

Hierarchical Faceted Metadata in Site Search Interfaces

Jennifer English, Marti Hearst, Rashmi Sinha, Kirsten Swearingen, Ka-Ping Yee

Group for User Interface Research

University of California, Berkeley

Berkeley, CA 94720-4600

{jenglish,hearst,sinha,kirstens,ping}@sims.berkeley.edu

ABSTRACT

One of the most pressing usability issues in the design of large web sites is that of the organization of search results. A previous study on a moderate-sized web site indicated that users understood and preferred dynamically organized faceted metadata over standard search. We are now examining how to scale this approach to very large collections, since it is difficult to present hierarchical faceted metadata in a manner appealing and understandable to general users. We have iteratively designed and tested interfaces that address these design challenges; the most recent version is receiving enthusiastic responses in ongoing usability studies.

Keywords: Search User Interfaces, Iterative Design, WWW

INTRODUCTION

Although general Web search is steadily improving, studies show that search is still the primary usability problem in web site design. We are examining how faceted metadata can be used to tie search results into the information architecture of a web site. We use the term *faceted metadata* to mean orthogonal descriptors within a metadata system. For example, different facets associated with architecture include Materials (concrete, brick, wood, etc.), Styles (Baroque, Gothic, Ming, etc.), View Types (interior view, exterior view, etc.), People (architects, artists, developers, etc.), and so on.

Although much study has been done on how to *create* metadata, very little has been done on how to *use* it explicitly in search user interfaces, and whether or not doing so is even a good idea. Our design goals are to use metadata to create the navigation structure of a web site; to organize the search results; and to suggest, via query previews [2], the next steps a user might take within a search session. The interface should allow users to use hyperlinks, rather than, or in addition to, keyword search, to flexibly explore a collection's contents along several different dimensions simultaneously. The interface should also allow an easy flow between refining and broadening the query, and should support both directed and undirected exploration.

As a prelude to designing our own interface, we conducted a small exploratory study of a commercial recipe interface offering three search options: a standard search form, a search form with faceted metadata, and a browsing interface that makes use of faceted metadata and dynamic query previews [1]. Our results suggest that for many tasks, the metadata-based approaches are superior to a standard search form and linear results listing. They also suggest that, when given a choice between a search-centric and a browse-centric interface, users recognize the utility of switching to the interface that best supports the current task.

Although the study results indicated the utility of explicit metadata, the collection used was of only moderate size (about 13,000 items and 90 metadata choices), and the metadata was organized into flat (non-hierarchical) facets. The question remains of how to make such an approach scale to larger collections; thus, we are applying iterative design principles to the problem of how to incorporate hierarchical faceted metadata and query previews to create an intuitive, accessible search interface. We are using our own software to allow for direct contrasts between different interface features. Although our software will work with any collection for which faceted metadata is available, the initial test collection consists of over 40,000 images from an architecture slide library. Each slide has approximately 10 items of metadata assigned to it, and the metadata hierarchy contains 9 facets and several thousand metadata terms.

INTERFACE DESIGN

We have iterated on the design of this interface several times. In August 2001 we examined three versions against a control search-only interface. The overall results revealed an interest in the metadata approach, as well as which features succeeded and which failed, but two of the preliminary interfaces elicited generally poor scores for ease of use and other preference measures. However, one of the experimental interfaces, which we called the Matrix view, generated somewhat more enthusiastic responses.

Informed by that study, we made significant changes to the Matrix view, and incorporated other results of the study. We highlight important features of this interface in Figure 1. Additional important features are: (1) When the user selects an image thumbnail, the full image is shown along with hyperlinks to its associated metadata. In the example above, click-



Figure 1: Site search interface with query previews on hierarchical faceted metadata (images are blanked out for intellectual property considerations). Two facets are selected (Locations and Structure Types) with results grouped by a third facet (View Types). The previews on the lefthand side are updated as constraints are added or removed. Clicking on the group heading “city details” refines the query into that category. The boxes at the top display the current query; clicking on “Western Europe” relaxes the Location constraint, and on the X button removes the location facet from the query. Images can also be sorted by the three most important metadata types: Architect, Location, and Building Name (this feature is available only in ungrouped view).

ing on the thumbnail labeled “Cannes” reveals the potentially surprising fact that there is a metadata category called “roof,” which can then be selected to form a new query, thus acting as an expansion (as opposed to refinement) feature. (2) Since keyword queries usually match multiple category terms, their results are shown as a list of hits contextualized by their location within the facet hierarchy. (3) Intermediate pages are used to handle large result sets, both as a result of keyword searches and of selecting category hyperlinks. For example, there are over 16,000 images associated with architects in the collection, so clicking on the architects link before refining with other facets brings up an A-Z style listing; selecting a category such as “Wright” within this listing returns the main view, showing only those designs by Wright. The innovation here is to always show an intermediate page when results of any form are too large. Additionally, these intermediate pages vary in their presentation depending on the number of items: 1 to 50 categories are arranged in a simple two-column listing; 51 to 500 categories are shown in alphabetized sections in two columns; more than 500 categories are arranged in alphabetized sections with two columns within each section and a hyperlinked A-Z index at the top. Usability studies (described below) reveal that users find this use of intermediate pages to be intuitive.

USABILITY STUDY

We are currently (December 2001) running new studies, again with architects, and the response (from 12 participants) is now highly enthusiastic. Participants are assigning an average of at least 6 on a 7 point scale on nearly every measure, including likelihood of future use, ease of use, and usefulness for: design work, exploring a collection, finding a particular image, and seeing relationships. Several participants have expressed a strong desire to have access to the interface immediately. This is especially significant since users are often reluctant to switch from a simple search form and results listings to more complex or innovative interfaces. Thus this work suggests that hyperlinked, faceted, dynamically-generated metadata can be used as a viable alternative to keyword based queries for web site search.

This research was supported by NSF 9984741.

REFERENCES

1. Jennifer English, Marti Hearst, Rashmi Sinha, Kirsten Swearington, and Ping Yee. Examining the usability of web site search, 2002. MS, in progress.
2. Catherine Plaisant, Ben Shneiderman, Khoa Doan, and Tom Bruns. Interface and data architecture for query preview in networked information systems. *ACM Transactions on Information Systems*, 17(3):320–341, 1999.