Augmenting Digital Library Search Interfaces with Visual Analysis Tools

Edward Clarkson and James D. Foley

Georgia Institute of Technology

ABSTRACT

Digital libraries commonly elide hierarchical metadata that might be used more effectively. This proposal presents the ResultMap concept, a tool that leverages that metadata for digital library search facilities; an initial study of its effectiveness; the concept of applying ResultMaps to faceted metadata, allowing visual detection of implicit correlations between facets; and proposals for further study of ResultMaps in both directed search and faceted browsing environments.

CR Categories and Subject Descriptors: H.5.4 [Hypertext/ Hypermedia]: Navigation, User issues.

Additional Keywords: user studies, exploratory search.

1 INTRODUCTION

Over the past 3 years, we have built two digital libraries for educational materials: one for Visual Analytics[1] and one for Human-Centered Computing[2] (HCC) [1]. A common theme among digital repositories, including ours, is the use of either pre-structured or generated hierarchy to organize content. Hierarchies like ours may be stand-alone or part of a *faceted metadata* scheme. Faceted metadata is a way of classifying an object along multiple independent dimensions. For example, a research paper might have facets for its topic, its publication venue or its length.

We are interested in how digital repositories can use visualization to leverage these existing structures, and studying how such techniques affect users. We focus in particular on the areas of traditional keyword search and faceted browsing interfaces. We have augmented a standard search engine with what we call *ResultMaps* [2] (RMs), and we further propose to study their applicability to faceted environments.

2 MOTIVATION AND RESEARCH QUESTIONS

There has been a variety of work in the area of visualization for digital libraries (DLs) and generalized search visualization. The standard paradigm for keyword search result listings (a list of relevance-ordered document hits) does little to contextualize the results within the overall information space, and existing hierarchical classifications are ideal for this purpose. Consequently, search environments do not fully leverage the utility of hierarchy's inherent metadata.

Several factors make this proposal distinct from previous work (too extensive to mention in this format): we are interested specifically in the digital library usage environment; we deploy our tools on *live* repositories; and we propose questions which have not previously been studied in such contexts.

2.1 Directed Search

ResultMaps are based on the popular treemap layout technique [3]. They take a digital repository that uses some form of hierarchical classification and uses that structure to map each document into a treemap, and search engines can highlight query result items. Representing the entire document space provides context for individual search results, and facilitates tasks like detecting outliers and clusters within search results by them making visual instead of textual processes. Moreover, this context is consistent between successive search queries.

Both the Visual Analytics Digital Library (VADL) and HCC Education Digital Library (HCC EDL) have implemented ResultMaps as a part of their search engine facilities. Space constraints prevent us from including an illustration; we encourage the reader to visit either the VADL or HCC EDL for an example. We have used these environments to address a number of larger research questions regarding directed search tasks:

1) How do RMs affect user performance on DL search tasks?
2) How do RMs affect subjective impressions of DL interfaces?
3) Do RMs yield a greater level of knowledge about the overall content in digital library?
4) How do RMs affect query string characteristics over sequences of queries?

2.2 Faceted Environments

Faceted browsing environments do an admirable job of supporting painless transition between directed and exploratory search activity. But one problem with these environments is that it can be difficult to diagnose the effects of an action after the fact—that is, how a selection in one facet affects the distribution of remaining items across the other facets. The only information present in current tools (such as Flamenco [5] among several) that might assist such diagnosis is the raw item counts remaining in each facet category. Indeed, designers of another faceted environment suggest "…representing temporary hierarchies to the user is important for…understanding the effects of facets on each other" [4]. The relationship between two facets can be complex, but can be quantified by the correlation between their respective metadata values. For example, a dataset may contain subtle relations like the connection between medium and country of origin in a visual artworks dataset. But the frequency and magnitude of inter-facet correlations is unknown, and forms an initial exploratory research question:

5) What is the extent of facet correlation in real-world and standardized sample datasets?

Another salient observation of our ResultMap environment is that it is essentially a degenerate case of a generalized faceted system in which a library has a single hierarchical facet. As a result, we can apply ResultMaps to a general faceted environment using one ResultMap per facet—what we call *faceted ResultMaps*. Faceted ResultMaps allow users to use visual analysis to intuitively discern correlations—users may notice that a selection in one facet non-randomly culls items in another facet. But faceted ResultMaps are merely one approach to representing dataset correlations: a simpler alternative is to link visually facets

801 Atlantic Drive, Atlanta, GA 30332-0280
{edcclark, foley}@cc.gatech.edu

[1] http:/vadl.cc.gatech.edu
[2] http://hcc.cc.gatech.edu

with strong correlations. Accordingly, our research questions concern the objective and subjective effectiveness of faceted ResultMaps in digital library environments:

6) How do faceted ResultMaps and other correlative displays affect performance on exploratory search tasks?
7) How do faceted ResultMaps and other correlative displays affect subjective attitudes of faceted DL interfaces?
8) How do faceted ResultMaps and other correlative displays affect query sequences and/or facet selection paths?

3 Methodology

We address our research questions through a series of controlled and field studies. We have addressed questions 1-3 in a recent study A1; study A2 approaches 1-3 with less experimental control but more ecological validity; questions 4 and 5-8 are addressed in studies B-D. We summarize the design of each study below.

3.1 Study A1 (Questions 1-3)

The design is between subjects with two levels: ResultMap vs. non-ResultMap (non-RM). The dependent measures include task completion time, accuracy, and post-tests of users' knowledge of and subjective attitudes about the DL environment.

Subjects are presented with a series of 12 search result listings (using 10 results per page) from the HCC EDL, randomly ordered for each subject. For each result list, subjects are asked to select a document satisfying a set of constraints. The RM condition is the search engine interface in use with the HCC EDL; the non-RM interface simply removes the RM interface piece, making it similar to a standard search engine (e.g., Google). Query strings for the tasks are the same for all subjects to control for search expertise. After subjects complete all tasks, they are asked to complete a set of questions testing their knowledge about characteristics of the repository as a whole and about their subjective impressions of the library.

3.2 Study A2 (Questions 1-3)

ResultMaps are in use with each of our digital libraries as mentioned above. That installation has yielded a large (ever-expanding) corpus of log data for analysis of *in-situ* user behavior. We are implementing a semi-controlled longitudinal field replication of study A1 by randomly presenting the ResultMap or the traditional search interface to users.

3.3 Study B (Question 4)

The design is within subjects with two levels of interface type: ResultMap (RM) vs. non-ResultMap (non-RM). The dependent measures include task time and accuracy, number of query strings, average query length, query string change between successive queries, and subjective user ratings.

Subjects are asked to use the library search engine to find two series of documents meeting a series of related constraints. The series are randomly assigned to condition for each subject, and conditions are counterbalanced between subjects. Subjects are pre-tested on their query string generation skill.

3.4 Study C (Question 5)

This is a non-experimental study aimed at gathering data about correlations in real-world datasets. Assuming there are effective techniques for showing correlation data, the prevalence of correlated facets provides a baseline for how often those techniques might be employed. Examples of real-world faceted datasets to examine include eBay auctions, the Flamenco dataset examples, and faceted versions of our own digital libraries.

3.5 Study D (Question 6-8)

The design is within-subjects with three levels of interface type: RM, linked, and baseline. The baseline control condition is an unaugmented faceted environment; the linked condition is a faceted environment that draws connections between facets that are significantly correlated; the RM condition is a faceted environment augmented with ResultMaps. We use a larger, non-visualization dataset than our DLs, allowing us to control for visualization expertise and to test on multiple corpora sizes.

The dependent measures include task time, accuracy, number of facet selections and queries, query string characteristics (length, etc.), and subjective ratings. Subjects are given both unstructured and structured tasks requiring them to interleave both faceted browsing and directed search behaviors.

4 Results

We have completed Study A1 (N=20). A preliminary analysis of the data yields mixed findings:

- There were no significant differences in performance or accuracy between groups (like many other studies of visualization tools).
- RM users scored significantly higher on some parts of the repository knowledge post-test that the control group.
- RM users rated the subjective impact of the interface on task difficulty significantly better than the control group.

More detailed analysis is underway to explore any effects of result list characteristics (e.g., clustering or outliers) on performance.

5 Future Work and Conclusions

Our early results show that ResultMaps have some positive benefits for search users, though they show no statistical benefits in many cases. Unfortunately, this is typical of many studies of visualization systems. But ResultMaps do make use of hierarchical metadata typically ignored by search interfaces, which at least provides the opportunity for users to gain unsought knowledge—a situation difficult to experimentally verify or quantify—of either ancillary topics or the digital library information space as a whole.

Faceted systems also do not typically reveal the full extent of the available hierarchical metadata; that fact also makes it difficult to identify patterns or linkages between facets. As a result, the application of ResultMaps to faceted environments is an intriguing possibility for further study and refinement. Here, we describe one such path of research in this problem space.

References

[1] Clarkson, E., Day, J. and Foley, J. An Educational Digital Library for Human-Centered Computing. In *CHI '06 Extended Abstracts*, pp. 646-651.

[2] Clarkson, E. and Foley, J. ResultMaps: Search Result Visualization for Hierarchical Information Spaces. In *Proceedings of the CHI '07 Workshop on Exploratory Search and HCI*, pp. 32 - 39.

[3] Johnson, B. and Shneiderman, B. Tree-maps: A space filling approach to the visualization of hierarchical information structures. In *Proc. of IEEE Visualization '91*, pp. 284-291.

[4] Wilson, M., schraefel, m. c. and White, R. Evaluating Advanced Search Interfaces using Established Information-Seeking Models. Technical Report (2007), School of Electronics and Computer Science, University of Southampton.

[5] Yee, P., Swearingen, K., Li, K. and Hearst, M. Faceted Metadata for Image Search and Browsing. In *Proc. of CHI '03*, pp. 401-408.