User Interfaces for Search
by Marti A. Hearst

Chapter 2 of the second edition of *Modern Information Retrieval*
Renamed *Modern Information Retrieval: The Concepts and Technology behind Search Engines (2nd Edition).*

Contents

0.1 Introduction

Much of this book describes the algorithms behind search engines and information retrieval systems. By contrast, this chapter focuses on the human users of search systems, and the window through which search systems are seen: the *search user interface*. The role of the search user interface is to aid in the searcher's understanding and expression of their information needs, and to help users formulate their queries, select among available information sources, understand search results, and keep track of the progress of their search.

In the first edition of this book, very little was known about what makes for an effective search interface. In the intervening years, much has become understood about which ideas work from a usability perspective, and which do not. This chapter briefly summarizes the state of the art of search interface design, both in terms of developments in academic research as well as in deployment in commercial systems. The sections that follow discuss how people search, search interfaces today, visualization in search interfaces, and the design and evaluation of search user interfaces.

0.2 How People Search

Search tasks range from the relatively simple (e.g., looking up disputed facts or finding weather information) to the rich and complex (e.g., job seeking and planning vacations). Search interfaces should support a range of tasks, while taking into account how people think about searching for information. This section summarizes theoretical models about and empirical observations of the process of online information seeking.

Information Lookup versus Exploratory Search

User interaction with search interfaces differs depending on the type of task, the amount of time and effort available to invest in the process, and the domain expertise of the information seeker. The simple interaction dialogue used in Web search engines is most appropriate for finding answers to questions or to finding Web sites or other resources that act as search starting points. But, as Marchionini [89] notes, the "turn-taking" interface of Web search engines is inherently limited and is many cases is being supplanted by speciality search engines – such as for travel and health information – that offer richer interaction models.

Marchionini [89] makes a distinction between *information lookup* and *exploratory search*. Lookup tasks are akin to fact retrieval or question answering, and are satisfied by short, discrete pieces of information: numbers, dates, names, or names of files or Web sites. Standard Web search interactions (as well as standard database management system queries) can work well for these.

Marchionini divides the exploratory search class of information seeking tasks into *learning* and *investigating* tasks. Learning searches require more than single query-response pairs, and involve the searcher spending time scanning and reading multiple information items, and synthesizing content to form new understanding. Investigating refers to a longer-term process that involves "multiple iterations that take place over perhaps very long periods of time and may return results that are critically assessed before being integrated into personal and professional knowledge bases."[89] Investigative search may be done to support planning, discover gaps in knowledge, and to monitor an on-going topic. Some kinds of investigative search are concerned with finding all or a large proportion of the relevant information available (high recall), such as for litigation research or academic research.

Supporting this view, a study of in-depth and recurring complex information needs (with a focus on business intelligence) by O'Day & Jeffries [97] found that the information seeking process consisted of a series of interconnected but diverse searches. They also found that search results for a goal tended to trigger new goals, and hence search in new directions, but that the context of the problem and the previous searches was carried from one stage of search to the next. They also found that the main value of the search resided in the accumulated learning and acquisition of information that occurred during the search process, rather than in the final results set.

More broadly, information seeking can be seen as being part of a larger process, referred to in the literature as *sensemaking* [103, 115, 114]. Sensemaking is an iterative process of formulating a conceptual representation from a large collection of information. Russell et al. [115] observe that most of the effort in sensemaking goes towards the synthesis of a good representation, or way of thinking about, the problem at hand. They describe the process of formulating and crystallizing the important concepts for a given task. Search plays only one part in this process; some sensemaking activities interweave search throughout, while others consist of doing a batch of search followed by a batch of analysis and synthesis. Examples of deep analysis tasks that require sensemaking in addition to search include the legal discovery process, epidemiology (disease tracking), studying customer complaints to improve service, and obtaining business intelligence.

The Classic versus the Dynamic Model of Information Seeking

Researchers have developed numerous theoretical models of how people go about doing search tasks. The classic notion of the information seeking process model as described by Sutcliffe & Ennis [124] is formulated as a cycle consisting of four main activities:

- problem identification,
- articulation of information need(s),

- query formulation, and
- results evaluation.

The standard model of the information seeking process contains an underlying assumption that the user's information need is static and the information seeking process is one of successively refining a query until all and only those documents relevant to the original information need have been retrieved. More recent models emphasize the dynamic nature of the search process, noting that users learn as they search, and their information needs adjust as they see retrieval results and other document surrogates. This dynamic process is sometimes referred to as the *berry picking* model of search [11].

The rapid response times of today's Web search engines allow searchers to adopt a strategy of issuing a somewhat general query to "test the waters," looking at the results that come back, and based on the words shown, reformulating their query in some way, to try to get "closer" to the desired target [12, 56, 90]. For example, rather than specifying a complex query such as `connectors for a 1/2 inch gas grill hose to a 3/8 inch home outlet with flare`, which is likely to fail, a typical searcher will opt instead to query on a more general term such as `gas grill hose connectors` or even `gas hoses`, see what the search engine returns, and then either reformulate the query or visit appropriate Web sites and browse to find the desired product.

This kind of behavior is a commonly-observed strategy within the berry-picking approach, sometimes referred to as *orienteering* [97, 128]. The orienteering information seeker issues a quick, imprecise query in the hopes of getting into approximately the right part of the information space, and then does a series of local navigation operations to get closer to the information of interest [12, 90]. Usability studies and Web search logs suggest this approach is common. Searchers are highly likely to reformulate their queries; one analysis of search logs found the proportion of users who modified queries to be 52% [64].

Some information seeking models cast the process in terms of *strategies* and how choices for next steps are made. In some cases, these models are meant to reflect conscious planning behavior by expert searchers. In others, the models are meant to capture the less planned, potentially more reactive behavior of a typical information seeker. Bates [10] suggests that searcher's behavior can be characterized by search strategies which in turn are made up of sequences of search *tactics*. Bates [10] also discusses the importance of monitoring the progress of the current search and weighing the costs and benefits of continuing with the current strategy or trying something else. Russell et al. [115] also cast the activity of monitoring the progress of a search strategy relative to a goal or subgoal in terms of a *cost structure analysis*, or an analysis of diminishing returns. This cost structure analysis method was subsequently expanded by Pirolli & Card into *information foraging theory* [102, 104], a theoretical framework that uses an evolutionary biology stance to model and make predictions about peoples' strategies for navigating within information structures.

Navigation versus Search

Not all searches start with typing a keyword query in a search box. Many Web sites and some search engines allow users to peruse an information structure of some kind to select a starting point for search. The terms *navigation* and *browsing* are used somewhat interchangeably to refer to a strategy in which the searcher looks at an information structure, and moving among views of the available information, in a sequence of scan and select operations. Searchers often prefer browsing over keyword searching when the information structure (such as hyperlinks on a Web site) are well-matched to their information needs, A study by Hearst et al. [53] found that self-described searchers converted to browsers after a few iterations of using a well-designed faceted category system.

Browsing is often preferred because it is mentally less taxing to *recognize* a piece of information than it is to *recall* or remember it. But there are diminishing returns to scanning links if it takes too long to find the label of interest, or if the desired information is not visible. That is, browsing works well only so long as appropriate links are available, and have meaningful cues (sometimes referred to as *information scent* [104]) about the underlying information.

With an appropriate navigation structure, an interface that takes several clicks to lead a searcher to their goal is not necessarily a bad thing. Spool [121] claims that as a general rule, searchers do not so much object to following many links as they object to having to follow links that do not appear to bring them closer to their goal. So long as the searcher does not lose the "scent" on the path towards finding the desired information object, the interface performs well. Spool discusses an example of a user looking for a software driver for a particular laser printer. Say the user first clicks on *printers*, then *laser printers*, then the following sequence of links:

HP laser printers
HP laser printers model 9750
software for HP laser printers model 9750
software drivers for HP laser printers model 9750
software drivers for HP laser printers model 9750 for the Win98 operating system

This kind of interaction is acceptable when each refinement makes sense for the task at hand, and nowhere is it necessary to back up and try another choice: the trail never gets "cold." But if at some point midway through the clicks the searcher does not see a link leading closer to the goal, then the experience is frustrating and the interface fails from a usability perspective.

Observations of the Search Process

Numerous studies have been made of people engaged in the search process, and the results can help guide the design of search interfaces. One common observation, mentioned above, is that users often reformulate their queries with slight modifications, since this can be easier than trying to specify the query precisely in the first attempt. Another is that searchers often search for information that they have previously accessed [68, 91], and that users' search strategies differ when searching over previously seen materials [9, 68]. Researchers have developed search interfaces specifically designed to take into account the likelihood of re-accessing information [26, 32], supporting both query history and revisitation of previously accessed information items.

Studies have shown that it is difficult for people to determine whether or not a document is relevant to a topic [23, 116, 143], and the less someone knows about a topic, the poorer judge they are about if a search result is relevant to that topic [120, 133]. For Web search engines, searchers tend to look at only the top-ranked retrieved results and are biased towards thinking the top one or two results are better than those beneath it simply by virtue of their position in the rank ordering [44, 66].

Studies also show that people are poor at estimating how much of the relevant material in a collection they have found, and the less expert they are about a topic, the more likely they are to feel confident that all of the relevant information has been accessed [126]. Furthermore, searchers often terminate the search after a few results are found even though there might be much better information in the collection [124].

Some search usability studies have assessed the effects of knowledge of the search process itself, contrasting expert and novice searchers, although there is no consensus on the criteria for these classifications [8]. Studies have observed that experts use different strategies than novices [61, 83, 143], but perhaps more tellingly, other studies have found interaction effects between search knowledge and domain expertise [61, 65]. In one study, the overall characteristic that distinguished analysts who discovered high-value documents from those who did not was the persistence of the analysts; those who read more documents and spent more time on the task did better than those who did not [100]. In another study, expert searchers were more patient than novices, and this along with a positive attitude led to better search outcomes [126].

0.3 Search Interfaces Today

At the heart of the typical search session is a cycle of query specification, inspection of retrieval results, and query reformulation. As the process proceeds, the searcher learns about their topic, as well as about the available information sources.

This section describes several user interface components that have become standard in search interfaces and which exhibit high usability. As these components are described, the design characteristics that they support will be underscored. Ideally, these components are integrated together to support the different parts of the process, but it is useful to discuss each separately.

Getting Started

How does an information seeking session begin? For many today, the Web has largely supplanted traditional physical information sources such as printed telephone books and encyclopedias. For a user of online information systems, the most common way to start a search session is to access a Web browser and use a Web search engine.

Another method for getting started, which used to be heavily used but has become less popular as search engines became faster and more accurate, is to select a Web site from a personal collection of already-visited sites, typically stored in a browser's bookmark saving facility [128]. Alternative bookmark systems, in which links to a user's preferred sites are stored on a Web site (and so accessible from any connected computer) and in which it is possible to see which Web sites other people have chosen to save, have become popular among a smaller segment of users. These sites (delicious.com and furl.com are examples) allow users to assign labels, or *tags* to the content, allowing for searching or browsing by topic as well as text search over the titles of the Web sites.

Web directories also used to be a common starting point. Yahoo.com's directory was the most popular starting point for navigating the Web in its early days, but directories have been largely replaced by search engines, both because the Web became too large to catalog manually, and because of the increased accuracy of Web search [101]. Still, some academics worry that users of search should be more cognizant of the sources that their information comes from, and feel that this information should be more prominently indicated in search results listings [111].

Query Specification

Once a search starting point has been selected, the primary methods for a searcher to express their information need are either entering words into a search entry form or selecting links from a directory or other information organization display. For Web search engines, the query is specified in textual form. Today this is usually done by typing text on a keyboard, but in future, query specification via spoken commands will most likely become increasingly common, using mobile devices as the input medium.

Typically in Web queries today, the text is very short, consisting of one to three words [64, 63]. Multiword queries are often meant to be construed as a phrase, but can also consist of multiple topics. Short queries reflect the standard usage scenario in which the user "tests the waters" to see what the search engine returns in response to their short query. If the results do not look relevant, then the user reformulates their query. If the results are promising, then the user navigates to the most relevant-looking Web site and pursues more fine-tuned queries on that site [12, 37, 56, 90]. This search behavior in which a general query is used to find a promising part of the information space, and then follow hyperlinks within relevant Web sites, is a demonstration of the orienteering strategy of Web search [97, 128]. There is evidence that in many cases searchers would prefer to state their information need in more detail, but past experience with search engines taught them that this method does not work well, and that keyword querying combined with orienteering works better [16, 106].

Commercial text search systems before the Web regularly supported Boolean operators and command-based syntax, and in fact did not generally support keyword queries well. However, Boolean operators and command-line syntax have been shown time and again to be difficult for most users to understand, and those who try to use them often do so incorrectly [30, 45, 47, 56, 58].

Although most Web search engines support some form of Boolean syntax, a recent Web query log study, conducted over 1.5M queries, found that 2.1% contained Boolean operators and 7.6% contained other query syntax, primarily double-quotation marks for phrases [63]. Another study examined interaction logs of nearly 600,000 users issuing millions of queries over a period of 13 weeks in 2006. They found that 1.1% of the queries contained at least one of the four main Web operators (double quotes, `+`, `-`, and `site:`) and only 8.7% of the users used an operator at any time [142].

Web ranking has gone through three major phrases. In the first phase, from approximately 1994–2000, statistical ranking was used by most search engines, but information about query term proximity within the page was not, nor was information about relative importance of Web pages. The Web was much smaller then, and so complex queries were less likely to have relevant information sources available to answer them. Pages would be retrieved that omitted a key word from the query, and many users found this behavior baffling. (AltaVista introduced the *mandatory* operator, indicated with a plus sign, that allowed users to indicate that a word following the plus was required in the query, but only the most sophisticated users make use of query operators.)

Around 1997, Google moved to conjunctive queries only, meaning that all terms in the query had to be found in a Web page in order for it to be retrieved. They also added term proximity information and page importance scoring (see PageRank in Chapter **??**), and this greatly improved the relevance for many queries, especially navigational queries; a query on "Toyota" retrieved the home page for the Toyota corporation as opposed to whichever page used the word "toyota" most often. The other Web search engines followed, and conjunctive ranking became the norm.

As the amount of information available on the Web increased, sophisticated searchers found that longer queries posed as phrases often produced highly relevant results. In the past, if a searcher had a complex information need and attempted to express it fully to a Web search engine, the attempt would fail. For instance, if a searcher wanted to known "where can I find good replacement hubcaps for my 1985 corolla?", a query written in this form would return empty results due to the conjunctive constraint. Today, Web search engines have become more sophisticated about dropping terms that would result in empty hits, while matching the important terms, ranking the hits higher that have these terms in close proximity to one another, and using the other aspects of ranking that have been found to work well for Web search.

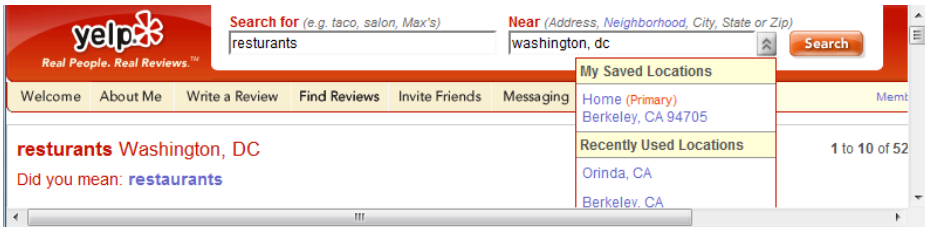


Figure 0.1 Query form, from yelp.com, illustrating support to facilitate structured queries, and stored information about past queries.

Query Specification Interfaces

The standard interface for a textual query is a search box entry form, in which the user types a query, activated by hitting the return key on the keyboard or selecting a button associated with the form. Studies suggest a relationship between query length and the width of the entry form; results find that either small forms discourage long queries or wide forms encourage longer queries [15, 41].

Some entry forms are divided into multiple components, allowing for a more general free text query followed by a form that filters the query in some way. For instance, at yelp.com, the user enters a general query into the first entry form and refines the search by location in the second form (see Figure 0.1). Forms allow for selecting information that has been used in the past; sometimes this information is structured and allows for setting parameters to be used in future. For instance, the yelp.com form shows the user's home location (if it has been indicated in the past) along with recently specified locations and the option to add additional locations.

An increasingly common strategy within the search form is to show hints about what kind of information should be entered into each form via greyed-out text. For instance, in zvents.com search (see Figure 0.2), the first box is labeled "what are you looking for?" while the second box is labeled "when (tonight, this weekend, ...)". When the user places the cursor into the entry form, the grey text disappears, and the user can type in their query terms.

This example also illustrates specialized input types that some search engines are supporting today. For instance, the zvents.com site recognizes that words like "tomorrow" are time-sensitive, and interprets them in the expected manner. It also allows flexibility in the syntax of more formal specification of dates. So searching for "comedy" on "wed" automatically computes the date for the nearest future Wednesday. This is an example of designing the interface to reflect how people think, rather than making how the user thinks conform to the brittle, literal expectations of typical programs. (This approach to "loose" query specification works better for "casual" interfaces in which getting the date right

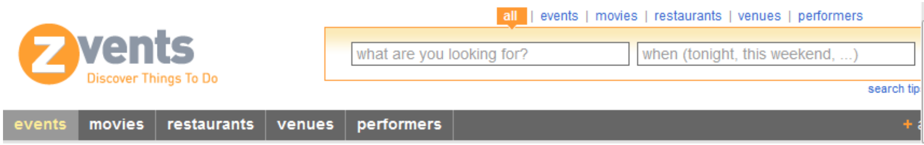


Figure 0.2 Query form, from zvents.com, illustrating greyed-out text that provides hints about what kind of information to type, directly in the form.

Figure 0.3 Dynamic query suggestions, from Netflix.com.

is not critical to the use of the system; casual date specification when filling out a tax form is not acceptable, as the cost of error is too high.)

An innovation that has greatly improved query specification is the inclusion of a dynamically generated list of query suggestions, shown in real time as the user types the query [141]. This is referred to variously as *auto-complete*, *auto-suggest*, and *dynamic query suggestions*. A large log study found that users clicked on dynamic suggestions in the Yahoo Search Assist tool about one third of the time they were presented [6].

Often the suggestions shown are those whose prefix matches the characters typed so far, but in some cases, suggestions are shown that only have interior letters matching. If the user types a multiple word query, suggestions may be shown that are synonyms of what has been typed so far, but which do not contain lexical matches. Netflix.com both describes what is wanted in grey and then shows hits via a dropdown (see Figure 0.3).

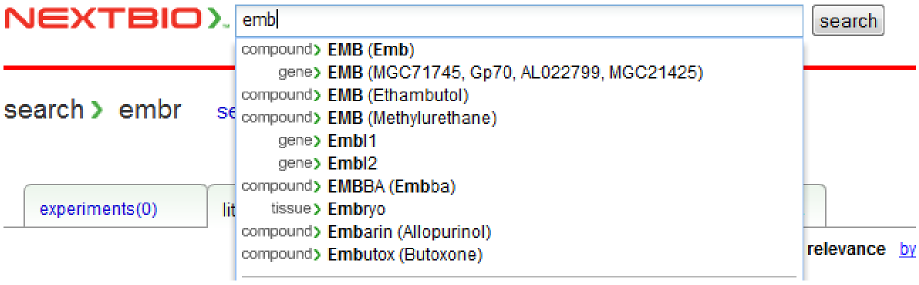


Figure 0.4 Dynamic query suggestions, grouped by type, from NextBio.com.

In dynamic query suggestion interfaces, the display of the matches varies. Some interfaces color the suggestions according to category information. In most cases the user must move the mouse down to the desired suggestion in order to select it, at which point the suggestion is used to fill the query box. In some cases the query is then run immediately; in others, the user must hit the Return key or click the Search button.

The suggestions can be derived from several sources. In some cases the list is taken from the user's own query history, in other cases, it is based on popular queries issued by other users. The list can be derived from a set of metadata that a Web site's designer considers important, such as a list of known diseases or gene names for a search over pharmacological literature (see Figure 0.4), a list of product names when searching within an e-commerce site, or a list of known film names when searching a movie site. The suggestions can also be derived from all of the text contained within a Web site.

Another form of query specification consists of choosing from a display of information, typically in the form of hyperlinks or saved bookmarks. In some cases, the action of selecting a link produces more links for further navigation, in addition to results listings. This kind of query specification is discussed in more detail in the section on organizing search results below.

Retrieval Results Display

When displaying search results, either the documents must be shown in full or else the searcher must be presented with some kind of representation of the content of those documents. The document *surrogate* refers to the information that summarizes the document, and is a key part of the success of the search interface. The design of document surrogates and retrieval results displays is an active area of research and experimentation.

The quality of the surrogate can greatly effect the perceived relevance of the

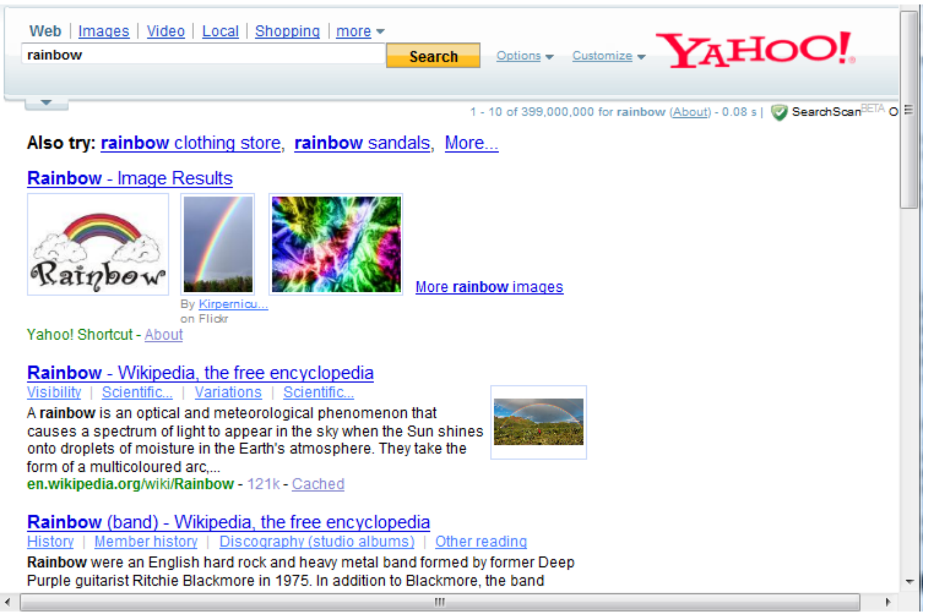


Figure 0.5 Yahoo search results page for a query on “rainbow.” The results display includes, from top to bottom, query refinement suggestions, links to images of rainbows, a link to an encyclopedia article on rainbows, with an illustrating image, and another encyclopedia article on a rock band called Rainbow.

search results listing. In Web search, the page title is usually shown prominently, along with the URL and sometimes other metadata. In search over information collections, metadata such as date published and document author are often displayed (but this kind of metadata is less applicable to Web pages). The text *summary* (also called the *abstract, extract, excerpt*, and *snippet*) containing text extracted from the document is also critical for assessment of retrieval results.

One study evaluated which attributes of a search result display led to more clicks, finding significant positive efforts for – among other things – longer versus shorter text summaries, titles containing the query words, combinations of title, summary and URL together containing the query as a phrase match, shorter URLs, and URLs that contain query terms in the domain name [21].

Currently the standard results display is a vertical list of textual summaries, and is sometimes referred to as the *SERP* (Search Engine Results Page). In some cases the summaries are excerpts drawn from the full text that contain the query terms. In other cases, specialized kinds of metadata are shown in additional to standard textual results, using a technique known as *blended* results (also called *universal search*). For example, a query on a term like “rainbow” may return

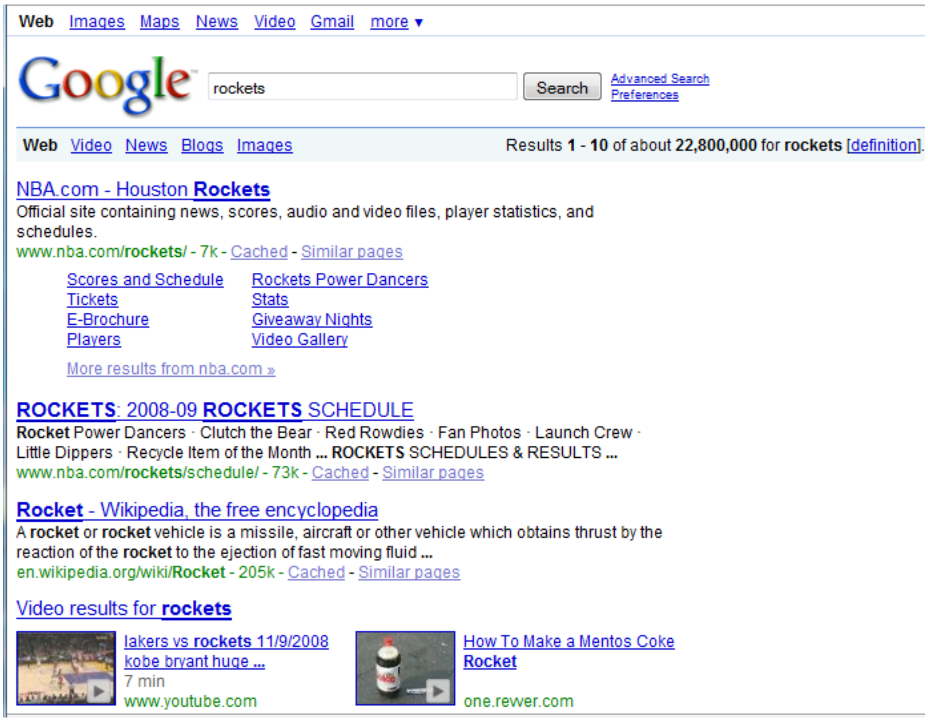


Figure 0.6 Google search results page for a query on "rockets." Several different kinds of information are shown for different senses of the word. The first hit is a link to the home page of a basketball team, along with "deep links" to popular pages within the site. This is followed by another link pertaining to the sports team, a link to an encyclopedia page about the projectile sense of the word, and links to videos of the sports team and a rocket how-to video.

sample images as one line in the results listing (see Figure 0.5), or a query on the name of a sports team might retrieve the latest game scores and a link to buy tickets or see the broadcast time of upcoming games (see Figure 0.6). Nielsen [95] notes that in some cases the information need is satisfied directly in the search results listing, thus making the search engine into an "answer engine."

Studies of ranking find that proximity information can be quite effective at improving precision of searches [22, 50, 127]. Usability studies suggest that showing the query terms in the context in which they appear in the document improves the user's ability to gauge the relevance of the results [130, 139]. This is sometimes referred to as *KWIC* (keywords in context), *query-biased summaries*, *query-oriented summaries*, or *user-directed summaries*.

The visual effect of query term *highlighting* has been known for decades to improve usability of search results listings [82, 84, 88, 90]. Highlighting refers to visually contrasting one part of the text with the rest, and can be achieved with boldfacing, changing the color of the text or its background, changing the size of the text, among other methods. Highlighting can usefully be shown both in document surrogates in the retrieval results and in the retrieved documents themselves. Some interfaces use visualizations to give an overview of the highlighted portions within a document [13, 19, 42, 59].

Determining which text to place in the summary, and how much text to show, is a challenging problem. Often the most relevant articles are those that contain all the query terms in close proximity to one another, but for less conforming results there is a tradeoff between showing contiguous sentences, to aid in coherence in the result, and showing sentences that contain the query terms. Some results suggest that it is better to show full sentences rather than cut them off [7, 113, 140], but on the other hand, very long sentences are usually not desirable in the results listing.

There is also evidence that the kind of information to display in search results summaries should vary according to the intent of the query and the goals of the search session. Several studies have shown that longer results are deemed better than shorter ones for certain types of information need [71, 87, 99]. When the searcher has decided to go directly to the home page of a known Web site, then an abbreviated listing is preferable to long, detailed information. In general, users engaged in *known-item searching* tend to prefer short surrogates that clearly indicate the desired information. Home page search is essentially address search; the user knows the name and wants to find the Web address (URL). Similarly, requests for factual information that can be stated briefly can be satisfied with a concise results display. By contrast, if the user has a complex information need, more in-depth document summaries can result in a better search experience. This is true for rich tasks such as advice seeking or for getting familiar with a subject.

Other kinds of document information can be usefully shown in the search results page. Figures 0.6 and 0.8 show the use of what are known as *sitelinks* or *deep links*, in which popular pages within a Web site are shown beneath that site's home page. As another example, a study of bioscience literature search found that most participants strongly favored showing figures extracted from journal articles alongside the search results [52]. In Figure 0.7, a screenshot from the BioText system shows this idea and also illustrates the use of highlighted and boldfaced query terms and a mechanism for the user to see more or less context around the query term hits.

Query Reformulation

After a query is specified and results have been produced, a number of tools exist to help the user reformulate their query, or take their information seeking process in a new direction. Analysis of search engine logs shows that reformulation is a common activity; one study found that more than 50% of searchers modified at

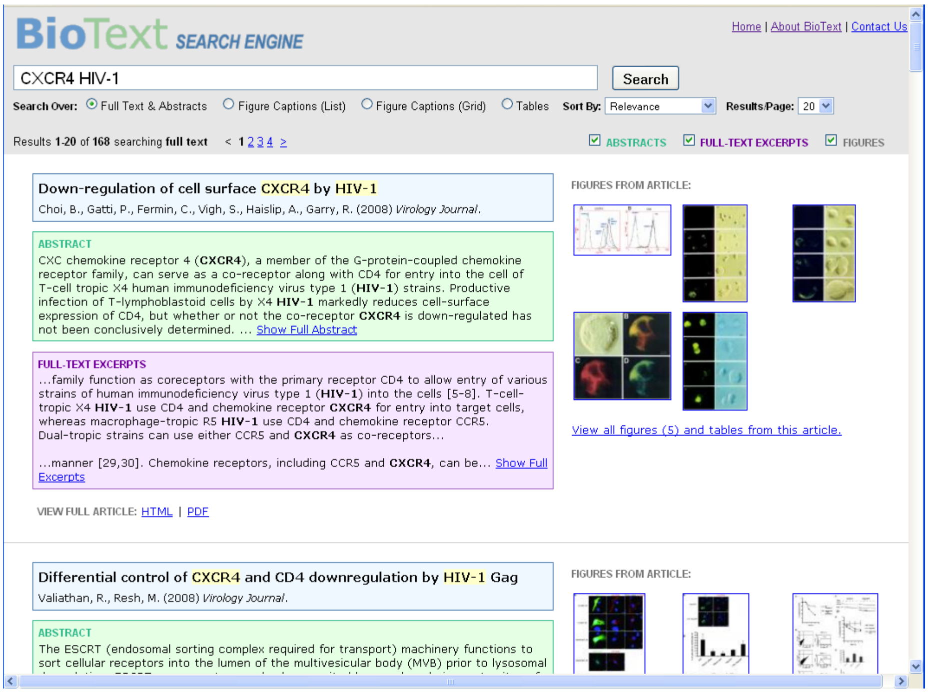


Figure 0.7 Search results from the BioText system, in which rich document surrogate information is shown, including figures extracted from the articles, query term highlighting and boldfacing, and an option to expand or shorten extracted document summaries. From http://biosearch.berkeley.edu.

least one query during a session, with nearly a third of these involving 3 or more queries [64].

One of the most important query reformulation techniques consists of showing terms related to the query or to the documents retrieved in response to the query. A special case of this is spelling corrections or suggestions; by one estimate, typographical errors occur in about 10-15% of queries [25]. In pre-Web search, spelling suggestions were primarily dictionary based [79]. In Web search, query logs have been used to develop highly accurate algorithms for detecting and correcting spelling and typographical errors [25, 85]. In the search interface, usually only one suggested alternative is shown; clicking on that alternative re-executes the query. In years back, the search results were shown using the purportedly incorrect spelling; today some search engines take the liberty of interweaving hits with the proposed correction among the hits with the original spelling, or showing hits with the correction separated from those with the original spelling.

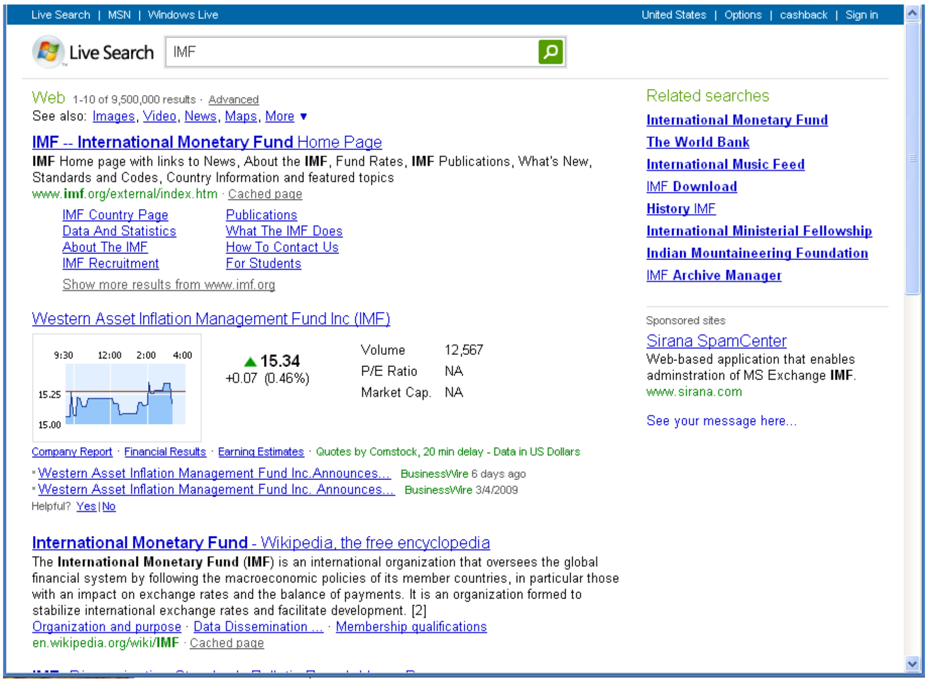


Figure 0.8 Microsoft Live's search results page on the query "IMF," including related terms suggestions (on the righthand side), links to alternative search "verticals" (images, videos, news, etc.), sitelinks under the first hit, financial statistics, and an encyclopedia article.

In addition to spelling suggestions, search interfaces are increasingly employing related term suggestions, a technique often referred to as *term expansion*. Log studies suggest that term suggestions, if presented well, are a somewhat heavily-used feature in Web search. One log study found that about 8% of queries were generated from term suggestions [63] (but it was not shown what percent of queries were shown such suggestions), while another found about 6% of users who were exposed to term suggestions chose to click on them [6].

Earlier attempts to use term expansions or suggestions in search interfaces showed more than a dozen thesaurus terms, and often forced the user to select among these before showing any search results [17, 29]. More recent studies suggest that a smaller number of suggestions, requiring only one click, to select or else grouping related terms together for selection with one click, is a preferable approach [4, 31, 138]. Figures 0.5 and 0.8 show examples of one-click term expansion suggestions.

Some query term suggestions are based on the entire search session of the particular user, while others are based on behavior of other users who have issued the same or similar queries in the past. One strategy is to show similar queries by other users; another is to extract terms from documents that have been clicked on in the past by searchers who issued the same query. In some cases, the same algorithms are used as for in the real-time query auto-suggest described above.

Relevance feedback is another method whose goal is to aid in query reformulation, as discussed in detail in Chapter **??**. The main idea is to have the user indicate which documents are relevant (and optionally, non-relevant) to their query. In some variations, users also indicate which terms extracted from those documents are relevant [76]. The system then computes a new query from this information, using any of a variety of algorithms, and shows a new retrieval set [116].

Relevance feedback has been shown in non-interactive or artificial settings to be able to greatly improve rank ordering [1, 74]. Nonetheless, this method has not been found to be successful from a usability perspective and does not appear in standard interfaces today [23, 116]. This stems from several factors: people are not particularly good at judging document relevance, especially for topics with which they are unfamiliar [133, 143], and the beneficial behavior of relevance feedback is inconsistent [24, 92], which is problematic from a usability perspective. Additionally, much of relevance feedback's benefits come from tasks in which a large number of relevant documents are desired, and this is less often the case in Web search; in fact there is some evidence that the benefits of relevance feedback disappear when compared to searching over the entire Web [129] (most benefits of relevance feedback have been seen when applied to smaller collections).

A variation on relevance feedback which has been shown to have positive responses in certain circumstances is an automatically computed "related documents" feature. In the biomedical literature search system PubMed, a handful of related articles are shown alongside a given journal article, and this features is well-received among bioscientists. One study showed that of the sessions in which related articles where shown, 18.5% included a click on a suggested article [86].

Organizing Search Results

Searchers often express a desire for a user interface that organizes search results into meaningful groups to help understand the results and decide what to do next. A longitudinal study in which users were provided with the ability to group search results found that users changed their search habits in response to having the grouping mechanism available [72]. Currently two methods for grouping search results are popular: *category systems*, especially *faceted categories*, and *clustering*. Both are described in more detail in this section, and their usability is compared.

A category system is a set of meaningful labels organized in such a way as to reflect the concepts relevant to a domain. They are usually created manually, although assignment of documents to categories can be automated to a certain degree of accuracy [117] (see Chapter **??**). Good category systems have the characteristics of being coherent and (relatively) complete, and their structure is predictable and consistent across search results for an information collection.

The most commonly used category structures, both for organizing search results, and for presenting information collection structure generally, are *flat, hierarchical,* and *faceted* categories. Flat categories are simply lists of topics or subjects. They can be used for grouping, filtering (narrowing), and sorting sets of documents in seach interfaces. Most Web sites organize their information into general categories, where selecting that category narrows the set of information shown accordingly. Some experimental Web search engines automatically organize results into flat categories; studies using this kind of design have received positive user responses [33, 80]. It can difficult to find the right subset of categories to use for the vast content of the Web; rather, category systems seem to work better for more focused information collections.

Hierarchical organization online is most familiarly seen in desktop file system browsers. In the early days of the Web, hierarchical directory systems such as Yahoo's were used to organize popular sites into a browsable structure. It is, however, difficult to maintain a strict hierarchy when the information collection is large, and cross-links were eventually allowed in these structures. Furthermore, the size of the Web overtook what was manageably browsed in such a system, and search engine use has largely replaced browsing of directory structures.

One context in which hierarchy can be effective is in a table-of-contents style presentation of results over a book or other small collection. The Superbook system [35, 36, 82] was an early search interface, which made use of the structure of a large document to display query term hits in context. After the user specified a query on the book, the search results were shown in the context of the table-of-contents hierarchy (see Figure 0.9). When the user selected a page from the table-of-contents view, the page itself was displayed in the righthand side and the query terms within the page were highlighted in reverse video. More recently, some research projects have applied this idea to organizing Web intranets [20, 94, 146].

There is increasing recognition that strictly hierarchical organization of categories is not ideal for navigating information structures. Hierarchy forces the user to begin with a particular category even though most information items can be classified by a multitude of attributes. Hierarchy also usually assumes an information item is located in one place in the category system, while computer interfaces can be more flexible than library bookshelves.

An alternative representation, known as faceted metadata, has become the primary way to organize Web site content and search results. Faceted metadata is intermediate in complexity between flat categories and full knowledge representation, and has been shown, when properly designed, to be understood by users and preferred over other organizations. Rather than creating one large category hierarchy, faceted metadata consists of a set of categories (flat or hierarchical), each of which corresponds to a different facet (dimension or feature type) relevant to the collection to be navigated. After the facets are designed,

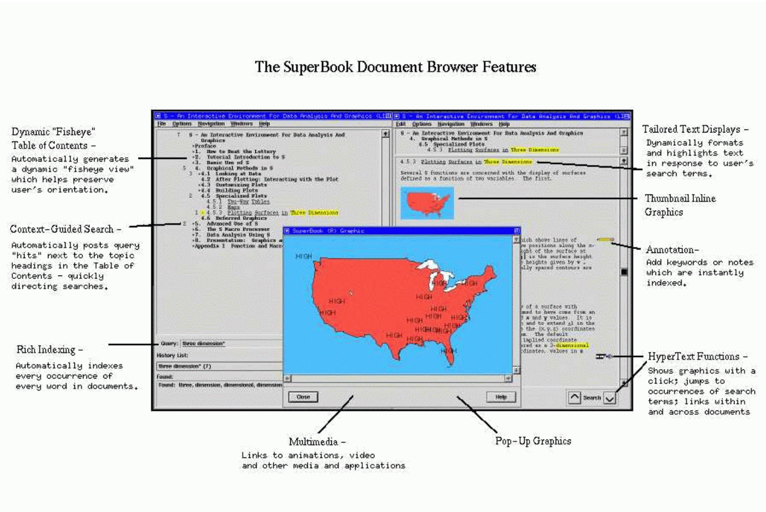


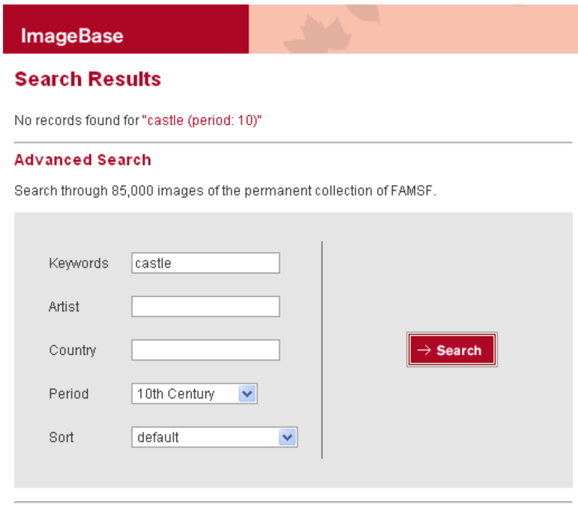
Figure 0.9 The SuperBook interface for showing retrieval results in context, using the table of contents from a large manual, from [82].

each item in the collection is assigned any number of labels from the facets.

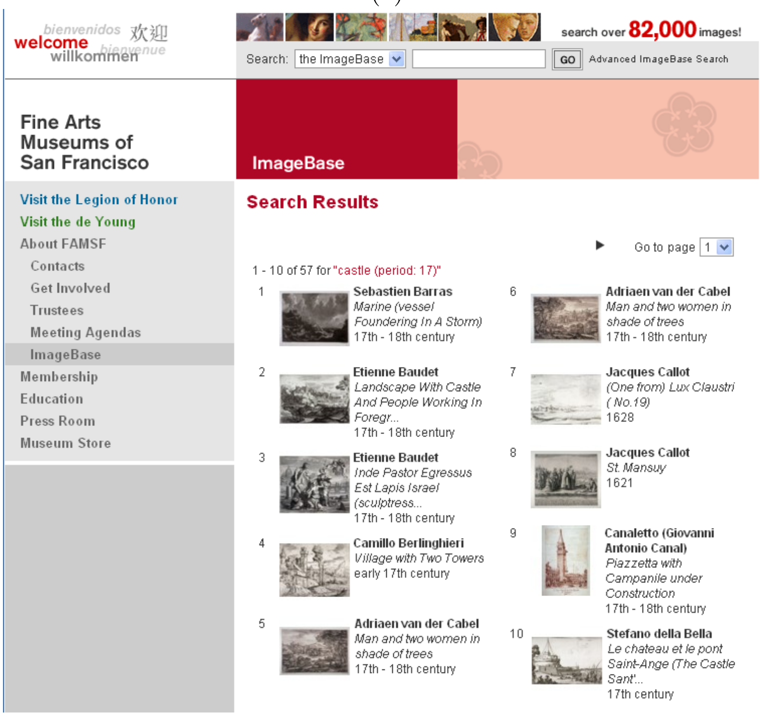
An interface using hierarchical faceted navigation simultaneously shows previews of where to go next and how to return to previous states in the exploration, while seamlessly integrating free text search within the category structure. Thus mental work is reduced by promoting recognition over recall and suggesting logical but perhaps unexpected alternatives at every turn, while at the same time avoiding empty results sets. This approach provides organizing context for results and for subsequent queries, which can act as important scaffolding for exploration and discovery.

Figure 0.10(a) and (b) show the search results for a typical image search interface in a hypothetical search session. The user entered the keyword "castle" on an advanced search form, and tried selecting "10th Century" as well, but received an error message indicating no records found. After trial and error, the user did find hits for "17th Century," and the results are shown in a fixed order that does not allow for organization and exploration.

Figure 0.11 shows the same kind of information made more accessible via faceted navigation using the Flamenco system [53, 148]. The user initially entered a keyword query of "castle," and then was shown 229 image search results, along with the structure on the lefthand side which allows the user to organize the results by Media type, Location, Objects (visible within the image), Building types (of which "castle" is one), Author, and so on. Empty results listings aren't a problem as long as the user selects hyperlinks, and query previews [105]



(a)



(b)

Figure 0.10 Typical image search interface, on a collection of images from the Fine Arts Museum of San Francisco. (a) Error message after the user tries a query on two fields in a typical "advanced search" form. A common problem with such forms is the production of empty results sets. (b) Standard search results listings after using an advanced search form to select keyword "castle" and time period "17th Century."

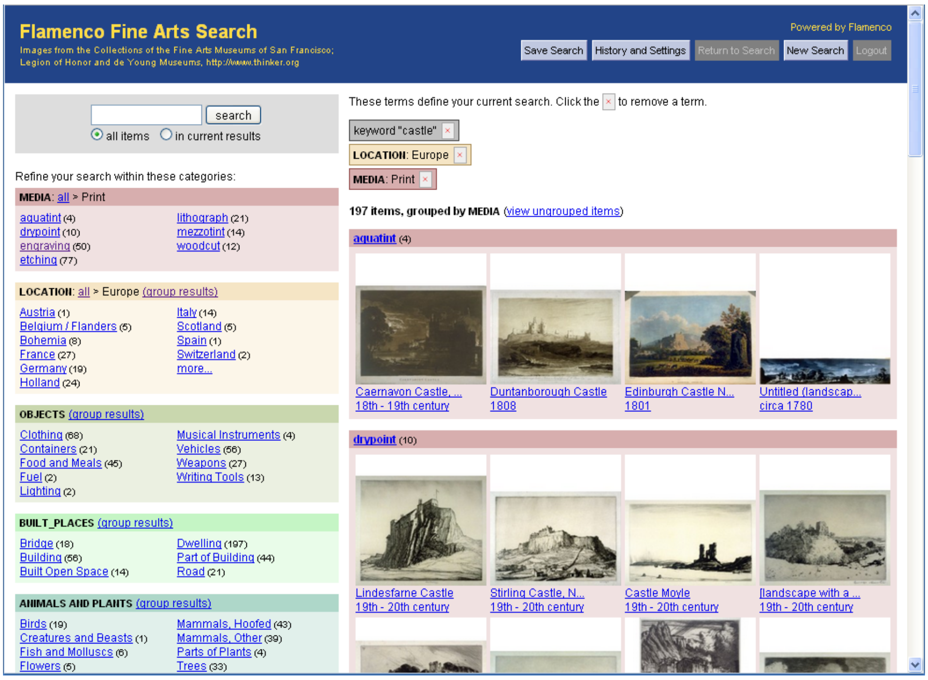


Figure 0.11 Faceted navigation in the Flamenco interface, applied to a subset of the San Francisco Fine Arts Museum collection.

show how many hits to expect to see after a link is clicked on. In this case the user first selected *Media* > *Prints*, then organized by *Location* > *Europe*, and then chose to re-organize by the subhierarchy beneath *Print*. The hierarchical faceted metadata at the left shows which European countries are represented, and how many times, by the remaining 197 images. Selecting an image shows the metadata associated with it, along with links to related concepts, such as "ruins" and "hill." Figure 0.12 shows a similar idea applied to a a digital library catalog, and Figure 0.13 shows it applied to a yellow pages and reviewing Web site.

Usability studies find that users like and are successful using faceted navigation, if the interface is designed properly, and that faceted interfaces are overwhelmingly preferred for collection search and browsing as an alternative to the standard keyword-and-results listing interface [53, 148].

Clustering refers to the grouping of items according to some measure of similarity (see Chapter **??**). In document clustering, similarity is typically computed using associations and commonalities among features, where features are typically words and phrases [27]. The greatest advantage of clustering is that it

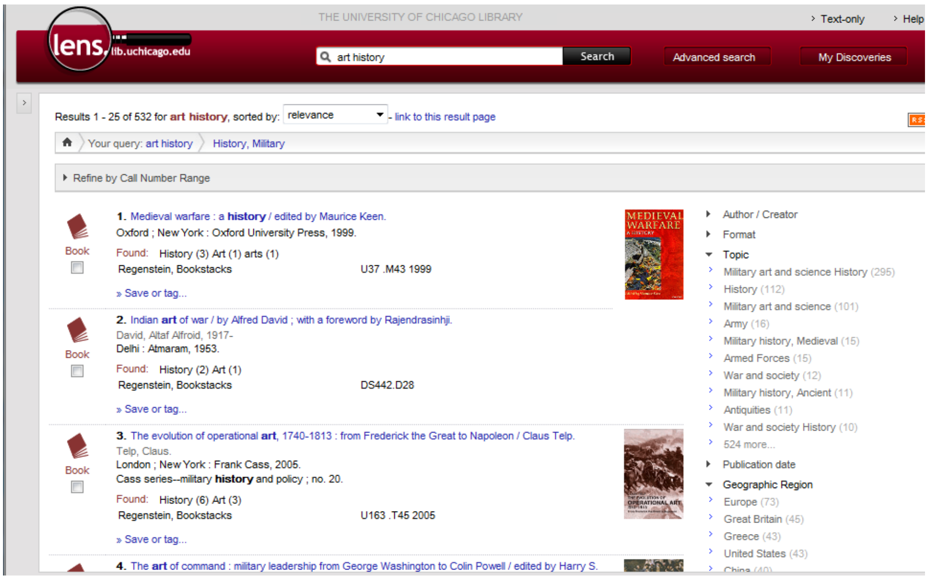


Figure 0.12 University of Chicago digital library faceted navigation, by AquaBrowser.

is fully automatable and can be easily applied to any text collection. Clustering can also reveal interesting and potentially unexpected or new trends in a group of documents, and groups together documents that are similar to one another but different from the rest of the collection, such as all the document written in Japanese that appear in a collection of primarily English articles.

The disadvantages of clustering include an unpredictability in the form and quality of results the difficulty of labeling the groups, and the counter-intuitiveness of cluster subhierarchies. Some algorithms [72, 150] build clusters around dominant phrases, which makes for understandable labels (see Figure 0.14), but the contents of the clusters are not necessarily coherent between and among the groups.

Figure 0.15 shows the output of clustering on Web search results, using the Clusty system by Vivisimo on the query “senate.” Two of the clusters are shown expanded to reveal their subhierarchies. The topmost cluster is labeled *Biography*, *Constituent Services*, with subclusters labeled *Photos*, *Issues/news*, *Visiting Washington*, *Voting record*, *Virginia*, *Maine*, among others. It is unclear exactly what this cluster is representing; if it is the U.S. Senate, there are also pages about the U.S. Senate in other clusters, and in any case, the top-most label is not particularly informative. The next top-level cluster, labeled *Senate Committee*, is selected in order to show its constituent documents (on

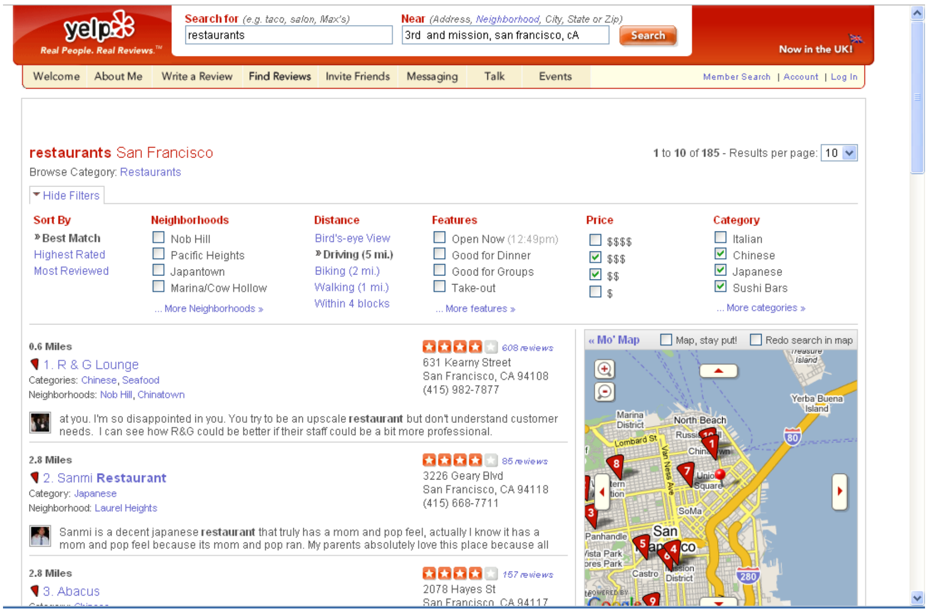


Figure 0.13 Faceted navigation at yelp.com.

the right side of the Figure), which range from the main U.S. Senate Web page (whose focus is not the various committees) to Web pages of some U.S. Senate committees, to pages from Kansas and Cambodia. The third main cluster, whose label is *Votes*, is also expanded, showing subclusters labeled *Constituent Services*, *Obama Budget*, *Expand*, and *Senate Calendar*.

This mixed bag of topics, including overlap among the groups, is typical for document clustering. Usability results show that users do not like disorderly groupings that are often produced by clustering, preferring understandable hierarchies in which categories are presented at uniform levels of granularity [107, 112].

One drawback of faceted interfaces versus clusters are that the categories of interest must be known in advance, and so important trends in the data may not be noticed. But by far the largest drawback is the fact that in most cases the category hierarchies are built by hand and automated assignment of categories to items is only partly successful, although attempts to build facet hierarchies are making progress [123].

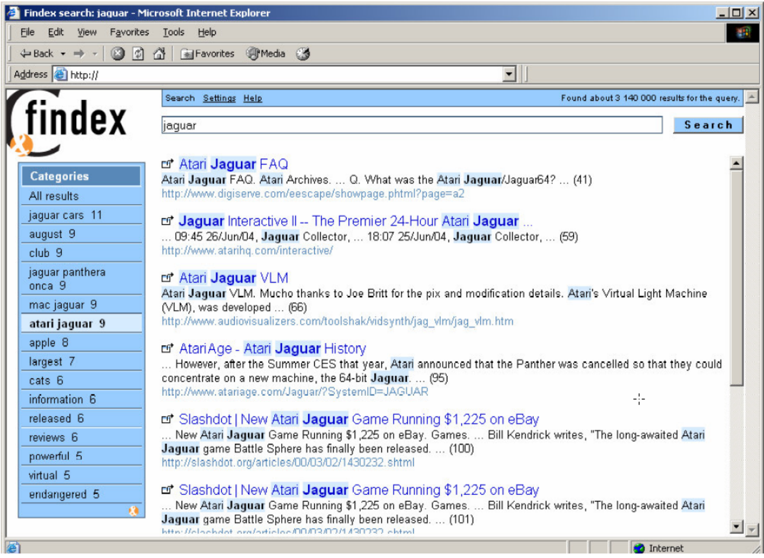


Figure 0.14 Output produced using Findex clustering [72].

Figure 0.15 Cluster output on the query "senate", from Clusty.com.

0.4 Visualization in Search Interfaces

This book is primarily about search of textual information. Text as a representation is highly effective for conveying abstract information, but reading and even scanning text is a cognitively taxing activity, and must be done in a linear fashion.

By contrast, images can be scanned quickly and the visual system perceives information in parallel. People are highly attuned to images and visual information, and pictures and graphics can be captivating and appealing. A visual representation can communicate some kinds of information much more rapidly and effectively than any other method. Consider the difference between a written description of a person's face and a photograph of it, or the difference between a table of numbers containing a correlation and a scatter plot showing the same information.

Over the last few years, information visualization has become a common presence in news reporting and financial analysis, and creative, innovative ideas about how to visualization information have blossomed and spread throughout the Web. Social visualization sites such as ManyEyes [135] allow people to upload their data and explore the contents using bar charts, bubble graphs, line plots and more, and data analysis tools such as Tableau help analysts slice and rearrange their data visually.

However, visualization of inherently abstract information is more difficult, and visualization of textually represented information is especially challenging. Language is our main means of communicating abstract ideas for which there is no obvious physical manifestation. Words and concepts do not have an inherent order, which makes words difficult to plot in a meaningful way along an axies.

Despite the difficulties, researchers are attempting to represent aspects of the information access process using information visualization techniques. Aside from using *icons* and *color highlighting*, spatial layout in the form of lines, bars, circles and canvas views of information are common. *Sparklines* [132] are miniaturized graphics, shown inline with text and tables. Interactivity seems to be an especially important property for visualizing abstract information; the main interactive information visualization techniques include *panning and zooming*, *distortion-based views* (including *focus plus context*), and the use of *animation* to retain context and help make occluded information visible.

Experimentation with visualization for search has been primarily applied in the following ways:

- Visualizing Boolean Syntax,
- Visualizing Query Terms within Retrieval Results,
- Visualizing Relationships among Words and Documents,
- Visualization for Text Mining.

Each is touched on in the discussion below.

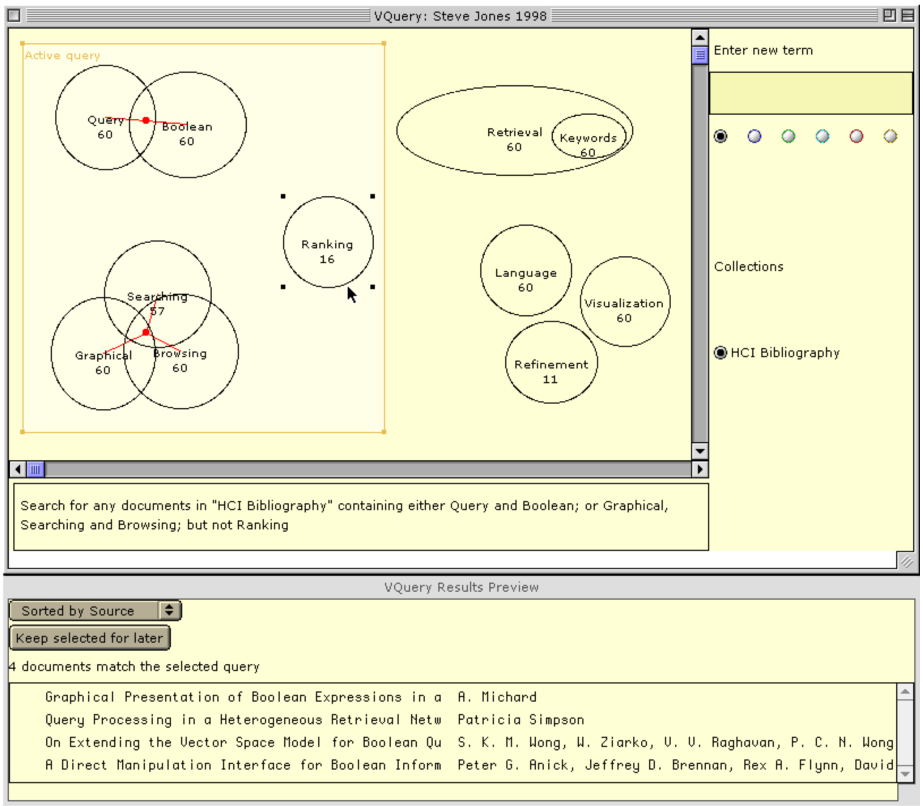


Figure 0.16 The VQuery [67] Venn Diagram interface for Boolean query specification.

Visualizing Boolean Syntax

As noted above, Boolean query syntax is difficult for most users and is rarely used in Web search. For many years, researchers have experimented with how to visualize Boolean query specification, in order to make it more understandable. A common approach is to show Venn diagrams visually; Hertzum & Frokjaer [56] found that a simple Venn diagram representation produced more accurate results than Boolean syntax. A more flexible version of this idea was seen in the VQuery system [67] (see Figure 0.16). Each query term is represented by a circle or oval, and the intersection among circles indicates ANDing (conjoining) of terms. VQuery represented disjunction by sets of circles within an active area of the canvas, and negation by deselecting a circle within the active area.

One problem with Boolean queries is that they can easily end up with empty results or too many results. To remedy this, the filter-flow visualization allows users to lay out the different components of the query, and show via a graphical flow how many hits would result after each operator is applied [149]. Other visual representations of Boolean queries include lining up blocks vertically and horizontally [5] and representing components of queries as overlapping "magic" lenses [40].

Visualizing Query Terms within Retrieval Results

As discussed above, understanding the role played by the query terms within the retrieved documents can help with the assessment of relevance. In standard search results listings, summary sentences are often selected that contain query terms, and the occurrence of these terms are highlighted or boldfaced where they appear in the title, summary, and URL. Highlighting of this kind has been shown to be effective from a usability perspective.

Experimental visualizations have been designed that make this relationship more explicit. One of the most well-known is the TileBars interface [49], in which documents are shown as horizontal glyphs with the locations of the query term hits marked along the glyph (see Figure 0.17). The user is encouraged to break the query into its different facets, with one concept per line, and then the horizontal rows within each document's representation shows the frequency of occurrence of query terms within each topic. Longer documents are divided into subtopic segments, either using paragraph or section breaks, or an automated discourse segmentation technique called TextTiling [48]. Grayscale implies the frequency of the query term occurrences. The visualization shows where the discussion of the different query topics overlaps within the document.

A number of variations of the TileBars display have been proposed, including a simplified version which shows only one square per query term, and color gradation is used to show query term frequency [60]. Two more-compressed versions showed grey-scaled hits within a document-shaped icon [54], or showing hits as colors in a pie chart [2], but these views do not show positions of term overlap.

Other approaches to showing query term hits within document collections include placing the query terms in bar charts, scatter plots, and tables. A usability study by Reiterer et al. [109] compared five views: a standard Web search engine-style results listing, a list view showing titles, document metadata, and a graphic showing locations of query term hits along the document, with height indicating frequency (see Figure 0.18), a color TileBars-like view, with document titles alongside the graphic, a color bar chart view like that of Veerasamy & Belkin [134], and a scatter plot view plotting relevance scores against date of publication.

When asked for subjective responses, the 40 participants on average chose the field-sortable view first, followed by the TileBars, followed by the Web-style listing. The bar chart and scatter plot received negative responses. When compared to the Web-style listing, there were no significant differences for task effectiveness for the other conditions, except for bar charts, which were significantly

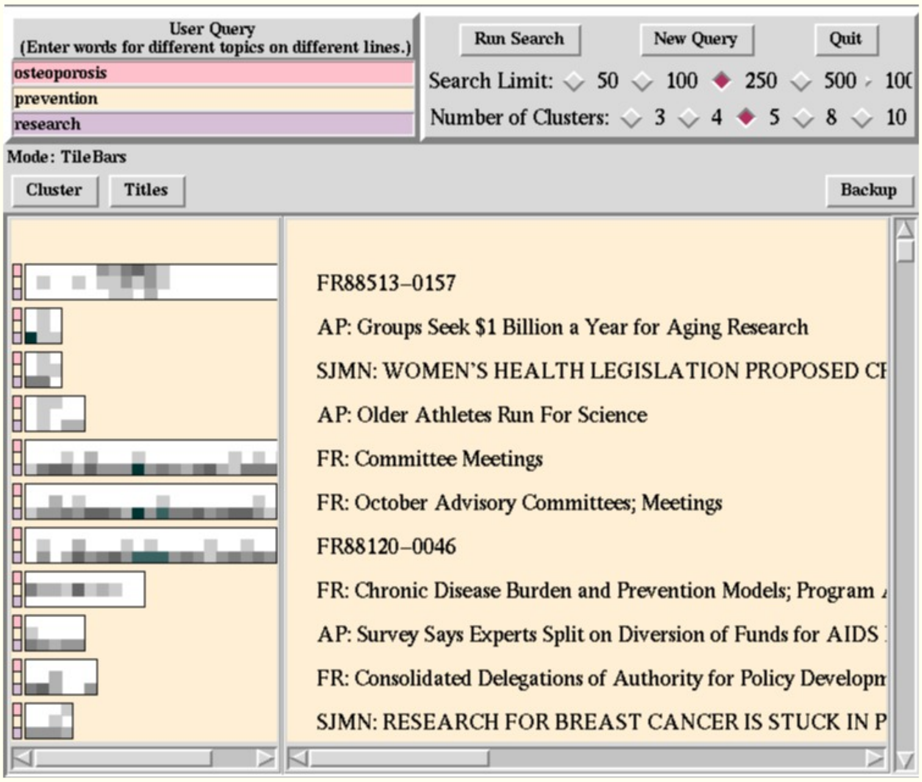
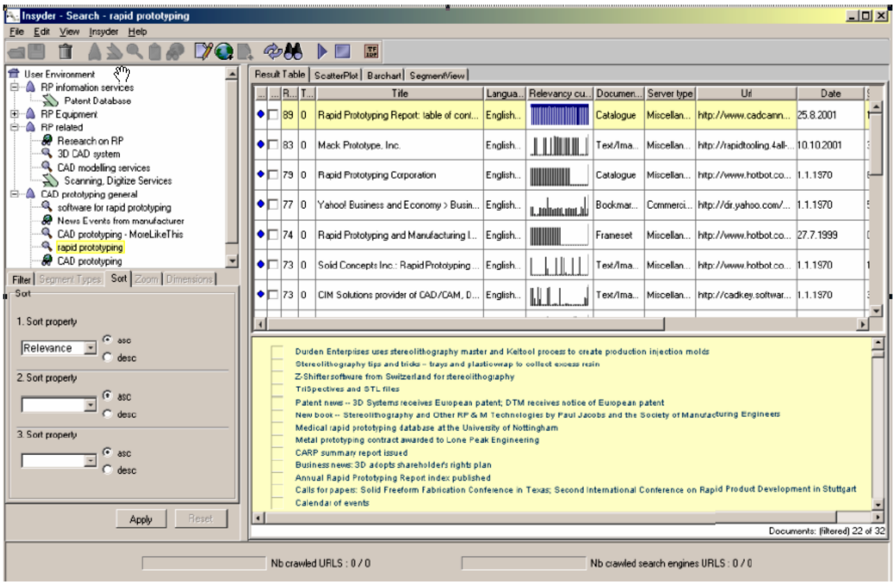


Figure 0.17 The TileBars visualization of query term hits within retrieved documents, from [49].

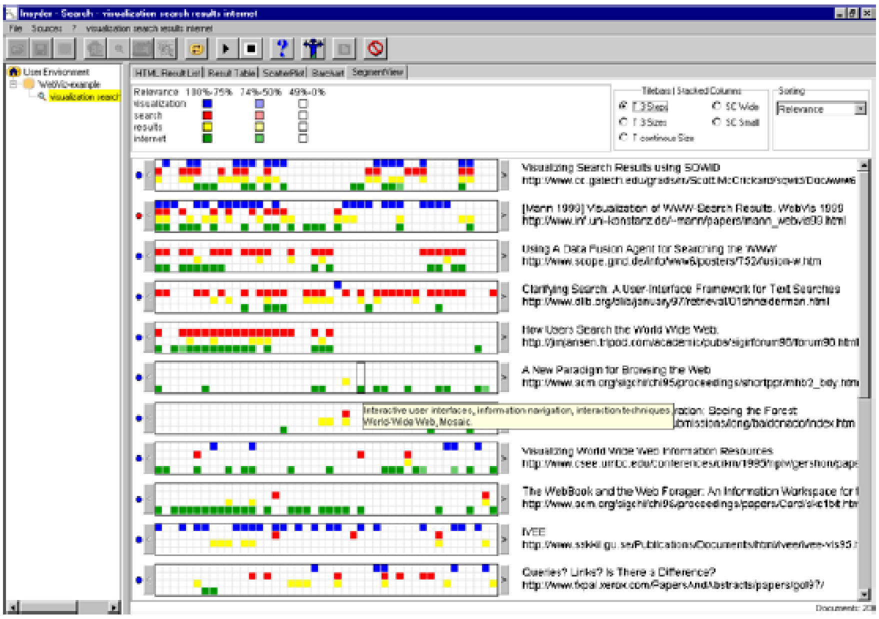
worse. All conditions had significantly higher mean task times than the Web-style listing. This last point echos a common result for search visualizations – even in those cases where the visualization is seen as being helpful, it usually causes search times to be longer than for text-only views. This may be caused by the time needed to switch from interpreting images to reading text, which are different cognitive functions.

Another variation on the idea of showing query term hits within documents is to show *thumbnails* – miniaturized rendered versions of the visual appearance of the document. One experiment using thumbnails found they were no better than blank squares for improving search results, [28] and another found participants were more likely to erroneously think documents were relevant when shown thumbnails plus titles vs. thumbnails alone [34]. That said, both studies found that the thumbnails had subjective appeal.

The negative study results may stem from a problem with the size of the



(a)



(b)

Figure 0.18 (a) Field-sortable search results view, including a sparklines-style graphic showing locations of query hits, which is a simplification of TileBars and (b) colored TileBars view from [109].

Figure 0.19 Textually enhanced thumbnails, from [147].

thumbnails; newer results suggest that increasing the size improves their use for search results display [70]. A related study shows that making the query terms larger and more visible via highlighting within the thumbnail improves its usability for search results for certain types of tasks [147] (see Figure 0.19). The use of larger thumbnails with enlarged text in place for search results is being explored in the SearchMe Web search engine using cover-flow animation.

Visualizing Relationships Among Words and Documents

Numerous visualization developers have proposed variations on the idea of placing words and documents on a two-dimensional canvas, where proximity of glyphs represents semantic relationships among the terms or documents. An early version of this idea is seen in the VIBE interface, where queries are laid out on a plane, and documents that contain combinations of the queries are placed mid-

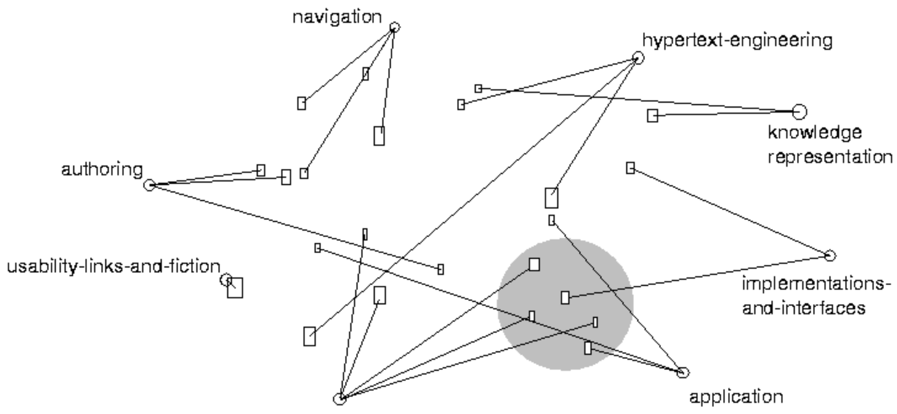


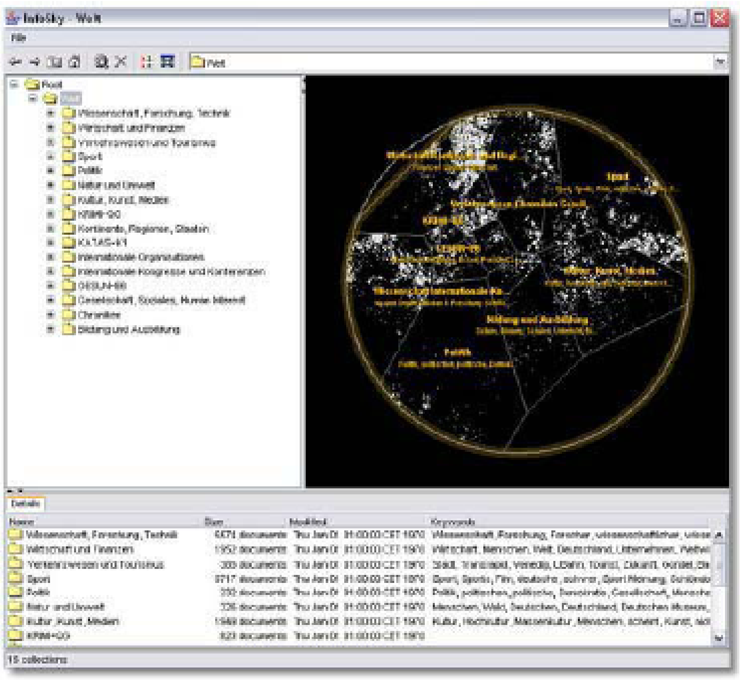
Figure 0.20 The VIBE display, in which query terms are laid out in a 2D space, and documents are arranged according to which subset of the text they share, from [98].

way between the icons representing those terms (see Figure 0.20) [98]. A more modern version of this idea is seen in the Aduna Autofocus product, and the Lyberworld project [55] presented a 3D version of the ideas behind VIBE.

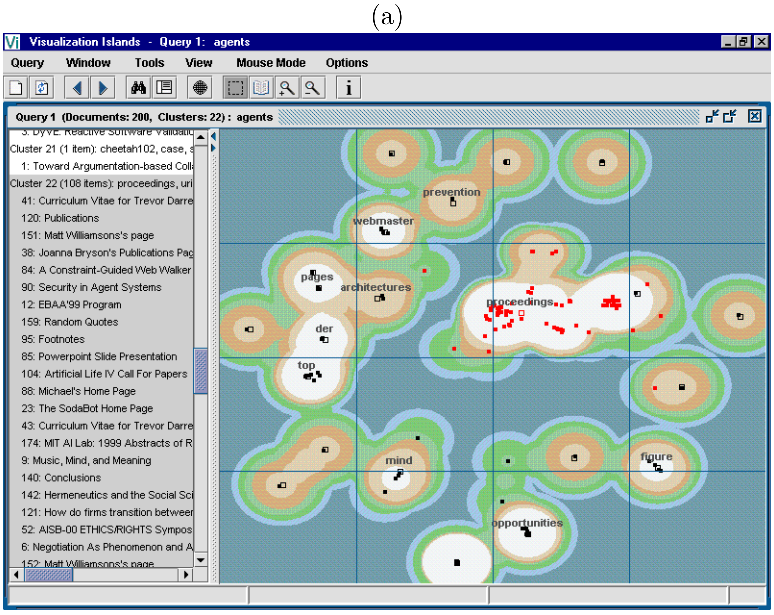
Another variation of this idea is to map documents or words from a very high-dimensional term space down into a two-dimensional plane, and show where the documents or words fall within that plane, using 2D or 3D [3, 43, 57, 144, 145]. This variation on clustering can be done to documents retrieved as a result of a query, or documents that match a query can be highlighted within a pre-processed set of documents. Figure 0.21 shows two variations on these *starfield* displays.

These views are relatively easy to compute and can be visually striking. However, evaluations that have been conducted so far provide negative evidence as to their usefulness [43, 62, 75, 114]. The main problems are that the contents of the documents are not visible in such views, and a 2D representation does not do justice to the complex ways in which meanings can interact.

A more promising application of this kind of idea is in the layout of thesaurus terms, in a small network graph, such as used in Visual Wordnet (see Figure 0.22). This view simplifies the large WordNet database by only showing nodes directly connected to the target node. This use of nodes and links views has not been evaluated in a published study, but its use for organizing search results does not seem to work well [125].



(a)



(b)

Figure 0.21 The idea of representing documents as a 2D or 3D mapping of glyphs has been proposed many times. Two examples are shown here: (a) InfoSky, from [43], (b) xFIND's VisIslands, from [3].

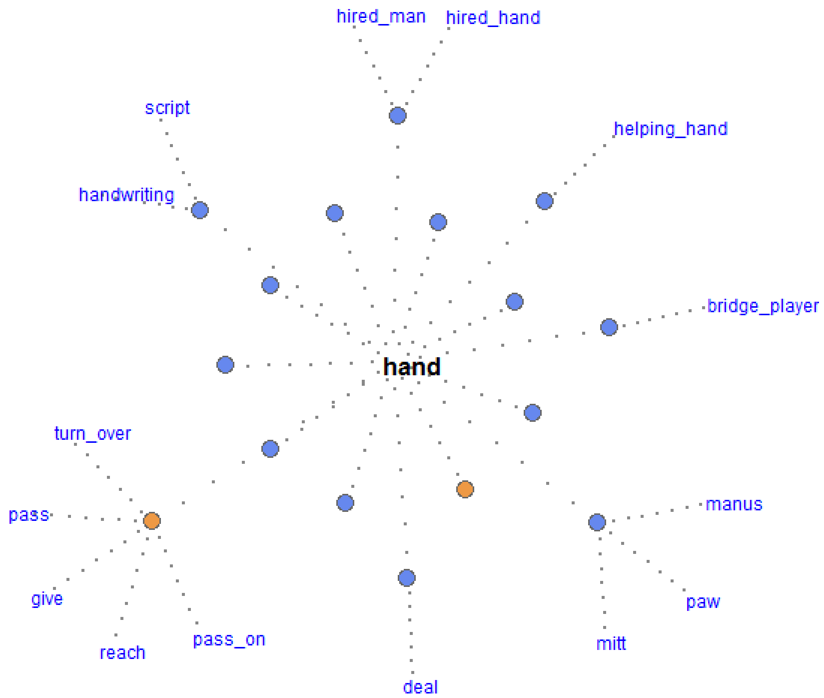


Figure 0.22 The Visual WordNet view of the WordNet lexical thesaurus, from http://kylescholz.com/projects/wordnet/.

Visualization for Text Mining

The subsections above show that usability results for visualization in search are not particularly strong. It seems in fact that visualization is better used for purposes of analysis and exploration of textual data. Most users of search systems are not interested in seeing how words are distributed across documents, or in viewing the most common words within a collection, but these are interesting activities for computational linguists, analysts, and curious word enthusiasts. Visualizations such as the Word Tree [137] show a piece of a text concordance, allowing the user to view which words and phrases commonly precede or follow a given word (see Figure 0.23), or the NameVoyager explorer [136], which shows frequencies of baby names for U.S. children across time (see Figure 0.24).

Visualization is also used in search interfaces intended for analysts. Figure 0.25 shows the TRIST information "triage" system [69, 108] which is intended to aid intelligence analysts in their work. Search results were represented as document icons; thousands of documents can be viewed in one display, and the

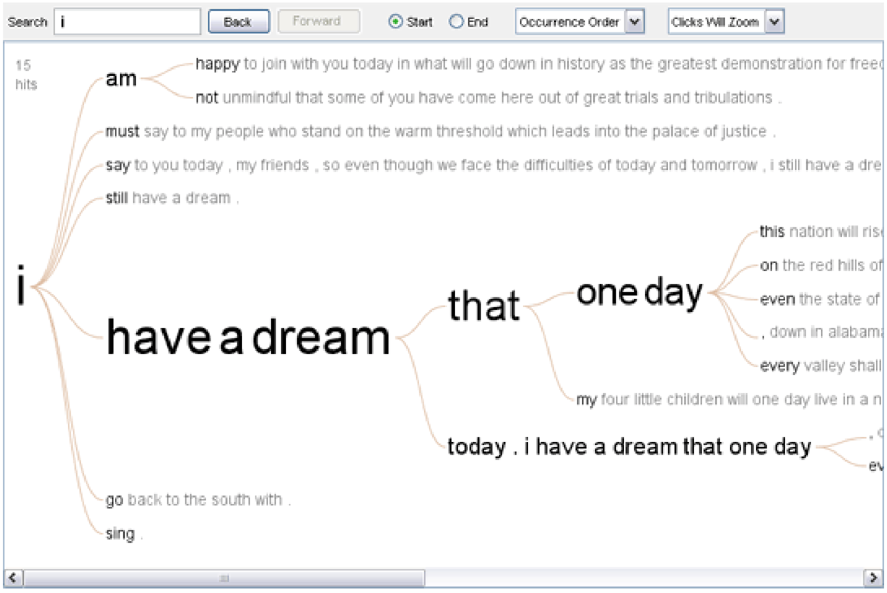


Figure 0.23 The Word Tree visualization, on Martin Luther King's *I have a dream* speech, from [137].

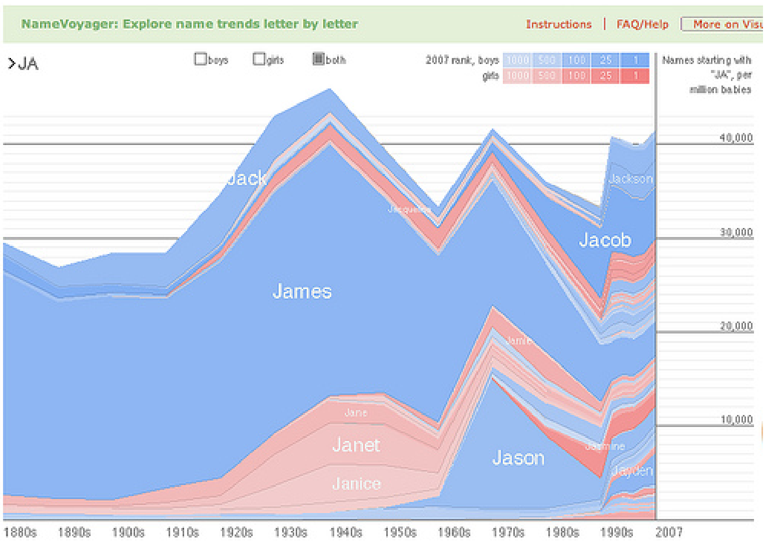


Figure 0.24 A visualization of the relative popularity of baby names over time, with names beginning with the letters JA, from babynameswizard.com.

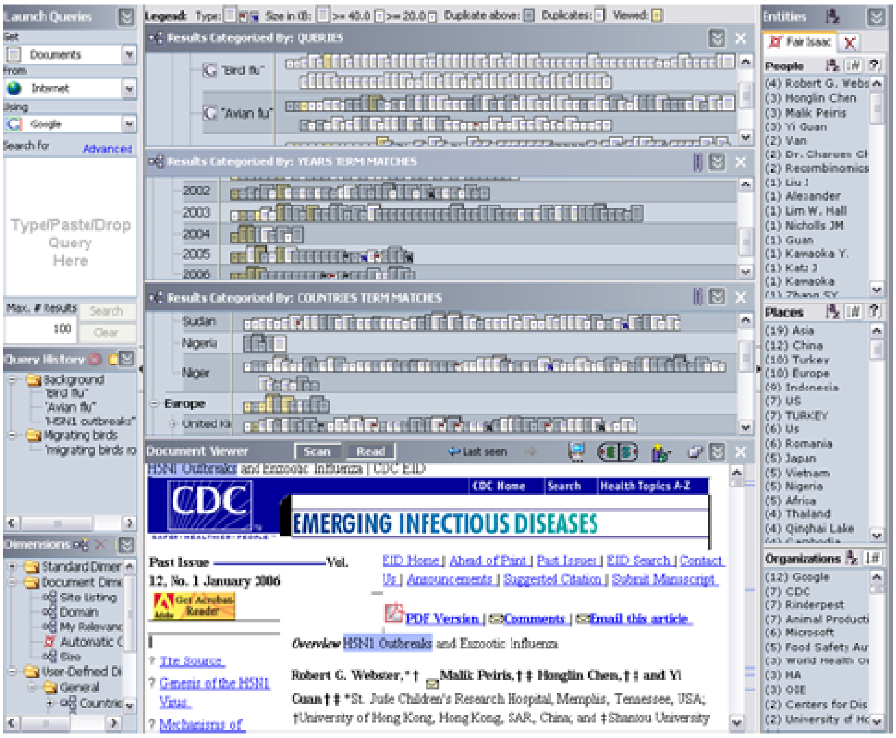


Figure 0.25 The TRIST interface shown responding to queries on related to Avian Flu, from [108].

system supports multiple linked dimensions that allow for finding characteristics and correlations among the documents. Icon color was used to show the user which documents they have already viewed, and icon size and shape show the length and type of the document, respectively. This seems to be an effective system, enabling its designers to win the IEEE Visual Analytics Science and Technology (VAST) contest [46] for two years running.

0.5 Design and Evaluation of Search Interfaces

User interface design is a practice whose techniques are encompassed by the field of Human-Computer Interaction (HCI). This field studies how people think about, respond to, and use technology, and how best to design user interfaces to suit people's needs and inclinations. Based on years of experience, a set of practices and guidelines have been developed to facilitate the design of successful interfaces. The practices are collectively referred to as *user-centered design*, which emphasizes building the design around people's activities and though processes, rather than the converse.

The design process begins by determining what the intended users' *goals* are, and then devising an interface that can help people achieve those goals by completing a series of *tasks*. Goals in the domain of information access can range quite widely, from finding a plumber to keeping informed about a business competitor, from writing a publishable scholarly article to investigating an allegation of fraud. Information access tasks are used to achieve these goals. These tasks span the spectrum from asking specific questions to exhaustively researching a topic.

A user interface is designed via an iterative process, in which the goals and tasks are elucidated via user research, and then initial designs are created – often based on existing designs, but potentially including new ideas. These initial designs are tested with prospective users, and then are assessed and re-designed, and evaluated again, in a cycle that can repeat numerous times.

The procedure for evaluating a user interface is often different than that of evaluating a ranking algorithm or a crawling technique. A crawler can be assessed by crisp quantitative metrics such as coverage and freshness. A ranking algorithm can be evaluated by precision, recall, and speed. But the quality of a user interface is determined by how people respond to it. Subjective responses are as, if not more, important than quantitative measures, because if a person has a choice between two systems, they will use the one they prefer. The reasons for preference may be determined by a host of factors, including speed, familiarity, aesthetics, preferred features, or perceived ranking accuracy. Often the preferred choice is the familiar one. In search interfaces particularly, a new interface idea must be perceived as being qualitatively better than an old one before users will switch. This observation may explain why search results displays have not changed markedly since Web search first appeared.

How best to evaluate a user interface depends on the current stage in the development cycle. When starting with a new design or idea, *discount* usability methods are typically used. One discount method involves showing a few potential users several different designs and allowing them to indicate which parts are promising and which are not. Often paper mockups are used for this task, both because they are fast to develop and easy to throw away, and because there is evidence that study participants are more willing to criticize something that is clearly unfinished than something that looks polished [14, 110]. Usually the design-test-redesign process is cycled through several times before an acceptable starting point for an interactive prototype found.

Next, low-functionality interactive designs are developed and tested with a small number of participants, to obtain both their subjective responses and to ascertain what, if anything, works well in the design and to determine which aspects are confusing or do not work well. If a participant looks at a screen and does not know what to do, that suggests a redesign is needed.

Another commonly used discount evaluation method is *heuristic evaluation*, in which usability experts "walk through" a design and evaluate the functionality in terms of how well it measures up against a set of design guidelines. This kind of evaluation can work very well at discovering usability problems, as seasoned experts can often anticipate problems accurately, but should be combined with responses from the target user audience.

After more design iteration, an interactive system can be developed and tested with study participants in a more formal experiment, comparing the new design against a competitive baseline, or comparing two candidate designs, according to a variety of metrics. In evaluating search interfaces, it is important for participants to be *motivated* to do well on the task [18, 122]. Asking people who do not care or know much about camera lenses to do extensive searches for camera lenses will most likely not produce realistic results. To help guarantee motivation, the study participants should be engaged by the subject matter of the queries and the information collection, and should drawn from a pool of people who would in practice be using the final system, but close surrogates can be used (e.g., nursing students rather than nurse practitioners may be more willing to test out a design).

Formal experiments are usually conducted to produce research results that are published and used by the wider community of practice, although some organizations conduct formal studies internally. Formal studies follow practices seen in scientific fields, by recruiting study participants, and comparing a control condition against experimental conditions. A formal experiment must be carefully designed to take into account potentially confounding factors, by balancing the order in which competing designs are shown, and the experimenters must avoid expressing bias towards one design over another. Such a study is used to answer specific questions about how well particular design elements work or if a new feature works better than the status quo [51].

This kind of study can uncover important subjective results, such as whether a new design is strongly preferred over a baseline, but the nature of search interfaces makes it difficult to find accurate quantitative differences with a small number of participants. There are many factors that contribute to this difficulty, including the strong effect of tasks or queries on system behavior – in many studies of search systems, be they interactive or batch analysis, the differences between tasks swamps out differences between systems or between participants. Another problem is that a search being conducted by a person can go so many different ways that it is difficult to compare quantitative outcomes directly. Finally, sometimes the timing variable is not the right measure for evaluating an interactive search session, because a tool that allows the searcher to learn about their subject matter as they search may be more beneficial, but take more time, than a competing design.

To counter some of these problems, two approaches to evaluating search interfaces have gained in popularity in recent years. One is to conduct a *longitudinal* study, meaning that participants use a new interface for an extended period of time, and their usage is monitored and logged [32, 118]. Evaluation is based both on the objective measures recorded in the logs as well as via questionnaires and interviews with the participants. In some cases, the benefits of a new approach only become clear to searchers after they interact with it over time, and so a long-term study can uncover such benefits. For example, a study by [72] found that participants changed their searching patterns in response to an interface over time, and also revealed under which circumstances the new feature was useful, and when it was not needed. Alternatively, an interface that is initially engaging (say, due to impressive graphics) may become tiresome to use over time, leading to a desire to move back to a baseline familiar interface.

Another major evaluation technique that has become increasingly popular in the last few years is to perform experiments on a massive scale on already heavily-used Web sites . This approach is often referred to as *bucket testing, A/B testing*, or *parallel flights*. A search engine that received hundreds of thousands, or even millions of queries per day can be used to perform studies in which a randomly selected subset of users is shown a new design, and their actions are compared to another randomly selected control group that continues to use the existing interface [6, 78, 77]. This differs from a formal usability study in that participants do not knowingly "opt in" to the study; rather, visitors to the site are selected to be shown alternatives without their knowledge or consent (note that this sort of service is allowed under most organization's Web sites' terms of use statements).

Studies conducted in this manner can often be completed as quickly as in a twenty-four hour period, although one-two weeks is typically recommended. Performance measures such as which links are clicked on are especially informative in tests like these. For instance, an interface that shows query suggestions can be compared against one that does not, and how often the suggestions are clicked on can be recorded. As another example, the effects of inserting multimedia results within a textual result listing can be assessed in terms of what kinds of changes are seen in the clicks on the surrounding links as well as how often the new information is clicked on. User behavior in the control versus the experimental condition are compared, and in some cases the behavior can be compared for identical queries in the two conditions, because the volume of searchers is so large. A potential downside to such studies is that users who are familiar with the control version of the site may initially reactive negatively to the unfamiliarity of the interface, and so some studies factor the initial reaction into their assessments. Also, subjective information usually is not gathered with this technique; many experimenters use follow-up surveys to determine subjective responses.

0.6 Trends and Research Issues

This chapter has discussed many ideas for improving the human-computer interaction experience for information seekers. This is still a rapidly developing area, and improvements in the interface are likely to lead the way toward better search results and better-enabled information creators and users. Important future areas of advancement are social search, mobile search interface, multimedia search, and natural language-oriented querying, as discussed in detail by Hearst [51].

0.7 Bibliographic Discussion

The book *Search User Interfaces* by Hearst [51] provides an in-depth treatment of the topics touched on in this chapter. See also a related book on information seeking behavior, *Information Seeking in Electronic Environments* by Marchionini [90]. Some discussion of search interfaces occurs as well in *Understanding Digital Libraries* by Lesk [84].

There are a number of excellent books on HCI and interface design generally, including *Designing the User Interface* by Shneiderman et al. [119], and *Observing the User Experience: A Practitioner's Guide to User Research* by Kuniavsky [81]. An old but nevertheless still excellent book on evaluating user interfaces is *Usability Engineering* by Nielsen [96].

Many books describe how to design Web sites, but those most relevant to this chapter are *Information Architecture for the World Wide Web, 3rd edition* by Morville and Rosenfeld [93], which includes two chapters on search, and *Designing Web Navigation, Optimizing the User Experience*, by Kalback [73], which discusses navigation design for Web sites. Additionally, there are dozens, if not hundreds, of Web sites dedicated to usability and user interface design.

There are now excellent books on how to design using information visualization, including *Now You See It: Simple Visualization Techniques for Quantitative Analysis* and *Information Dashboard Design: The Effective Visual Communication of Data* by Few [38, 39] and *The Visual Display of Quantitative Information*, by Tufte [131], although these do not focus on visualizing text or search.

References

[1] J. Allan. HARD Track overview in TREC 2004 high accuracy retrieval from documents. *Proceedings of the Thirteenth Text REtrieval Conference (TREC'04)*, 2005.

[2] T. Anderson, A. Hussam, B. Plummer, and N. Jacobs. Pie charts for visualizing query term frequency in search results. *Proceedings of the Fifth International Conference on Asian Digital Library*, pages 440–451, 2002.

[3] K. Andrews, C. Gütl, J. Moser, V. Sabol, and W. Lackner. Search Result Visualisation with xFIND. *Proceedings of User Interfaces to Data Intensive Systems (UIDIS 2001)*, pages 50–58, 2001.

[4] P. Anick. Using Terminological Feedback For Web Search Refinement: A Log-Based Study. *Proceedings of the 26th Annual International ACM SIGIR Conference on Research and development in information retrieval (SIGIR'03)*, pages 88–95, 2003.

[5] P. Anick, J. Brennan, R. Flynn, D. Hanssen, B. Alvey, and J. Robbins. A direct manipulation interface for boolean information retrieval via natural language query. In *Proceedings of the 13th Annual International ACM SIGIR Conference on Research and development in information retrieval (SIGIR'90)*, pages 135–150, Brussels, Belgium, 1990.

[6] P. Anick and R.G.P. Kantamneni. A longitudinal study of real-time search assistance adoption. In *Proceedings of the 31st Annual International ACM SIGIR Conference on Research and development in information retrieval (SIGIR'08)*, pages 701–702. ACM, 2008.

[7] A. Aula. Enhancing the readability of search result summaries. *Proceedings of HCI 2004*, pages 6–10, 2004.

[8] A. Aula. *Studying user strategies and characteristics for developing web search interfaces.* PhD thesis, University of Tampere, Finland, Ph.D. Dissertation, Dissertations in Interactive Technology, Number 3., 2005.

[9] D. Barreau and B.A. Nardi. Finding And Reminding: File Organization From The Desktop. *Acm Sigchi Bulletin*, 27(3):39–43, 1995.

[10] M.J. Bates. Information search tactics. *Journal of the American Society for Information Science*, 30(4):205–214, 1979.

[11] M.J. Bates. The design of browsing and berrypicking techniques for the on-line search interface. *Online Review*, 13(5):407–431, 1989.

[12] M.J. Bates. Where should the person stop and the information search interfaces start? *Information Processing and Management*, 26(5), 1990.

[13] P. Baudisch, B. Lee, and L. Hanna. Fishnet, a fisheye web browser with search term popouts: a comparative evaluation with overview and linear view. *Proceedings of the working conference on Advanced Visual Interfaces (AVI'04)*, pages 133–140, 2004.

[14] M. Beaudouin-Lafon and W. Mackay. Prototyping Tools And Techniques. In *Human-Computer Interaction Handbook*. Lawrence Erlrbaum Associates, 2003.

[15] N.J. Belkin, D. Kelly, G. Kim, J.Y. Kim, H.J. Lee, G. Muresan, M.C. Tang, X.J. Yuan, and C. Cool. Query length in interactive information retrieval. *Proceedings of the 26th Annual International ACM SIGIR Conference on Research and development in information retrieval (SIGIR'03)*, pages 205–212, 2003.

[16] D. Bilal. Children's use of the Yahooligans! Web search engine. I. Cognitive, physical, and affective behaviors on fact-based search tasks. *Journal of the American Society for Information Science*, 51(7):646–665, 2000.

[17] P. Bruza, R. Mcarthur, and S. Dennis. Interactive Internet Search: Keyword, Directory And Query Reformulation Mechanisms Compared. *Proceedings of the 23th Annual International ACM SIGIR Conference on Research and development in information retrieval (SIGIR'00)*, pages 280–287, 2000.

[18] G. Buchanan, S.J. Cunningham, A. Blandford, J. Rimmer, and C. Warwick. Information Seeking by Humanities Scholars. *Proceedings of the European Conference on Digital Libraries (ECDL'05)*, 2005.

[19] D. Byrd. A Scrollbar-based Visualization for Document Navigation. *Proceedings of the Fourth ACM International Conference on Digital Libraries*, 1999.

[20] M. Chen, M.A Hearst, J. Hong, and J. Lin. Cha-Cha: A System For Organizing Intranet Search Results. *Proceedings of the 2nd conference on USENIX Symposium on Internet Technologies and Systems*, pages 11–14, 1999.

[21] C.L.A. Clarke, E. Agichtein, S. Dumais, and R.W. White. The influence of caption features on clickthrough patterns in web search. In *Proceedings of the 30th Annual International ACM SIGIR Conference on Research and development in information retrieval (SIGIR'07)*, pages 135–142. ACM Press New York, NY, USA, 2007.

[22] C.L.A. Clarke, G.V. Cormack, and F.J. Burkowski. Shortest substring ranking (multitext experiments for TREC-4). In Donna Harman, editor, *Proceedings of the Fifth Text REtrieval Conference (TREC-5)*, 1996.

[23] W.B. Croft, S. Cronen-Townsend, and V. Larvrenko. Relevance Feedback And Personalization: A Language Modeling Perspective. *Delos Workshop: Personalisation And Recommender Systems In Digital Libraries*, 2001.

[24] S. Cronen-Townsend, Y. Zhou, and W.B. Croft. A Language Modeling Framework For Selective Query Expansion. Technical Report Technical Report Ir-338, Center For Intelligent Information Retrieval, University Of Massachusetts, 2004.

[25] S. Cucerzan and E. Brill. Spelling Correction As An Iterative Process That Exploits The Collective Knowledge Of Web Users. *Proceedings Of Conference On Empirical Methods In Natural Language Processing (Emnlp'04)*, pages 293–300, 2004.

[26] E. Cutrell, D. Robbins, S. Dumais, and R. Sarin. Fast, flexible filtering with Phlat. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI'06)*, pages 261–270, 2006.

[27] D.R. Cutting, J.O. Pedersen, D. Karger, and J.W. Tukey. Scatter/Gather: A cluster-based approach to browsing large document collections. In *Proceedings of the 15th Annual International ACM SIGIR Conference on Research and development in information retrieval (SIGIR'92)*, pages 318–329, Copenhagen, Denmark, 1992.

[28] M. Czerwinski, M. van Dantzich, G. Robertson, and H. Hoffman. The contribution of thumbnail image, mouse-over text and spatial location memory to web page retrieval in 3D. *Proceedings of Human-Computer Interaction (INTERACT'99)*, 99:163–170, 1999.

[29] S. Dennis, R. Mcarthur, and P. Bruza. Searching The World Wide Web Made Easy? The Cognitive Load Imposed By Query Refinement Mechanisms. *Proceedings Of The Australian Document Computing Conference*, pages 65–71, 1998.

[30] J. Dinet, M. Favart, and J.M. Passerault. Searching for information in an online public access catalogue(OPAC): the impacts of information search expertise on the use of Boolean operators. *Journal of Computer Assisted Learning*, 20(5):338–346, 2004.

[31] A. Divoli, M.A. Hearst, and M.A. Wooldridge. Evidence for showing gene/protein name suggestions in bioscience literature search interfaces. *Pacific Symposium on Biocomputing*, 568:79, 2008.

[32] S. Dumais, E. Cutrell, Jj Cadiz, G. Jancke, R. Sarin, and D.C. Robbins. Stuff I've Seen: A System For Personal Information Retrieval And Re-Use. *Proceedings of the 26th Annual International ACM SIGIR Conference on Research and development in information retrieval (SIGIR'03)*, pages 72–79, 2003.

[33] S. Dumais, E. Cutrell, and H. Chen. Optimizing search by showing results in context. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI'01)*, pages 277–284, 2001.

[34] S. Dziadosz and R. Chandrasekar. Do Thumbnail Previews Help Users Make Better Relevance Decisions about Web Search Results. *Proceedings of the 25th Annual International ACM SIGIR Conference on Research and development in information retrieval (SIGIR'02)*, pages 365–366, 2002.

[35] D.E. Egan, J.R. Remde, L.M. Gomez, T.K. Landauer, J. Eberhardt, and C.C. Lochbaum. Formative design evaluation of SuperBook. *Transaction on Information Systems*, 7(1), 1989.

[36] D.E. Egan, J.R. Remde, T.K. Landauer, C.C. Lochbaum, and L.M. Gomez. Behavioral evaluation and analysis of a hypertext browser. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI'89)*, pages 205–210, May 1989.

[37] G. Eysenbach and C. Kohler. How do consumers search for and appraise health information on the world wide web? Qualitative study using focus groups, usability tests, and in-depth interviews. *British Medical Journal*, 324(7337):573–577, 2002.

[38] S. Few. *Information Dashboard Design: The Effective Visual Communication of Data*. O'Reilly, 2006.

[39] S. Few. *Now You See It: Simple Visualization Techniques for Quantitative Analysis*. Analytics Press, 2009.

[40] Ken Fishkin and Maureen C. Stone. Enhanced dynamic queries via movable filters. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI'95)*, volume 1, pages 415–420, 1995.

[41] K. Franzen and J. Karlgren. Verbosity and interface design. Technical report, Technical Report T2000, 2000.

[42] J. Graham. The reader's helper: a personalized document reading environment. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI'99)*, pages 481–488, 1999.

[43] M. Granitzer, W. Kienreich, V. Sabol, K. Andrews, and W. Klieber. Evaluating a System for Interactive Exploration of Large, Hierarchically Structured Document Repositories. *Proceedings of the IEEE Symposium on Information Visualization (INFOVIS'04)*, pages 127–133, 2004.

[44] L.A. Granka, T. Joachims, and G. Gay. Eye-tracking analysis of user behavior in WWW search. *Proceedings of the 27th Annual International ACM SIGIR Conference on Research and development in information retrieval (SIGIR'04)*, pages 478–479, 2004.

[45] S. L. Greene, S.J. Devlin, P.E. Cannata, and L.M. Gomez. No ifs, ands, or ors: A study of database querying. *International Journal of Man [sic] -Machine Studies*, 32(3):303–326, 1990.

[46] G. Grinstein, T. O'Connell, S. Laskowski, C. Plaisant, J. Scholtz, and M. Whiting. VAST 2006 Contest–A Tale of Alderwood. *Proceedings of the IEEE Symposium on Visual Analytics Science and Technology (VAST'06)*, pages 215–216, 2006.

[47] E. Hargittai. Classifying And Coding Online Actions. *Social Science Computer Review*, 22(2):210–227, 2004.

[48] M.A. Hearst. Multi-paragraph segmentation of expository text. In *Proceedings of the 32nd Meeting of the Association for Computational Linguistics (ACL'94)*, pages 9–16, Las Cruces, NM, June 1994.

[49] M.A. Hearst. TileBars: Visualization of Term Distribution Information in Full Text Information Access. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI'95)*, Denver, CO, May 1995.

[50] M.A. Hearst. Improving full-text precision using simple query constraints. In *Proceedings of the Fifth Annual Symposium on Document Analysis and Information Retrieval (SDAIR'96)*, Las Vegas, NV, 1996.

[51] M.A. Hearst. *Search User Interfaces.* Cambridge University Press, 2009.

[52] M.A. Hearst, A. Divoli, H. Guturu, A. Ksikes, P. Nakov, M.A. Wooldridge, and J. Ye. BioText Search Engine: beyond abstract search. *Bioinformatics*, 23(16):2196, 2007.

[53] M.A. Hearst, J. English, R. Sinha, K. Swearingen, and K.-P. Yee. Finding the flow in web site search. *Communications of the ACM*, 45(9), September 2002.

[54] T. Heimonen and N. Jhaveri. Visualizing Query Occurrence in Search Result Lists. *Proceedings of the IEEE Symposium on Information Visualization (INFOVIS'05)*, pages 877–882, 2005.

[55] Matthias Hemmje, Clemens Kunkel, and Alexander Willett. LyberWorld – a visualization user interface supporting fulltext retrieval. In *Proceedings of the 17th Annual International ACM SIGIR Conference on Research and development in information retrieval (SIGIR'94)*, pages 249–259, Dublin, Ireland, July 1994.

[56] Morten Hertzum and Erik Frokjaer. Browsing and querying in online documentation: A study of user interfaces and the interaction process. *ACM Transactions on Computer-Human Interaction (ToCHI)*, 3(2):136–161, 1996.

[57] E. Hetzler and A. Turner. Analysis Experiences Using Information Visualization. *IEEE Computer Graphics and Applications*, 24(5):22–26, 2004.

[58] C.R. Hildreth. OPAC research: laying the groundwork for future OPAC design. *The Online Catalogue: Development and Directions*, pages 1–24, 1989.

[59] W.C. Hill, J.D. Hollan, D. Wroblewski, and T. McCandless. Edit wear and read wear. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI'92)*, 92:3–9, 1992.

[60] O. Hoeber and X. D. Yang. A comparative user study of web search interfaces: HotMap, Concept Highlighter, and Google. *IEEE/WIC/ACM International Conference on Web Intelligence*, 2006.

[61] C. Hölscher and G. Strube. Web search behavior of Internet experts and newbies. *Computer Networks*, 33(1-6):337–346, 2000.

[62] K. Hornbæk and E. Frøkjær. Do Thematic Maps Improve Information Retrieval. *Human-Computer Interaction (INTERACT'99)*, pages 179–186, 1999.

[63] B.J. Jansen, A. Spink, and S. Koshman. Web searcher interaction with the Dogpile.com metasearch engine. *Journal of the American Society for Information Science and Technology*, 58(5):744–755, 2007.

[64] B.J. Jansen, A. Spink, and J.O. Pedersen. A Temporal Comparison Of Altavista Web Searching. *Journal Of The American Society For Information Science And Technology*, 56(6):559–570, 2005.

[65] C. Jenkins, C.L. Corritore, and S. Wiedenbeck. Patterns of information seeking on the Web: a qualitative study of domain expertise and Web expertise. *IT & Society*, 1(3):64–89, 2003.

[66] T. Joachims, L. Granka, B. Pan, H. Hembrooke, and G. Gay. Accurately Interpreting Clickthrough Data As Implicit Feedback. *Proceedings of the 28th Annual International ACM SIGIR Conference on Research and development in information retrieval (SIGIR'05)*, pages 154–161, 2005.

[67] Steve Jones and S. McInnes. Graphical query specification and dynamic result previews for a digital library. In *Proceedings of the 11th annual ACM symposium on User Interface Software and Technology (UIST'98)*, November 1998.

[68] W. Jones, S. Dumais, and H. Bruce. Once Found, What Then? A Study Of "Keeping" Behaviors In The Personal Use Of Web Information. *Proceedings Of The American Society For Information Science And Technology*, 39(1):391–402, 2002.

[69] D. Jonker, W. Wright, D. Schroh, P. Proulx, and B. Cort. Information Triage with TRIST. *Proceedings of the International Conference on Intelligence Analysis*, 2005.

[70] S. Kaasten, S. Greenberg, and C. Edwards. How People Recognise Previously Seen Web Pages from Titles, URLs and Thumbnails. *People and Computers*, pages 247–266, 2002.

[71] M. Kaisser, M.A. Hearst, and J.B. Lowe. Improving Search Results Quality by Customizing Summary Lengths. *Proceedings of the 46th Annual Meeting of the Association for Computational Linguistics: Human Language Technlogies (ACL-HLT'08)*, 2008.

[72] M. Käki. Findex: Search Result Categories Help Users When Document Ranking Fails. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI'05)*, pages 131–140, 2005.

[73] J. Kalbach. *Designing web navigation.* O'Reilly, 2007.

[74] D. Kelly, V.D. Dollu, and X. Fu. The Loquacious User: A Document-Independent Source Of Terms For Query Expansion. *Proceedings of the 28th Annual International ACM SIGIR Conference on Research and development in information retrieval (SIGIR'05)*, pages 457–464, 2005.

[75] A.J. Kleiboemer, M.B. Lazear, and J.O. Pedersen. Tailoring a retrieval system for naive users. In *Proceedings of the Fifth Annual Symposium on Document Analysis and Information Retrieval (SDAIR '96)*, Las Vegas, NV, 1996.

[76] J. Koenemann and N. Belkin. A case for interaction: A study of interactive information retrieval behavior and effectiveness. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI'96)*, volume 1, pages 205–212, 1996.

[77] R. Kohavi, R.M. Henne, and D. Sommerfield. Practical guide to controlled experiments on the web: listen to your customers not to the hippo. In *Proceedings of the 13th ACM SIGKDD international conference on Knowledge discovery and data mining (KDD'07)*, pages 959–967. ACM Press New York, NY, USA, 2007.

[78] R. Kohavi, L. Mason, R. Parekh, and Z. Zheng. Lessons and Challenges from Mining Retail E-Commerce Data. *Machine Learning*, 57(1):83–113, 2004.

[79] K. Kukich. Techniques for Automatically Correcting Words in Text. *ACM Computing Surveys*, 24(4), 1992.

[80] B. Kules and B. Shneiderman. Users can change their web search tactics: Design guidelines for categorized overviews. *Information Processing and Management*, 44(2):463–484, 2008.

[81] M. Kuniavsky. *Observing the User Experience: A Practitioner's Guide to User Research.* Morgan Kaufmann, 2003.

[82] T.K. Landauer, D.E. Egan, J.R. Remde, M.E. Lesk, C.C. Lochbaum, and D. Ketchum. Enhancing the usability of text through computer delivery and formative evaluation: the superbook project. In C. McKnight, A. Dillon, and J. Richardson, editors, *Hypertext: A Psychological Perspective*, pages 71–136. Ellis Horwood, 1993.

[83] A.W. Lazonder, H.J.A. Biemans, and I.G.J.H. Wopereis. Differences Between Novice And Experienced Users In Searching Information On The World Wide Web. *Journal Of The American Society For Information Science*, 51(6):576–581, 2000.

[84] Michael Lesk. *Practical Digital Libraries; Books, Bytes, & Bucks.* Morgan Kaufman, 1997.

[85] M. Li, M. Zhu, Y. Zhang, and M. Zhou. Exploring Distributional Similarity Based Models for Query Spelling Correction. In *Annual Meeting-Association For Computational Linguistics (ACL'06)*, 2006.

[86] J. Lin, M. DiCuccio, V. Grigoryan, and W.J. Wilbur. Navigating information spaces: A case study of related article search in PubMed. *Information Processing and Management*, 2008.

[87] J. Lin, D. Quan, V. Sinha, K. Bakshi, D. Huynh, B. Katz, and D.R. Karger. What Makes a Good Answer? The Role of Context in Question Answering. *Proceedings of Human-Computer Interaction (INTERACT'03)*, 2003.

[88] H.P. Luhn. *Keyword-in-context Index for Technical Literature (KWIC Index)*. International Business Machines Corp., Advanced Systems Development Division, 1959.

[89] G. Marchionini. Exploratory Search: From Finding To Understanding. *Communcations Of The Acm*, 49(4):41–49, 2006.

[90] Gary Marchionini. *Information Seeking in Electronic Environments*. Cambridge University Press, 1995.

[91] N. Milic-Frayling, R. Jones, K. Rodden, G. Smyth, A. Blackwell, and R. Sommerer. Smartback: supporting users in back navigation. *Proceedings of the 13th International Conference on World Wide Web (WWW'04)*, pages 63–71, 2004.

[92] M. Mitra, A. Singhal, and C. Buckley. Improving Automatic Query Expansion. *Proceedings of the 21st Annual International ACM SIGIR Conference on Research and development in information retrieval (SIGIR'98)*, pages 206–214, 1998.

[93] P. Morville and L. Rosenfeld. *Information Architecture for the World Wide Web*. O'Reilly, 2006.

[94] A. Nation. Visualizing websites using a hierarchical table of contents browser: Webtoc. In *Proceedings of the Third Conference on Human Factors and the Web*, Denver, CO, 1997.

[95] J. Nielsen. When Search Engines Become Answer Engines, 2004. http://www.useit.com/alertbox/20040816.html.

[96] Jakob Nielsen. *Usability Engineering*. Academic Press, 1993.

[97] 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 Conference on Human Factors in Computing Systems (CHI'93)*, Amsterdam, April 1993. IOS Press.

[98] K.A. Olsen, R.R. Korfhage, K.M. Sochats, M.B. Spring, and J.G. Williams. Visualization of a Document Collection with Implicit and Explicit Links-The Vibe System. *Scandinavian Journal of Information Systems*, 5:79–95, 1993.

[99] T. Paek, ST Dumais, and R. Logan. WaveLens: A new view onto internet search results. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI'04)*, pages 727–734, 2004.

[100] ES Patterson, EM Roth, and DD Woods. Predicting Vulnerabilities in Computer-Supported Inferential Analysis under Data Overload. *Cognition, Technology & Work*, 3(4):224–237, 2001.

[101] S. Piontek and K. Garlock. Creating a World Wide Web resource collection. *Internet Research: Electronic Networking Applications and Policy*, 6(4):20–26, 1996.

[102] P. Pirolli and S.K. Card. Information foraging. *Psychological Review*, 106(4):643–675, 1999.

[103] P. Pirolli and S.K. Card. The sensemaking process and leverage points for analyst technology as identified through cognitive task analysis. In *Proceedings of the 2005 International Conference on Intelligence Analysis*, McClean, VA, May 2005.

[104] Peter Pirolli. *Information Foraging Theory*. Oxford University Press, 2007.

[105] C. Plaisant, B. Shneiderman, K. Doan, and T. Bruns. Interface and data architecture for query preview in networked information systems. *ACM Transactions on Information Systems (TOIS)*, 17(3):320–341, 1999.

[106] A. Pollock and A. Hockley. What's Wrong with Internet Searching. *D-Lib Magazine*, 1997. www.dlib.org.

[107] W. Pratt, M.A. Hearst, and L. Fagan. A knowledge-based approach to organizing retrieved documents. In *Proceedings of 16th Annual Conference on Artificial Intelligence(AAAI 99)*, Orlando, FL, 1999.

[108] P. Proulx, S. Tandon, A. Bodnar, D. Schroh, W. Wright, D. Schroh, R. Harper, and W. Wright. Avian Flu Case Study with nSpace and GeoTime. In *Proceedings of the IEEE Symposium on Visual Analytics Science and Technology (VAST'06)*. IEEE, 2006.

[109] H. Reiterer, G. Tullius, and T.M. Mann. Insyder: a content-based visual-information-seeking system for the web. *International Journal on Digital Libraries*, pages 25–41, Mar 2005.

[110] M. Rettig. Prototyping For Tiny Fingers. *Communications Of The Acm*, 37(4), 1994.

[111] S.Y. Rieh. Judgment of information quality and cognitive authority in the Web. *Journal of the American Society for Information Science and Technology*, 53(2):145–161, 2002.

[112] K. Rodden, W. Basalaj, D. Sinclair, and K. R. Wood. Does organisation by similarity assist image browsing? In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI'01)*, pages 190–197, 2001.

[113] D.E. Rose, D. Orr, and R.G.P. Kantamneni. Summary attributes and perceived search quality. In *Proceedings of the 16th International Conference on World Wide Web (WWW'07)*, pages 1201–1202. ACM Press New York, NY, USA, 2007.

[114] D.M. Russell, M. Slaney, Y. Qu, and M. Houston. Being literate with large document collections: Observational studies and cost structure tradeoffs.

In *Proceedings of the 39th Annual Hawaii International Conference on System Sciences (HICSS'06)*, 2006.

[115] D.M. Russell, M.J. Stefik, P. Pirolli, and S.K. Card. The cost structure of sensemaking. In *Proceedings of the INTERCHI Conference on Human Factors in Computing Systems (CHI'93)*, Conceptual Analysis of Users and Activity, pages 269–276, 1993.

[116] I. Ruthven and M. Lalmas. A Survey On The Use Of Relevance Feedback For Information Access Systems. *The Knowledge Engineering Review*, 18(02):95–145, 2003.

[117] F. Sebastiani. Machine learning in automated text categorization. *ACM Computing Surveys (CSUR)*, 34(1):1–47, 2002.

[118] B. Shneiderman and C. Plaisant. Strategies for evaluating information visualization tools: multi-dimensional in-depth long-term case studies. *Proceedings of the 2006 conference Advanced Visual Interfaces (AVI'04), Workshop on Beyond time and errors: novel evaluation methods for information visualization*, pages 1–7, 2006.

[119] B. Shneiderman, C. Plaisant, M. Cohen, and S. Jacobs. *Designing the user interface: strategies for effective human-computer interaction, 5/E.* Addison Wesley, 2009.

[120] A. Spink, H. Greisdorf, and J. Bateman. From Highly Relevant To Not Relevant: Examining Different Regions Of Relevance. *Information Processing And Management*, 34(5):599–621, 1998.

[121] Jared Spool. *Web Site Usability: A Designer's Guide.* Morgan Kaufmann, 1998.

[122] Jared Spool. Usability beyond common sense, 2002. http://www.bcs-hci.org.uk/talks/Spool/UIE-BeyondCommonSense.pdf.

[123] E. Stoica, M.A. Hearst, and M. Richardson. Automating Creation of Hierarchical Faceted Metadata Structures. In *Human Language Technologies: The Annual Conference of the North American Chapter of the Association for Computational Linguistics (NAACL-HLT 2007)*, pages 244–251, 2007.

[124] A.G. Sutcliffe and M. Ennis. Towards a cognitive theory of information retrieval. *Interacting with Computers*, 10:321–351, 1998.

[125] R.C. Swan and J. Allan. Aspect Windows, 3-D Visualizations, and Indirect Comparisons of Information Retrieval Systems. *Proceedings of the 21st Annual International ACM SIGIR Conference on Research and development in information retrieval (SIGIR'98)*, pages 173–181, 1998.

[126] D. Tabatabai and B.M. Shore. How experts and novices search the Web. *Library & Information Science Research*, 27(2):222–248, 2005.

[127] T. Tao and C. Zhai. An exploration of proximity measures in information retrieval. In *Proceedings of the 30th Annual International ACM SIGIR Conference on Research and development in information retrieval (SIGIR'07)*, pages 295–302. ACM Press, 2007.

[128] J. Teevan, C. Alvarado, M.S. Ackerman, and D.R. Karger. The perfect search engine is not enough: a study of orienteering behavior in directed search. *Proceedings of the 27th Annual International ACM SIGIR Conference on Research and development in information retrieval (SIGIR'04)*, pages 415–422, 2004.

[129] J. Teevan, S.T. Dumais, and E. Horvitz. Personalizing Search Via Automated Analysis Of Interests And Activities. *Proceedings of the 28th Annual International ACM SIGIR Conference on Research and development in information retrieval (SIGIR'05)*, pages 449–456, 2005.

[130] A. Tombros and M. Sanderson. Advantages of query biased summaries in information retrieval. *Proceedings of the 21st Annual International ACM SIGIR Conference on Research and development in information retrieval (SIGIR'98)*, pages 2–10, 1998.

[131] Edward Tufte. *The Visual Display of Quantitative Information.* Graphics Press, Chelshire, CT, 1983.

[132] E.R. Tufte. Beautiful Evidence. *Information Design Journal*, 15(2):188–191, 2007.

[133] P. Vakkari. Relevance And Contributing Information Types Of Searched Documents In Task Performance. *Proceedings of the 23th Annual International ACM SIGIR Conference on Research and development in information retrieval (SIGIR'00)*, pages 2–9, 2000.

[134] A. Veerasamy and N.J. Belkin. Evaluation of a tool for visualization of information retrieval results. *Proceedings of the 19th Annual International ACM SIGIR Conference on Research and development in information retrieval (SIGIR'96)*, pages 85–92, 1996.

[135] F.B. Viégas, M. Wattenberg, F. van Ham, J. Kriss, and M. McKeon. Many Eyes: A Site for Visualization at Internet Scale. *IEEE Transactions On Visualization And Computer Graphics*, pages 1121–1128, 2007.

[136] M. Wattenberg and J. Kriss. Designing for Social Data Analysis. *IEEE Transactions on Visualization and Computer Graphics*, 12(4):549–557, 2006.

[137] M.V. Wattenberg and B. Fernanda. The Word Tree, an Interactive Visual Concordance. *Visualization and Computer Graphics, IEEE Transactions on*, 14(6):1221–1228, 2008.

[138] R.W. White, M. Bilenko, and S. Cucerzan. Studying the Use of Popular Destinations to Enhance Web Search Interaction. *Proceedings of the 30th Annual International ACM SIGIR Conference on Research and development in information retrieval (SIGIR'07)*, 2007.

[139] R.W. White, J. Jose, and I. Ruthven. A task-oriented study on the influencing effects of query-biased summarisation in web searching. *Information Processing and Management*, 39(5):707–733, 2003.

[140] R.W. White, J.M. Jose, and I. Ruthven. Using Top-Ranking Sentences for Web Search Result Presentation. *Proceedings of the 12th International Conference on World Wide Web (WWW'03)*, 2003.

[141] R.W. White and G. Marchionini. Examining the effectiveness of real-time query expansion. *Information Processing and Management*, 43(3), 2007.

[142] R.W. White and D. Morris. Investigating the querying and browsing behavior of advanced search engines users. *Proceedings of the 30th Annual International ACM SIGIR Conference on Research and development in information retrieval (SIGIR'07)*, 2007.

[143] R.W. White, I. Ruthven, and J.M. Jose. A Study Of Factors Affecting The Utility Of Implicit Relevance Feedback. *Proceedings of the 28th Annual International ACM SIGIR Conference on Research and development in information retrieval (SIGIR'05)*, pages 35–42, 2005.

[144] M.A. Whiting and N. Cramer. WebTheme: Understanding Web Information Through Visual Analytics. In *Proceedings of the First International Semantic Web Conference (ISWC'02)*, pages 460–468. Springer-Verlag London, UK, 2002.

[145] James A. Wise, James J. Thomas, Kelly Pennock, David Lantrip, Marc Pottier, and Anne Schur. Visualizing the non-visual: Spatial analysis and interaction with information from text documents. In *Proceedings of the IEEE Symposium on Information Visualization (INFOVIS'95)*, pages 51–58. IEEE Computer Society Press, 1995.

[146] Kent Wittenburg and Eric Sigman. Integration of browsing, searching, and filtering in an applet for web information access. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI'97)*. ACM, 1997.

[147] A. Woodruff, A. Faulring, R. Rosenholtz, J. Morrison, and P. Pirolli. Using thumbnails to search the Web. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI'01)*, pages 198–205, 2001.

[148] K.-P. Yee, K. Swearingen, K. Li, and M.A. Hearst. Faceted metadata for image search and browsing. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI'03)*, pages 401–408. ACM New York, NY, USA, 2003.

[149] Degi Young and Ben Shneiderman. A graphical filter/flow model for Boolean queries: An implementation and experiment. *Journal of the American Society for Information Science*, 44(6):327–339, July 1993.

[150] O. Zamir and O. Etzioni. Grouper: A Dynamic Clustering Interface to Web Search Results. *Proceedings of the 8th International Conference on World Wide Web (WWW'99)*, 31(11-16):1361–1374, 1999.