Examining the Usability of Web Site Search

Jennifer English, Marti Hearst, Rashmi Sinha, Kirsten Swearington, and Ping Yee
School of Information Management & Systems
University of California, Berkeley
Berkeley, CA 94720-4600
{jenglish,hearst,kirstens,sinha,ping}@sims.berkeley.edu

ABSTRACT

One of the most pressing usability issues in the design of web sites is that of how to improve navigation and search. We are conducting a series of usability studies to address this problem, focusing on web sites that consist of large collections of loosely organized information. This article describes our method and presents preliminary results which suggest that use of faceted metadata can be useful both for the initial stages of highly constrained search and for the intermediate stages of less constrained browsing tasks. We also find that users state an interest in using different search interface types to support different search strategies.

INTRODUCTION

Although general Web search is steadily improving [14], studies show that search is still the primary usability problem in web site design. A recent report by Vividence Research analyzing 69 web sites found that the most common usability problem was poorly organized search results, affecting 53% of sites studied. The second most common problem was poor information architecture, affecting 32% of sites [12].

There have been few usability studies that reveal how to build effective web site search interfaces, despite the fact that the essence of navigating a website is the interplay between hyperlinks and search results. To address this gap, we are engaged in a series of usability studies whose goal is to develop empirically grounded recommendations for this problem. We are also developing a suite of software tools that allow for the flexible design of different variations of the user interface, to facilitate comparison and isolation of the effects of particular features. Although we have hypotheses about which interfaces should work best and why (informed by several years experience in designing and evaluating search user interfaces) this setup allows us to determine objectively and systematically which features work well, and which do not.

Currently we are examining how metadata can be used to tie search results into the information architecture of a web site [7]. Metadata is used both to create the navigation structure of the web site and to organize search results. It is also used for suggesting, via query previews [10], the next steps a user might take within a search session.

Below we describe how metadata can be used in search interfaces, our methodology for assessing web site search interfaces, and the results of a small usability study that compared three interfaces: a standard search form, a search form with faceted metadata, and a browsing interface that makes use of faceted metadata and dynamic query previews. Our initial results suggest that for some tasks, the metadata-based approaches are superior to the standard search form, and furthermore suggest that users prefer different search interface types for different search strategies.

CHARACTERISTICS OF THE SEARCH PROCESS

Many researchers have observed that there are several different types of search goals, engendering different types of search strategies [9]. The logical extension of this idea, recognized by this research effort, is that different interfaces are suited to different types of search strategies. Indeed, our initial results support this hypothesis. In particular, "known-item" searches and highly constrained searches tend to work better with a keyword-based search interface, whereas less constrained or more open-ended tasks tend to work better with interfaces that support navigating the information structure, or a combination of keyword querying and navigation. This need to differentiate the search interface according to the search strategy has long been suspected, but few interfaces have been promoted that support multiple types of search behavior (for a review of studies of this kind, and of search user interfaces in general, see [6]).

Another frequently mentioned inadequacy of search user interfaces is their inability to support the intermediate

stages of search. Like the game of chess, the search process has an opening, a middlegame, and an endgame. The choice of a chess opening is very important, well-studied, and has a profound effect on the overall structure of the game. The endgame is well-defined in terms of the pieces and positions necessary to deliver checkmate. The middlegame is more difficult to characterize and requires a flexible, opportunistic combination of tactics and strategies.

The opening in a web search user interface usually consists of one of the following: one or more entry fields in which the user can type query terms; a set of hyperlinks that allow navigation within subsets of the collection; or a set of checkboxes and list boxes that allow specification of constraints for the items to be selected. The endgame usually consists of the detailed display of the content of one or a few items – an article, a web page describing a product, or a small set of images, for example.

The middlegame that links these two is generally not well understood. The standard search interface middlegame is the rank-ordered display of a list of items which match the query to some degree. This one-shot, stateless approach can be adequate for well-constrained tasks, but there is general agreement that current search interfaces do not support intermediate results well [4]. It is often argued that richer interfaces are needed for situations in which users are less certain of what they are looking for, and to help users learn about the collection and the subject matter as the session progresses [1, 8].

One of the questions we are attempting to answer is how best to support the middlegame – the intermediate stages of a search session. We are doing this by comparing different approaches and by probing the participants' mental state while engaged in the process. This includes asking questions about whether or not the participants feel they know where to go next, how confident they are about the path they are currently pursuing, how often they back up or get empty results, and other metrics.

We are also investigating the notion of "scent" for the design of web site information architectures. We can draw an analogy between the process of hunting for information and the process of tracking an animal. If the scent disappears, the trail grows cold, and we are no longer sure where to go next. Within web site navigation structures, a site has good scent if the user always has an idea of what link to click on next, given where they currently are and what their goal currently is. As long as it is clear which step to take next, the user will not become disoriented or frustrated. Jared Spool argues that the number of links a users needs to follow to find some information is less important than whether or not the scent is adequate at all stages of the search [13]. Work at Xerox PARC is also studying the role of scent in web user interface design [3]. However, these researchers study log files of existing web sites to try to infer users' tasks and the paths they take to reach their goals.

item: A piece of content in a collection.
metadata: The classification data about a particular item, usually expressed as a set of *attributes*.
attribute: A predicate about a particular *item*, such as "costs $57" or "was painted in 1923".
facet: A subjective category under which attributes are organized, such as "price" or "period" or "location". (Attributes are intrinsic, whereas facets are designed.)
facet value: An attribute in the context of a particular facet. For example, "less than $100" could be an allowed value of the "price" facet for an item that costs $57.
flat facet: A facet whose values are organized in a list.
hierarchical facet: A facet whose values are organized in a hierarchy.

Table 1: Terminology used throughout the paper.

Our research, in contrast, consists of live usability studies in which the participants' goals are known and different interfaces are compared directly. Thus we can ask questions about which interface type is preferred, or more effective, for each type of task.

USING METADATA TO INTEGRATE SEARCH AND NAVIGATION

Previous research suggests that users often wish to see the results of a keyword search organized in some manner. Many researchers have applied text clustering to retrieval results; this method is attractive from the point of view of implementation, because the groupings can be determined automatically. Unfortunately, studies show that ordinary users find the results of clustering to be difficult to interpret. Instead, they prefer the predictable organization of category hierarchies [11, 2].

A different way to organize search results is to arrange them according to categorical metadata [7]. Much work has been done on automating the mapping of query terms into metadata categories, for example, converting the query "heart attack" into the pre-defined term "myocardial infarction," and thus increasing the likelihood of making relevant matches [5]. However, here we are interested in a different problem; namely, how to use the metadata directly in the web site user interface, both as a starting point for the search and as a structure upon which to organize the results of a keyword-based query. Table 1 defines metadata and related terminology used throughout the rest of this article.

Content-oriented category metadata has become more

prevalent in the last few years, and many people are interested in standards for describing content in various fields. Web directories such as Yahoo and the Open Directory Project are familiar examples of the use of metadata for navigation structures. Web search engines have begun to interleave search hits on category labels with other search results. In addition to the web as a whole, many individual collections already have rich metadata assigned to their contents; for example, biomedical journal articles have on average have a dozen or more content attributes attached to them.

Usage of metadata for organizing content collections can be classified in a few ways. It is worth calling attention to a number of properties that may be present or not present in a metadata classification system.

- First, the metadata may be *faceted*, that is, composed of orthogonal sets of categories. For example, in the domain of architectural images, some possible facets might be Materials (concrete, brick, wood, etc.), Styles (Baroque, Gothic, Ming, etc.), View Types (interior view, exterior view, etc.), People (architects, artists, developers, etc.), Locations, Periods, and so on.
- Second, the metadata (or an individual facet) may be *hierarchical* ("located in Berkeley, California, United States") or *flat* ("by Ansel Adams").
- Third, the metadata (or an individual facet) may be *single-valued* or *multi-valued*. That is, the data may be constrained so that at most one value can be assigned to an item ("measures 36 cm tall") or it may allow multiple values to be assigned to an item ("uses oil paint, ink, and watercolor").

We believe that the full potential for the use of metadata in search results has not been realized. We suspect that, if properly applied, faceted metadata can markedly improve web site search, especially for large collections of similar-style items (such as product catalog sites, or sites consisting of collections of images or text documents on a topic such as medicine or law). However, there is little empirical evidence examining when and how ordinary users find such organizations useful. In the next section we describe a web site that makes use of faceted metadata in the search interface in a flexible manner.

Thus, we are investigating how best to incorporate descriptive metadata into the search process, both in the initial phrasing of the search and in integration into the results listings. We are particularly intrigued by the following questions: are flexible reorganizations of metadata from different hierarchies preferable to (and as understandable as) static hierarchies? Are faceted hierarchies preferable to strict hierarchies? Should retrieval results be organized according to the structure of this kind of metadata, and if so, how? Do people prefer to follow category-based hyperlinks? Do they prefer to issue a keyword-based query and sort through the results listings? Are there individual differences? Does the best method depend on the task at hand?

We note that there are a number of issues associated with creation of metadata itself which we are not addressing here. The most pressing problem is how to decide which descriptors are correct or at least most appropriate for a collection of information. Another problem relates to how to assign metadata descriptors to items that currently do not have metadata assigned. We will not be addressing these issues, in part because many other researchers already are, and because the fact remains that there are a number of existing collections with hierarchical metadata already assigned, usually by hand but sometimes by semi-automated text categorization methods.

EPICURIOUS

We wanted to ensure that the idea of using metadata to organize search results is a feasible one before writing code of our own to support this functionality. For this reason, we began the investigation by using an existing award-winning web site, called Epicurious, which contains a large collection of recipes. This is a good demonstration collection because the meaning of recipes and their associated metadata are familiar to most people. The facets used by Epicurious are Main Ingredient, Cuisine, Preparation Method, Season/Occassion, and Course/Dish. Each of these has one level of categories; for example, categories for Course/Dish include Appetizers, Bread, Desserts, Sandwiches, Sauces, Sides and Vegetables. The collection contains more than 13,000 recipes.

The Epicurious site provides three interfaces for searching for recipes:

- A Basic Search form that allows users to enter any number of words and shows the results as a list of titles in an (apparently) arbitrary order (see Figure 1).
- An Enhanced Search form that exposes much of the metadata to the user in the form of checkboxes and drop-down lists, augmented by text entry boxes for keyword terms (see Figure 2). Radio buttons indicate whether selections within a facet are to be ANDed or ORed together. There is an implicit AND across all facet selections. Results are shown as a list of titles, as in basic search.
- A Browse interface that allows the user to navi-

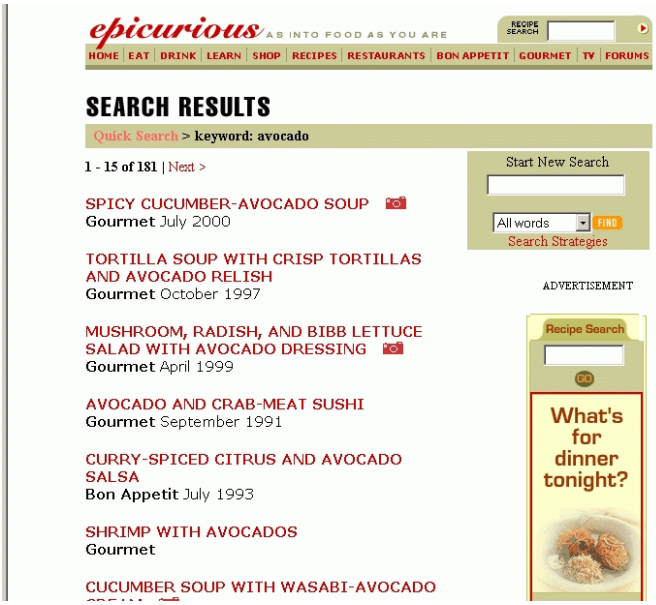

Figure 1: The Epicurious Basic search form and results listing.

gate through the collection, implicitly building up a query consisting of an AND across facets (see Figures 3 – 5). Selecting a category within a facet (e.g., Pasta within Main Ingredient) narrows the set of recipes shown. At each stage in the process, the user sees a preview of the number of recipes that are assigned to each of the remaining (not yet selected) facets, along with a list of the titles of matching recipes. For example, after the Pasta ingredient has been selected, the Cuisine facet of the Browse interface shows that 202 of the 694 recipes are Italian, 58 are Asian, 5 are Middle Eastern, and so on. The user can choose to switch the preview to the Preparation facet, which reveals that 234 recipes are Quick, 117 are Saute, 7 are Grill, and so on. If the user selects Saute, the results are refined to include only those recipes that have been assigned this preparation style. Now the interface shows the results according to a different facet. If the current view is Cuisine, the user sees that for Italian style recipes, the number has been reduced from 694 to 35. The user may also enter keywords within an entry form to search within the refined results, but the results of a search are shown as a list of titles without any associated preview information; thus running a query ends the ability to view results organized by metadata (see Figure 5). Finally, the user can see the history of selected facets and keywords in a "breadcrumb" shown at the top of the interface (e.g., Pasta > Saute > Italian).

The Browse interface has some interesting characteris-

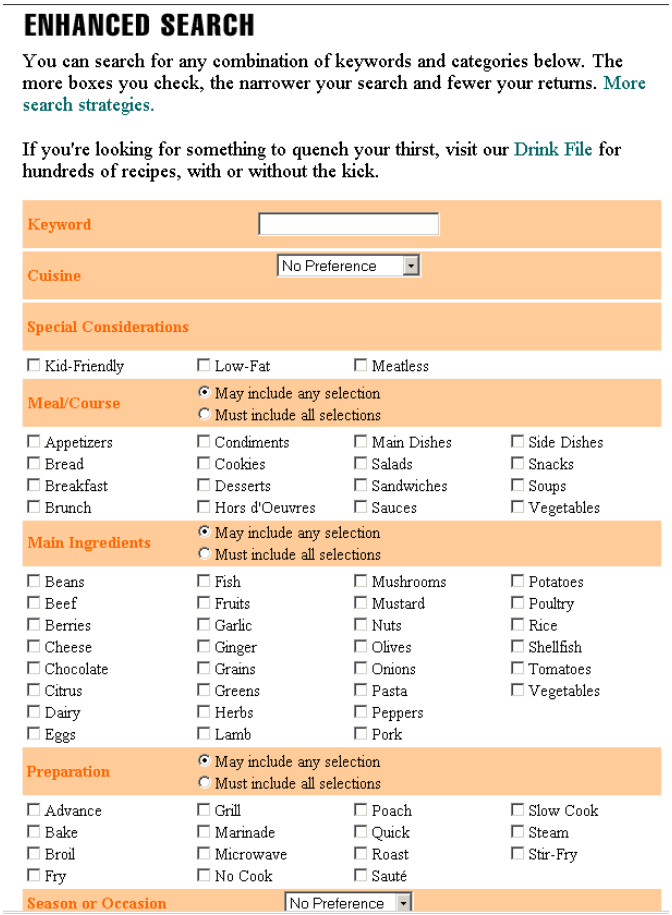

Figure 2: The Epicurious Enhanced search form.

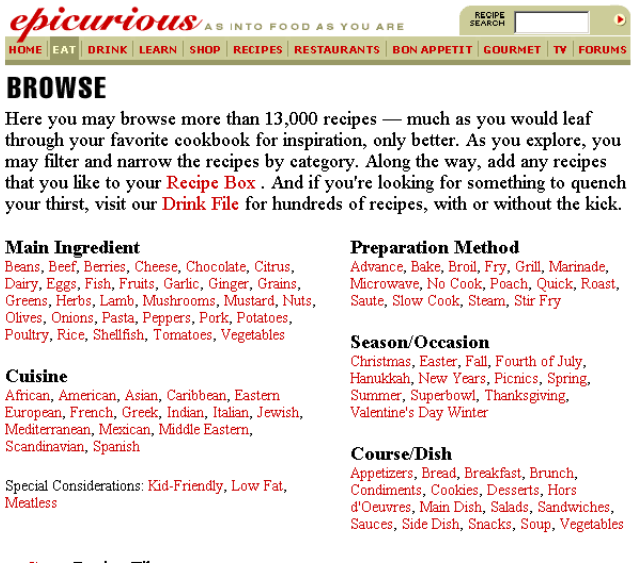

Figure 3: The opening of the Epicurious browse interface.

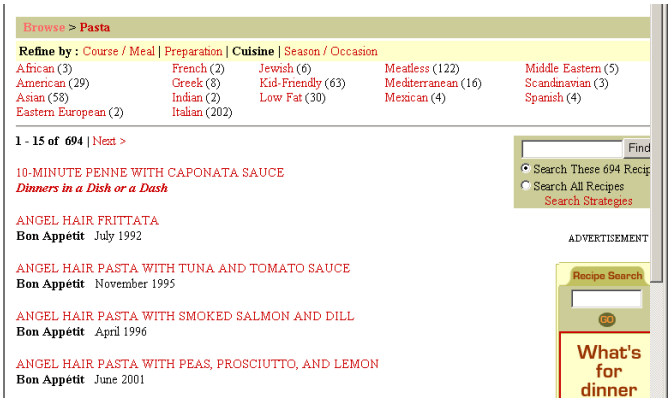

Figure 4: A view of the middlegame of the Epicurious Browse interface after one refinement operation.

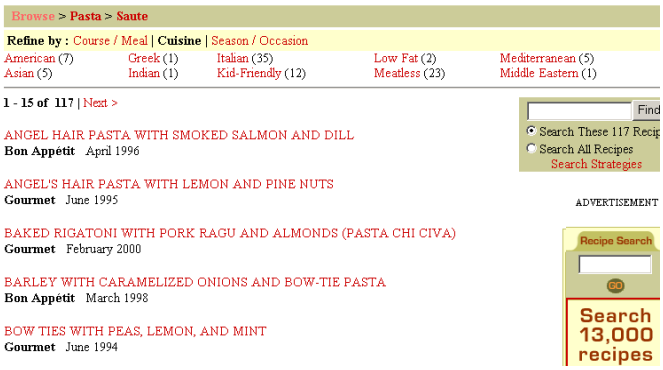

Figure 5: A view of the middlegame of the Epicurious Browse interface after another refinement operation.

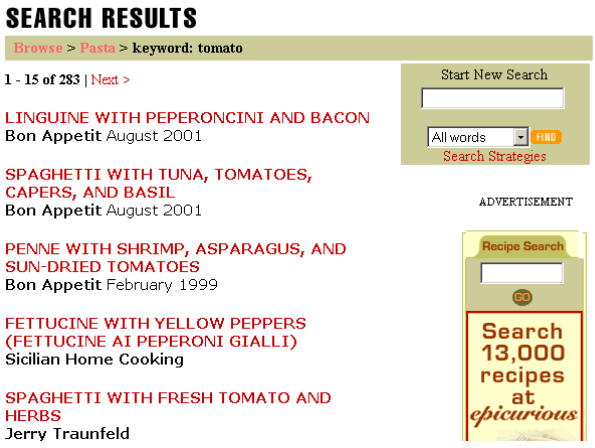

Figure 6: The results of searching within a Browse session.

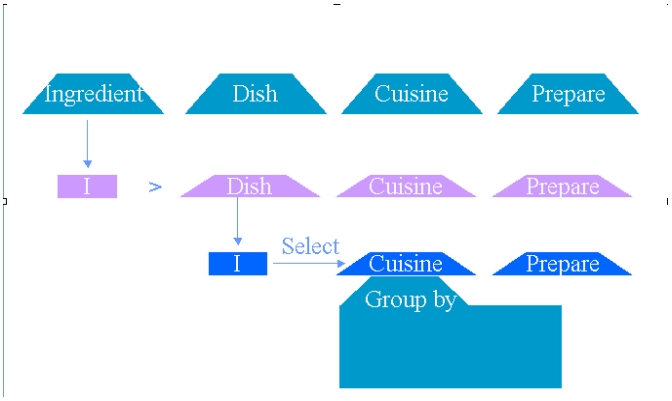

Figure 7: An illustration of the architecture of metadata use and query previews in the Epicurious Browse interface.

tics. Unlike when using standard directories such as Yahoo, the users can determine the order in which the facets are selected, and the previews are updated dynamically to reflect their choices. Each selection of a category from a facet narrows the result set, imposing an implicit AND across facets. To expand the query, the user in essence has to back up, either by selecting an earlier state from the history-based "breadcrumb" at the top of the display, or using the Back button. Users can also switch the view while holding the result set constant, thus viewing the same subcollection according to different facets. However, as Figure 7 shows, users cannot select multiple items from within one facet when using Browse mode (although they can do this in Enhanced Search).

This design choice makes the interface simpler than one that allows the users to select more than one item from each facet; however, it has the consequence of sometimes forcing users to do a keyword search within the Browse interface, since recipe search often benefits from specifying more than one ingredient. The search-within-results interface does not group the resulting recipes according to the metadata previews. We suspect that this is a flaw in the design of the system, and that users would prefer that the results would be grouped according to the metadata previews.

Finally, Epicurious simplifies the browsing-through- metadata problem by limiting each facet to only one level of hierarchy. The Dish facet, for example, has only one level, so that Desserts and Cookies are shown at the same level of description, and there is no Cakes category. This can be problematic, as past research shows that users are sensitive to this kind of blurring of levels [2], but the lack of hierarchy within facets greatly simplifies the interface design.

THE USABILITY STUDY

Study Goals

In this study we compared the three Epicurious interfaces. We wanted to know if the Browse interface is understood by users, if users like it, if they switch among the views, if they find the previews helpful. We also wanted to determine if users find different interfaces useful for different types of tasks. After some pilot testing, we hypothesized that the Enhanced Search interface would be more useful for highly constrained tasks, while the Browse interface would be more useful for open-ended, less constrained tasks. Finally, we wanted to develop hypotheses about how to best design search interfaces of this kind.

Method

A total of nine people participated in this pilot study, all of whom use the Internet on a regular basis to search for information, but otherwise vary widely in terms of their technical ability. Seven of the nine study participants were women, and all participants ranged in age from early 20s to late 40s. Motivated in part by Spool's exhortation to ensure that participants in a usability study actually care about what they are looking for [13], we made sure that all of the participants were interested in cooking and finding recipes. They were paid $12/hour for their participation.

Procedure

We asked participants to use the Epicurious interfaces before coming to the test session, thus allowing them to formulate opinions about the interfaces on their own, in their own environments and at their own pace. This helped to reduce the artificiality and pressure that may accompany testing an interface in a session with usability testers watching. We asked participants to complete three "treasure hunt" tasks (one task with each of the three interfaces) and then e-mail the recipes to the test facilitators. (A treasure hunt task is one in which the user is given the goal of finding a particular kind of recipe. For example, "Look for a recipe for a grilled chicken sandwich" or "Look for potato salad recipes.")

Next, participants came to our lab for the test session. Realistic study of search interfaces is difficult to conduct in a lab setting because of the problem of motivation [13]. If participants are not motivated to look for an item and are simply fulfilling a task requirement, then they are likely to be satisfied with the first set of results they find, rather than persevering to find items they truly want. To address this problem, which surfaced during our pilot user tests, we modified our testing protocol in two ways. First, we had participants generate their own search goals. The facilitator began by asking several general questions about the participant's cooking habits and use of recipe websites. Then the facilitator worked with participants to develop one or two scenarios describing situations in which they would be likely to search for new recipes, and to name three types of dishes they would like to prepare for the given scenarios. (To maintain a level of experimental control, the facilitator guided the participants to develop scenarios that would be appropriate for testing.) The second method for motivating participants was allowing them to save the recipes they found while searching. (We promised to create a formatted booklet of each participant's recipes.) Through these two simple techniques, the participants became highly motivated in a realistic setting.

Participants then used the Epicurious website to conduct searches for the dishes specified during the scenario-building step. Participants searched for each dish three times, once using each of the three different interfaces (Basic, Enhanced, Browse). The order of use of the search interfaces was randomized.

Whenever a participant received a set of intermediate results, they were asked if they felt as though they were getting closer, further away, or not changing position in relation to their search objectives. The facilitator recorded the number of recipes in each set of intermediate results.

After participants had completed a search, the facilitator asked questions about their level of satisfaction with the results, as well as what they liked, disliked, and wished to change about that search method.

In addition to the self-generated searches, participants also completed several structured tasks, using the two metadata-based interfaces (Enhanced and Browse). The structured tasks varied in terms of the level of constraint and the number of results required. An example of a high-constraint, single-result task is "Find a recipe for cooking a grilled eggplant sandwich with cheese." An example of a low-constraint, multiple-results task is "Your kitchen is being remodeled and for the next month, you will only have access to a grill for cooking. Find some recipes you can make this month." The participants were also given four hypothetical tasks, with different levels of constraint, and asked which interface they would prefer to use for searching.

Finally, the facilitator conducted a post-test interview, asking participants about the likelihood of using the Epicurious website and each of the three interfaces in the future. The interview also focused on specific system features, asking if users noticed the features and, if so, how helpful they felt them to be.

Measures

The independent variables for the study were as follows:

1. Epicurious Interface (Basic vs. Enhanced vs. Browse)
2. Task type (known-item search vs. browsing for inspiration
3. Degree of constraint of query
4. Number of results required (one vs. many)

The dependent variables are discussed in the results section.

Results and Discussion

Our results fall into three categories: participants' understanding of metadata, their use of metadata, and their preferences for different interfaces. We also use our study findings to profile the three interfaces we tested and to offer design recommendations for improving each one's effectiveness.

Understanding of Metadata

From the perspective of the search interface designer, metadata-based search is very different from keyword-based search. From the user's perspective, however, there is often no difference between searching with keywords and or with metadata. Some interfaces reveal the difference between keyword and metadata but most do not. Two examples show this confusion in the mind of our test participants. One participant attempted to use "kosher" as a keyword and received a recipe for a creamy chicken dish with bacon (which included "kosher salt" as an ingredient). The participant was understandably confused and subsequently distrustful of the keyword search. In this case, the participant really needed to use the word "kosher" as a recipe descriptor (or metadata).

Several participants were also confused by the keyword search box in the Enhanced interface. Typically participants would enter a query in it (e.g., sandwiches) in the search box, and then notice the same word as a checkbox option further down on the page. Participants were unsure if they should type the keyword and check the box or do both.

We think that the above two examples illustrate that, from the user's perspective, using keywords and metadata are very similar ways of constraining a search. For this reason, although our primary interest is in the use of metadata to constrain search, we have not separated metadata usage and keyword usage in the analysis below.

In addition, a few participants wanted more explanation of the metadata's meaning. For example, two participants were interested in the fat content and nutritional value of each recipe they viewed; they wanted to know exactly how checking "low-fat" would restrict their search results.

It is important to note that a metadata-based search interface can only be as good as the metadata behind it. We had several participants who were interested in facets that the metadata did not support. For such users, determining the appropriate metadata becomes a guessing game, and keyword search becomes the most effective strategy.

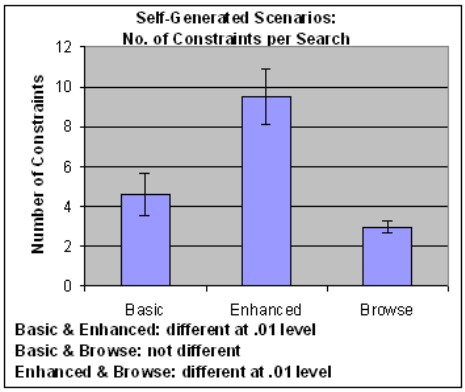

Figure 8: Number of constraints used on average, per search, for each interface type.

Use of Metadata

We asked participants to stop searching as soon as they felt satisfied that they had a reasonable set of results. While the three interfaces provided different ways of setting constraints, we wanted to learn if there was a difference in the total number of constraints set per search. (To determine the number of constraints for all three interfaces, we counted all keywords used and facets selected.)

Figure 8 indicates that in order to reach their goals, participants needed to apply more constraints to the Enhanced interface than the Browse interface. The high usage of constraints in Enhanced search was due to the fact that the search screen offered 9 facets, with a total of 68 options. Participants were initially excited about this degree of control, and began their searches by checking many boxes. 27% of the time participants constrained their search too much and received empty result sets. (By contrast, for Basic Search, participants received empty results only 12% of the time, and for Browse they received empty results only 4% of the time, from searching within results.)

Most likely in order to avoid empty results sets, for the Enhanced Search interface, participants soon learned to reduce the number of constraints per query. The average number of constraints for the first query the first time a participant used Enhanced Search was 5.13, as compared to the average number of constraints in the first query of subsequent searches (3.39).

Preferences for Search Interfaces

Overall Preferences: When asked which search method they preferred best, the participants were divided fairly

equally: 4 preferred Enhanced Search, while the other 5 preferred Browse. This can be explained by the fact that people indicated similar levels of satisfaction with the Browse and Enhanced interfaces (see Table 2)

Interface	Very Satis.	Satis.	Neither	Dissatis
Basic	32%	50%	9%	9%
Enhanced	43%	43%	4%	9%
Browse	35%	52%	4%	9%

Table 2: Satisfaction with Search Results

The satisfaction results for all three search types are high. We think that this is partly due to the fact that Epicurious is an award-winning site and since most of our participants liked cooking they were simply happy to have discovered Epicurious.

Task-Based Preferences: We were also interested in finding out if the participants found particular search interfaces more suitable for particular tasks. Towards the end of the testing phase, we presented participants with hypothetical search scenarios, and asked them which interface they would like to use. We had varied the hypothetical scenarios in terms of the degree of constraint. Table 3 shows preferences for various interfaces as a function of task constraint. Participants preferred the Enhanced interface for high constraint tasks, and the Browse interface for low constraint tasks. Only one participant preferred the Basic Search interface for the High Constraint task. This result indicates that participants understand the value of different interfaces for different tasks even after a short (one-hour) testing session. Also, participants were almost unanimous in this judgment. We think that this finding illustrates that different metadata-based search interfaces are needed for different search strategies.

	Basic	Enhanced	Browse
Low Constraint	0	1	8
High Constraint	1	8	2

Table 3: Preference for Interface as a Function of Task Type.

Reaction to System Features: We asked the participants about particular aspects of the interface-whether they noticed the features, and if so, the degree to which they found the features to be helpful while searching. Table 4 presents the summary of the participants' reaction to specific system features. It is interesting to note that only one participant found search options such as "Boolean queries" and "exact phrases" useful, and only two participants found useful the "may include" and "must include" options. The only feature that some participants felt interfered with the search process was the "set all criteria from one screen option" from the Enhanced search interface. This may be related to the frequency with which overly-constrained Enhanced searches returned empty results.

Interaction Profiles for the Three Interfaces

Basic Search Interface

The Epicurious Basic Search interface operated similarly to a typical search engine interface, and most participants found it easy to use. As the satisfaction data shows, participants liked this interface. However, they also recognized its limitations, since when we asked about task-based preferences for low and high constraint tasks, only one participant thought it was more suitable than either Enhanced Search or Browse.

The Basic Search interface of Epicurious was very limited since it did not offer any sort of a middlegame. Participants were only allowed to generate new queries and accept or reject the results set. The Back button was the only "refine search" option available.

We think that Basic Search can offer an easy entry point into metadata-based search interfaces. It is an interface that users are familiar with, and allows one to start the search without much though. However, the middlegame for Basic Search should offer options to make up for its impoverished opening.

Enhanced Search Interface

Using multiple constraints: The Epicurious Enhanced Search interface encourages the use of many constraints. Participants tried out an average of 9.5 constraints (including keywords) during each Enhanced Search session. With all of the metadata choices set before them in a menu, participants felt obliged to check two or more boxes. This constrained the search, so that it frequently resulted in empty results. The participants would return to the enhanced screen and gradually strip away constraints until an acceptable number of results were returned. Some participants found this very frustrating; others did not seem to mind. As the Table above shows, participants ended up with empty results 28% of the time.

Time: The average time per search was 168 seconds. This was longer than that for Browse and Basic search though the differences were not statistically significant. A possible reason for this result is that it took participants longer to carefully develop queries for the Enhanced interface. Average time per step was also the highest for Enhanced Search (68 seconds). However it required fewer steps than Browse.

Perception of Control: Our hypothesis is that Basic Search, Enhanced Search and Browse lead to different perceptions of control. A number of participants mentioned that one of the reasons they liked the Enhanced Search

	Did not notice	Helpful	Not Helpful	Interfered
Browse: Query previews	1	5	1	–
Enhanced & Browse: Having a complete list of ingredients	–	7	1	–
Browse: Search within results	–	6	1	–
Browse: Refine using hyperlinks	2	5	–	–
Enhanced: Set all criteria from one screen	–	3	2	2
Enhanced: "May include" & "must include" options	1	2	4	–
Basic/Enhanced: "All words," "any words," "exact phrase," and "Boolean" options	3	1	3	–

Table 4: Responses to Specific Features.

Interface was the degree of control it offered. It allowed them to choose the specific facets they were interested in (by clicking the checkboxes). Another reason for the feeling of control might be that changing a search or starting a new one were only one click away from the results page.

We think that the Enhanced Search interface can be an intuitive and powerful interface. However, it is important to incorporate some kind of a feedback mechanism into this interface so that participants are informed if their searches are going to lead to empty results. Part of the problem is that search forms (both Basic and Enhanced) do not offer any insight into the dataset being searched. As such the user has no way to learn about the dataset except by trial and error.

One possible remedy is for the interface to offer some constraint-specific feedback, perhaps in the form of query previews. Another possible design solution would be to offer fewer facets in the first search form, and use the remaining facets to organize the search results. Another problem with the Epicurious Enhanced Search interface was that it did not offer any way to do error correction. When participants choose too many or too few facets, they must go back to the start screen to make revisions. A strong middlegame could allow users to revise or to refine a query without having to start over entirely.

Browse Interface

The Epicurious Browse interface had a much stronger middle game than the Basic / Enhanced Search interface. Participants could start off by choosing one facet. Once they had made that initial choice, they had a whole range of options in front of them. They could search within results, backtrack a step, refine, and switch facets. Participants made use of all these options, though some were more easily understood than others.

Searching within Results: Participants chose that option frequently (31% of the time). This is the most easily understood option.

Backtracking: Participants could go back using either the Back button or the breadcrumb. In most cases participants used the Back button; the breadcrumb was clicked on only a few times. Generally the back button was used to go back a step or two and then the results were refined again using a different facet.

Switching Facets: Epicurious also offers the users to switch facets, in effect allowing them to view the same set of results from according to query previews for different facets. Initially participants did not notice or understand this feature until they had used the Browse interface a number of times. However, once they understood the feature they found it very useful.

The Epicurious Browse interface presents a very effective, well-designed example of a search middlegame, and we suspect that the Basic and Enhanced Search interface could be improved by incorporating some of the middlegame features from the Browse interface (such as organizing the output after searching within results).

However, there are a number of issues that this interface does not address. The most pressing issue is that of handling faceted metadata for which each facet is hierarchical.

CONCLUSIONS AND FUTURE WORK

The results of this preliminary study suggest that a metadata-based interface with dynamically generated metadata-based query previews can be an effective interface for navigating and searching large collections of loosely linked information. They also suggest that users recognize the utility of switching to the search interface that best supports the most appropriate search strategy. Finally, they suggest that, at least for a coherent collection like recipes, the use of metadata both to formulate the query and to navigate intermediate results is a promising avenue for the design of search user interfaces, and superior to standard keyword-based search paired with lists of results.

The study is too small to draw certain conclusions from; in particular, conclusions about preferences overall. Furthermore, the nature of the Epicurious site leaves us with many unanswered questions. The collection and metadata are of only moderate size, and the metadata

Figure 9: A view of a metadata-centric interface for architectural images.

is organized into flat facets. Furthermore, users cannot select multiple items from one facet while in Browse mode, and little support is given for expansion of the search. We are interested in learning how to make this kind of approach scale to very large collections (thus requiring hierarchical metadata facets), supporting selection of multiple items from multiple facets, supporting search expansion as well as refinement, and examining richer information needs.

For these reasons, we are in the midst of conducting another, larger study, using software developed ourselves, to allow us to directly contrast different interface features. (Figure 9 shows one of the interfaces being compared.) Although the software should 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, and the tasks are designed to be realistic from the point of view of an architect's information seeking needs. Each slide has approximately 10 items of metadata assigned it, and the metadata hierarchy contains several thousand terms and nine facets. These studies should further our understanding of how to improve web site search user interfaces.

Acknowledgements. This research is supported by an NSF CAREER grant grant, NSF9984741. We thank Ame Elliott for contributing to the earlier aspects of this work.

REFERENCES

1. Marcia J. Bates. The berry-picking search: User interface design. In Harold Thimbleby, editor, *User Interface Design*. Addison-Wesley, 1990.
2. Hsinchen Chen, Andrea L. Houston, Robin R. Sewell, and Bruce R. Schatz. Internet browsing and searching: User evaluations of category map and concept space techniques. *Journal of the American Society for Information Sciences (JASIS)*, 49(7), 1998.
3. Ed H. Chi, Peter Pirolli, Kim Chen, and James Pitkow. Using information scent to model user information needs and actions on the web. In *Proceedings of the ACM SIGCHI Conference on Human Factors in Computing Systems*, Seattle, WA, April 2001. ACM.
4. Steve B. Cousins and Ken Pier. A task-oriented interface to a digital library. In *Proceedings of the ACM SIGCHI Conference on Human Factors in Computing Systems, Conference Companion*, Vancouver, Canada, May 1996. ACM.
5. Fredric Gey, Hui-Min Chen, Barbara Norgard, Michael Buckland, Youngin Kim, Aitao Chen, Byron Lam, Jacek Purat, and Ray Larson. Advanced search technologies for unfamiliar metadata. In *Meta-Data '99 Third IEEE Meta-Data Conference*, Bethesda, MD, April 1999.
6. Marti A. Hearst. User interfaces and visualization. In Ricardo Baeza-Yates and Berthier Ribeiro-Neto, editors, *Modern Information Retrieval*, pages 257–323. Addison Wesley, ACM Computing Series, 1999.
7. Marti A. Hearst. Next generation web search: Setting our sites. *IEEE Data Engineering Bulletin*, 23(3), September 2000.
8. Gary Marchionini. *Information Seeking in Electronic Environments*. Cambridge University Press, 1995.
9. Vicki L. O'Day and Robin Jeffries. Orienteering in an information landscape: how information seekers get from here to there. In *Proceedings of the INTERCHI '93*, Amsterdam, April 1993. IOS Press.
10. 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.
11. Wanda Pratt, Marti Hearst, and Larry Fagan. A knowledge-based approach to organizing retrieved documents. In *Proceedings of 16th Annual Conference on Artificial Intelligence (AAAI 99)*, Orlando, FL, 1999.
12. Vividence Research. Tangled web 2001. http://www.vividence.com, June 2001.
13. Jared Spool. *Web Site Usability: A Designer's Guide*. Morgan Kaufmann, 1998.
14. Danny Sullivan. Npd search and portal site study. http://searchenginewatch.internet.com/reports/npd.html, July 6 2000. NPD's URL is http://www.npd.com.