rose, & Grau, 2000) and WebTOC (Nation, Plaisant, Marchionini, & Komlodi, 1997) prototypes display search results using hierarchical categories, allowing users to drill down for details . WebTOC displays an expandable/collapsible outliner (similar to a tree widget), with embedded colored histograms showing quantitative variables such as size or number of documents under the branch. GRiDL uses a grid to display two categorical attributes of a collection of documents. The List and Matrix Browsers (Kunz, 2003) provide similar functionality. SuperTable integrates several information visualization techniques using linking and brushing to coordinate multiple tiled windows. Informal evaluations of these interfaces have been promising, but no extensive studies of the techniques have been published. Evaluations often indicate that interface effectiveness is dependent on the specific information-seeking task. (Risden, Czerwinski, Munzner, & Cook, 2000; Sebrechts, et al., 1999)

PRINCIPLES OF SEARCH RESULT VISUALIZATION

We are developing a set of search result visualization principles, based on the premise that consistent, comprehensible visual displays built on meaningful and stable classifications will better support user understanding of search results. As users explore search results, they are grappling with multiple simultaneous information problems: Their conceptualization of the high-level information need is imperfect and evolving; their understanding of the relevant concepts and terminology is limited, and their understanding of the presentation and interactions available in the interface is incomplete (Marchionini, 1995). We claim that helping users to incrementally solve these problems enables them to make more effective progress toward their high-level objective. The principles described here begin to address these problems.

Represent a sufficient number of results

The visual-information-seeking mantra (Shneiderman & Plaisant, 2004) prescribes, “Overview first…” and a search result visualization must provide a sufficient quantity of results to start with. We postulate that at least 100 -1000 results will be required to form the basis of an overview, but the ideal number will certainly depend on many factors, including the task domain, the quality of documents and search engine capabilities.

Organize results by meaningful, stable classifications

Gaining an overview of search results involves a number of cognitive subtasks, including interpretation of the results within the context of the user’s internal mental model of the knowledge domain. Using meaningful, stable classifications to organize results will place each result in a known context. The closer the classification matches the information need, the better, but even if the selected classification is not a direct match, we believe that the context provided will be valuable.

Provide examples of documents for each category

Within a category, example results help to clarify ambiguous or unfamiliar labels, and often provide an indication of relevance, quality, etc. Dumais, Cutrell, & Chen (2001) noted that individual page titles helped disambiguate category names in their study of search results.

Use a stable visual substrate

Stable, consistent and meaningful displays have been shown to promote success in user interfaces (Tullis, 1988; Shneiderman & Plaisant, 2004). Niemela & Saariluoma (2003) demonstrated the importance of both visual syntax (spatial layout) and semantics (labels) in learning a visual display. Organizing the results in a manner consistent with the classification will minimize unnecessary cognitive activity and allow users to focus on the task at hand rather than interpreting the presentation of the results.

Visually encode quantitative attributes

Information visualization principles are grounded in our understanding of human perception and cognition, particularly the structure, functions, strengths and limitations of these systems. Visualization techniques augment human perceptual and cognitive systems by encoding data into visual constructions that are appropriate to and supportive of processing at each perceptual stage and within the cognitive system (Card, Mackinlay, & Shneiderman, 1999). Quantitative attributes such as dates or document counts and nominal attributes with a small range of values such as document types can be visually encoded.

Arrange important text for scanning/skimming

Users of search results attempt to rapidly ingest large amounts of text. In an initial user study, users scanned titles and snippets of text, quickly selecting specific pages to view. They skimmed the pages, and returned to the list to repeat this cycle. Arranging these elements in linear lists, columns, or matrices and ensuring that they are visible

(rather than requiring interaction such as moving the pointer over an item) will support fast scanning and skimming. Appropriate use of font weights, styles, sizes and colors will also help.

Support multiple visual presentations and classifications

No single classification or presentation is likely to be effective for all domains. Users should be able to select the desired form of data display (Shneiderman & Plaisant, 2004). Supporting multiple visual presentations and classifications will allow users to view and explore the search results from the perspectives most appropriate to their needs.

EXPLORATORY STUDY

This exploratory study, a first step to validate and refine these principles, used the context of government web sites (Kules & Shneiderman, 2004). The study focused on the process of exploring search results, rather than query reformulation, query refinement or relevance measurements. This experiment compared presentations of search results with and without hierarchical overviews. The control condition displayed search results in a manner similar to Google, adding the government department and agency. It provided no overview. Two experimental conditions used overview+detail interfaces (an expandable outliner and a treemap, figures 1 and 2) allowed users to limit the displayed lis t of results by selecting (clicking on) a single category.

A 1x3 between groups design (N=18) was used, with interface type as the independent variable. Participants were asked to perform 3 tasks for each of 3 scenarios. Each scenario was used to motivate a pre-specified query and its corresponding search results (for the queries “breast cancer”, “alternative energy” and “urban sprawl”). We used the government agency hierarchy to categorize the search results. Each result set has approximately 200 result items. All results were categorized into the leaf nodes of a broad, shallow, 2-level government agency hierarchy. Tasks included finding the agency with the most results, finding web pages that illustrate different perspectives within the results, and an open-ended task to identify any unusual findings. We hypothesized that users would be more successful and more satisfied with the overview conditions. We also hypothesized that the overviews (and particular the treemap) would be judged as more complex and mo re difficult to learn. Finally, we expected that the overview interfaces would help users find unusual results more effectively, but that the hierarchy would bias the kinds of unusual results identified (corresponding to category differences).

The results showed that the overview conditions produced significantly higher successful completion rates for the task of identifying the agency with the most pages. The subjective measures showed that the overview treatments were preferred and this was supported by user comments. They found the overviews significantly easier to use, more helpful, and more satisfying than the control, and they were more confident of their own success. They agreed more strongly that they had gained a good overview and found good examples of different perspectives. The results did not support the hypothesis that the overview tasks would be considered more complex and difficult to learn. Users found all three interfaces very understandable and fairly simple.

Figure 1. Detail of the expandable outliner condition, showing the tree control used to display an overview of all search results, limiting the list to the National Park Service.

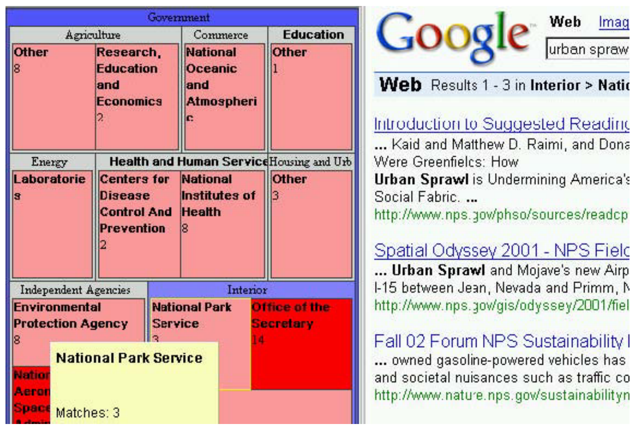

Figure 2. Detail of the treemap interface, which uses nesting to show both top and second-level categories simultaneously.

Having the overview available helped users to notice areas not covered by the search results. Only one of the six control users found it surprising that an agency had few or no results, whereas nine of the 12 overview users at some time found this surprising. Several users made comments like "what I found informative was… what didn't show up, which I wouldn't know if the hierarchy wasn't there." People used their prior knowledge to interpret search results. When interfaces provide affordances to organize search results – even when the organizing structure is not optimal for the task – users adapt their tactics to take advantage of their knowledge. Users also appeared to

become more familiar with the hierarchy over the course of the experiment, suggesting a learning effect that would be beneficial in successive searches.

Observations and subject comments confirm that text is important, even with the overviews available. As one person noted, the overview is a starting point. But users still needed to scan substantial amounts of text. Simply having category information available is not enough to help users gain an overview of search results. It must be appropriately integrated with the text.

Most users preferred the expandable outliner, but several users found the treemap more appealing. Based on subject comments, it is clear that additional user control of the overview is desirable. This includes allowing users to select the desired presentation, as well as creating or selecting the categorization scheme used.

It should be noted that these results must be interpreted within the context of the specified tasks and domain. Users who had been asked to perform narrower tasks (e.g. known-item retrieval) would probably have produced different results. The hierarchy used was limited in size and the specific tasks were somewhat artificial, and may not be fully representative of the types of tasks users perform in real-world topic searches. However, the types of behaviors they elicit do transfer: Examining large number of results, evaluating them in the context of current knowledge and identifying unusual items are characteristic of complex information-seeking tasks.

FUTURE WORK

The principles described above do not yet address a number of issues, including how much stability is needed in the visual substrate versus how much variability can be tolerated, what the permissible trade-offs are, and how much context is needed when navigating search results. These issues and others will be investigated as the principles are tested and refined during the course of our research.

The results of this exploratory study will be used to refine our understanding of the role of spatial layout, textual elements and dynamic interactions of categorized search result visualization. Additional domains and hierarchies will be investigated, including web search with the Open Directory (DMOZ) subject hierarchy and a search interface for a large collection of oral history videos that utilizes a specially developed thesaurus.

CONCLUSION

Categorizing search results into comprehensible visual displays using meaningful and stable classifications can support user exploration and understanding of large sets of search results while alleviating information overload. This extended abstract presents a set of principles that we are developing for search result visualization. Results of an exploratory study have begun to validate these principles and suggest additional refinements. As these principles are extended, refined and confirmed, they can be used by practitioners – the designers and developers of Web search engines – to help realize more effective interfaces for learning and research on the Web.

ACKNOWLEDGEMENTS

This research is supported by an AOL Fellowship in Human-Computer Interaction and National Science Foundation Digital Government Initiative grant (EIA 0129978) “Towards a Statistical Knowledge Network.” We thank the anonymous reviewers for their helpful comments.

REFERENCES

1. 1. Allen, R. (1995). Two digital library interfaces that exploit hierarchical structure. *Proc. DAGS95: Electronic Publishing and the Information Superhighway*.
2. 2. Card, S., Mackinlay, J., & Shneiderman, B. (1999). Readings in Information Visualization: Using Vision to Think. San Francisco: Morgan Kaufmann.
3. 3. Drori, O., & Alon, N. (2003). Using documents classification for displaying search results list. *Journal of Information Science, 29*(2), 97-106.
4. 4. Du mais, S., Cutrell, E., & Chen, H. (2001). Optimizing search by showing results in context. *Proc. SIGCHI Conference on Human Factors in Computing Systems*, 277-284.
5. 5. Ginsburg, M. (2004). Visualizing digital libraries with open standards. *Communications of the Association for Information Systems, 13*, 336-356.
6. 6. Hearst, M., Elliot, A., English, J., Sinha, R., Swearingen, K., & Yee, P. (2002). Finding the flow in web site search. *Communications of the ACM, 45*(9), 42-49.
7. 7. Kules, B., & Shneiderman, B. (2004). Empirical comparisons of categorized graphical overviews for web search results. *In process*.
8. 8. Kunz, C. (2003). SERGIO - An interface for context driven knowledge retrieval. *Proceedings of eChallenges, Bologna, Italy, 2003* .
9. 9. Lamping, J., & Rao, R. (1996). The hyperbolic browser: A focus + context technique for visualizing large hierarchies. *Journal of Visual Languages and Computing, 7*(1), 33-55.
10. 10. Marchionini, G. (1995). Information Seeking in Electronic Environments: Cambridge University Press.

1. 11. Nation, D. A., Plaisant, C., Marchionini, G., & Komlodi, A. (1997). Visualizing websites using a hierarchical table of contents browser: WebTOC. *Proc. 3rd Conference on Human Factors and the Web*.
2. 12. Niemela, M., & Saariluoma, P. (2003). Layout attributes and recall. *Behaviour & Information Technology, 22*(5), 353-363.
3. 13. Pollitt, S. (1997). Interactive Information Retrieval based on Faceted Classification using Views in Knowledge Organization for Information Retrieval.

Proc. 6th International Study Conference on Classification Research.

1. 14. Risden, K., Czerwinski, M., Munzner, T., & Cook, D. (2000). An initial examination of ease of use for 2D and 3D information visualizations of Web content. *International Journal of Human-Computer Studies*, 695-714.
2. 15. Rivadeneira, W., & Bederson, B. B. (2003). A Study of Search Result Clustering Interfaces: Comparing Textual and Zoomable User Interfaces. *University of Maryland HCIL Technical Report HCIL-2003-36*.
3. 16. Sebrechts, M., Vasilakis, J., Miller, M., Cugini, J., & Laskowski, S. (1999). Visualization of search results: a comparative evaluation of text, 2D, and 3D interfaces.

Proc. 22nd AnnualInternational ACM SIGIR Conference on Research and Development in Information Retrieval, 3-10.

1. 17. Shneiderman, B., Feldman, D., Rose, A., & Grau, X. F. (2000). Visualizing dig ital library search results with categorical and hierarchial axes. *Proc. 5th ACM International Conference on Digital Libraries*, 57-66.
2. 18. Shneiderman, B., & Plaisant, C. (2004). Designing the User Interface: Strategies for Effective Human-Computer Interaction (4th ed.). Boston: Pearson/Addison-Wesley.
3. 19. Soergel, D. (1999). The rise of ontologies or the reinvention of classification. *Journal of the American Society for Information Science and Technology, 50*(12), 1119-1120.
4. 20. Spink, A., Wolfram, D., Jansen, B. J., & Saracevic, T. (2001). Searching the web: The public and their queries. *Journal of the American Society for Information Science, 52*(3), 226-234.
5. 21. Tullis, T. (1988). Screen design. In M. Helander (Ed.), *Handbook of Human-Computer Interaction* (377-411). Amsterdam, The Netherlands: Elsevier Science Publishers.