

From Keyword Search to Exploration: Designing Future Search Interfaces for the Web

By Max L. Wilson, Bill Kules,
m. c. schraefel and Ben Shneiderman

Contents

1 Introduction	4
1.1 A Challenge Faced by Designers of Future Search Systems	9
1.2 A Taxonomy to Overcome the Challenge	11
1.3 The Scope of Our Approach	12
2 A Model of Search	14
2.1 A Context Model of Search	15
2.2 Information Retrieval	16
2.3 Information Seeking	17
2.4 Work Context	19
2.5 Summary	20
3 Survey of Search Systems	21
3.1 Adding Classifications	22
3.2 Result Organization	41

3.3	Additional Functions	57
3.4	Summary	64
4	Evaluations of Search Systems	65
4.1	Information Retrieval Evaluations	65
4.2	Information-Seeking Evaluations	67
4.3	Work-Context Evaluations	71
4.4	Summary	74
5	Taxonomy of Search Visualizations	75
5.1	The Taxonomy	75
5.2	Using the Taxonomy	79
6	Conclusions	81
	Acknowledgments	83
	References	84

From Keyword Search to Exploration: Designing Future Search Interfaces for the Web

Max L. Wilson¹, Bill Kules²,
m. c. schraefel³ and Ben Shneiderman⁴

¹ *Intelligence, Agents, Multimedia (IAM) Group, School of Electronics and Computer Science, University of Southampton, Southampton, SO17 1BJ, UK, mlw05r@ecs.soton.ac.uk*

² *School of Library and Information Science, The Catholic University of America, Washington, DC 20064, USA, kules@cua.edu*

³ *Intelligence, Agents, Multimedia (IAM) Group, School of Electronics and Computer Science, University of Southampton, Southampton, SO17 1BJ, UK, mc+ft@ecs.soton.ac.uk*

⁴ *Department of Computer Science and Human-Computer Interaction Laboratory, University of Maryland at College Park, College Park, MD 20742, ben@cs.umd.edu*

Abstract

This monograph is directed at researchers and developers who are designing the next generation of web search user interfaces, by focusing on the techniques and visualizations that allow users to interact with and have control over their findings. Search is one of the keys to the Web's success. The elegant way in which search results are returned has been well researched and is usually remarkably effective. However,

the body of work produced by decades of research into information retrieval continues to grow rapidly and so it has become hard to synthesize the current state-of-the-art to produce a search interface that is both highly functional, but not cluttered and distracting. Further, recent work has shown that there is substantial room for improving the support provided to users who are exhibiting more exploratory forms of search, including when users may need to learn, discover, and understand novel or complex topics. Overall, there is a recognized need for search systems to provide effective user experiences that do more than simply return results.

With the aim of producing more effective search interfaces, human computer interaction researchers and web designers have been developing novel interactions and features that enable users to conveniently visualize, parse, manipulate, and organize their Web search results. For instance, while a simple set of results may produce specific information (e.g., the capital of Peru), other methods may let users see and explore the *contexts* of their requests for information (more about the country, city, and nearby attractions), or the properties that associate groups of information assets (grouping hotels, restaurants, and attractions by their type, district, or price). Other techniques support information-seeking processes that may last weeks or months or may even require collaboration between multiple searchers. The choice of relevant result visualization strategies in new search systems should reflect the searchers and the higher-level information needs that motivate their searches. These examples provide further motivation for supporting designers, who are challenged to synthesize and understand the breadth of advances in search, so that they can determine the benefits of varied strategies and apply them appropriately to build better systems.

To support researchers and designers in synthesizing and understanding the advances in search, this monograph offers a structured means to think about web search result visualization, based on an inclusive model of search that integrates information retrieval, information seeking and a higher-level context of tasks and goals. We examine each of these levels of search in a survey of advances in browsers and related tools by defining search-related cognitive processes and

analyzing innovative design approaches. We then discuss evaluations at each of these levels of search, presenting significant results and identifying both the traditional and novel means used to produce them. Based on this examination, we propose a taxonomy of search result visualization techniques that can be used to identify gaps for future research and as a reference for designers of next generation web search systems.

1

Introduction

This monograph is for designers thinking about how they might enhance the experience of discovering, exploring and putting to work information they can access over the Web. This monograph is also for researchers who may be interested in how search interaction approaches have developed over the past decade. In both cases, a fundamental question is at play: what else users could possibly need besides Google to search the Web? That's a fair question, and readers of this survey may have the same question. So to tackle that head on, let us agree: Google is really good. For what it does.

Our monograph considers the approaches for exploring information spaces that Google's elegant keyword search cannot do. Over the past decade, research on alternative search paradigms has emphasized Web front ends on single or unified databases. Such work is productive when the designer has the luxury of working with well-curated documents from a single source, but the elegant visualization or well-tailored faceted browser may not scale to the size and diversity of Web-based information, links, scientific data sets, personal digital photos, creative videos, music, animations, and more. So why does keyword search scale? Because (a) huge resources have been thrown at the problem

and (b) textual data have a satisfying and compliant order to it. We hope this monograph shows that the research and evaluations of alternative approaches to data exploration for knowledge building are the best preparation we have for the next generation Web, the Web of Linked Data.

In a little more than a decade the Web has become the default global repository for information. Many factors have contributed to this remarkable result, from the success of the robust technologies that enable its networked connections, to the commercialization of the backbone that enticed business to support and utilize it, to the ease with which ordinary citizens can publish information to it. But perhaps the key technology that took the Web from a useful supplement of current information practice to become the default communication medium is search. Web search, as provided by Google, Microsoft, Yahoo, etc., enables users to find the information they want via the simplest of interaction paradigms: type some keywords into a box and get back an informative result list that is ranked, as if by magic, so that the first results most likely match what we're trying to find.

Search engines automated what was initially a community-based and commercially coordinated Easter egg hunt: category systems were proposed and documents as they were found either by humans recommending them to such sites, or discovered by human trawlers and some early web crawlers, were assigned to categories. The Web was set up like a giant Yellow Pages. Further, before these nascent directories, the Web was explored by the link. As recently as 2004, surfing the Web was still a common trope for browsing the Web, following from link to link, from one site to another. Not unlike blogs today, web sites might publish best new finds on a topic, and away one would go. Only five years later, who "surfs" or presumes to browse "the Web"? It has grown beyond that scale of the surfable, with its pockets of Dark Web and ice caps of Public Web, where so much more than is indexed by search engines is below the visible water line of documents. This growth is itself related to the existence of search: because it **can** be found by search, rather than relying on recommendations alone, it is worth publishing on the Web; indeed it is necessary to publish on the Web. Because conversely, if it cannot be found on the Web, does it exist?

Search as embodied by the text box and keyword has framed our understanding of what the Web is [145]. It has become so ubiquitously associated with the Web that it is difficult to find a web browser that does not have as a default tool, a keyword search box built into upper right of the browser window, right next to the box for the location. In some cases the URL address bar is already polymorphic: acting as either address area or search bar. The prominence of web search based on the fundamental efficacy of keyword search makes it difficult to imagine what an entire monograph on search interaction may be about. It turns out that this elegant paradigm is especially effective for the 1-min search — find the address for Chez Panisse, get a biography of Madame Curie, or locate the home page for a local car rental. But many users have come to the Web for substantive research that takes hours or weeks — find all the songs written about poverty in the 1930s, prove that there are no patents that cover my emerging innovation, or locate the best legal precedents for my brief.

A second motivator for new search strategies is that the next generation web will offer fresh possibilities that go well beyond finding documents by way of keyword search. Hall and O’Hara [69] stress that what we know as the Web today is the Document Web, and not the Web of Linked Data that is imminently upon us. The older Document Web is about the information, readable by us, written by us, and framed for our use. It is this very human-readable orientation of the Web and it is the presentation technologies in the browser that have enabled keyword search engines to become so very good: the words *in* the documents are all a search engine has to go on to find appropriate results. It is because the search engine is searching *in* documents that we get a list of documents back: we may only want a sentence in the middle of a document, but we get the whole thing (the document) back.

By contrast, in the newer Web of Linked Data, often called the Semantic Web, the idea is to give the search engine designers more to work with than making best guesses about what results to return based on keyword frequency and number of pages linked to a document. Imagine if instead of a list of results, the machine simply returned “the answer”? Some queries have specific answers: “mc’s phone number at work” or “British obesity rate in 2009?” There may be several sources

for this information, but if they concur, why not just present that result, and give the six places the result occurs? Right now, this kind of search result is not feasible because the Web for the most part does not have consistent tags to indicate what is a phone number or where to find obesity rates. Search engines check if a term requested maps to the term in a document, and does very effective ranking very fast. That is of course an oversimplification of all the sophistication to make that experience as fast and accurate as it appears to be.

The goal of the Web of Linked Data is to have information about information available so that not only can designers provide better, clearer answers to such simple queries about phone numbers and statistics, but also users can resolve questions that are difficult to handle without metadata. Some researchers are conducting intriguing research that attempts to create this metadata automatically to derive semantics from the documents themselves. In this monograph we are less concerned with how metadata becomes available. We are concerned with the question of what designers can do with it, once it exists.

While the power of Semantic Web of Linked Data is that it can enhance large diverse, unorganized, and heterogeneous datasets, the unique affordances also challenge our assumptions about how we access information [176]. As the links between data can be numerous, endless, and of any granularity, the assumptions about carefully structured classifications, for example, breakdown. Similarly, while web searches are typically for web pages, it is not clear whether searching at the data level should return any object [21], specific types of objects [146], object relationships [21, 76], portions of RDF [47], entire ontologies [2, 63], and so on.

Further, as the work on semantically linked data has separated the data from presentation, designers and users are able to represent the data however they like [21]. The flipside, however is that someone, either the interface designer or the end user, has to decide how to represent the data. In summary, the freedom enabled by semantically organized data sets has in turn broadened the options and increased the number of decisions that designers and end users have to make. Recent work has shown, however, that increasing numbers of options can make designers and users feel less confident in their decisions, and less happy

with their results [130, 149], rather than making them feel empowered. What effect, then, does this have on confidence during search interface design, given that designers and users now have more freedom.

These issues are becoming national policy issues, especially as the United States, United Kingdom, and other governments intend to release increasing amounts of information onto the Web as data with sufficient metadata to support more automatic analysis. Metadata tags will indicate that a number represents the reported cases of obesity at a given time and place. The number also has a source and date of creation associated with it so users can verify accuracy and timeliness. It is not simply a number in a document, instead it comes with well-associated meanings.

As soon as this information is available new opportunities for representation beyond document lists become possible. Users will be able to see quickly for themselves: are obesity levels rising at the rates presented in the media, or, by mapping these data from several sources, are they too conservative or aggressive? Imagine being able to look at such sources to ask these kinds of questions with the same facility as we use keyword search now. Now *that* is an interaction challenge.

Designers already offer such representations on smaller than Web scale information sources; that is what most of the literature we will review considers. In that sense we have some preparation for what is to come. But there are also entirely new interaction challenges. There will be many data sources from many places that map to given meanings, like obesity statistics. How can these disparate sources be gathered, inspected, and explored?

Right now, we are on the cusp of a transition from the Document Web to the Document/Data Web. It is an incredible time to be interested in designing or researching how to engage with these immense volumes of data. Are designers and researchers ready for this transition? The findings presented here may act as guideposts for this near future. We may also look back in ten years at these nascent efforts to imagine exploring data at scale and say either that was clever or that was naïve. It will be more than intriguing to see what principles remain, and what have yet to be imagined. In the meantime, we look back in order to leap ahead.

In the remainder of this monograph, Section 2 identifies and explains a model of search that is used to structure the discussion in the following sections and forms the basis of the taxonomy of advances in interactive search. In short, the model describes search in increasing layers of context, from simple retrieval, to broader strategies, and to the wider tasks that motivate them. Section 3 identifies specific advances in interactive search, and prominent examples of their use, organized by the different layers of the model to which they apply. Section 4 explains how advances at each layer of the model have been evaluated. Section 5 then presents the final taxonomy, and identifies areas of relatively little research as potential focal points for future research. For search interface designers, the taxonomy provides a list of potential features to include in designs, describing (a) how they support users, (b) how their support has been evaluated, and (c) how prevalent they have become on the Web.

1.1 A Challenge Faced by Designers of Future Search Systems

Understanding how search interfaces and visualization affect searcher success is a hard challenge, and cannot be as easily measured as speed, document similarity, and result accuracy. In the early 1970s, Cooper [43] suggested that instead of speed metrics, search evaluations should be based on subjective user satisfaction with the results. Later, Robertson and Hancock-Beaulieu [143] noted that the recent, at the time, revolutions of IR research had begun to focus more on users and less on systems. Even more recently, though, researchers have identified just how inadequate the familiar keyword search paradigms, provided by environments such as Google and Bing¹ (Microsoft's search engine), might be for users who need to do more than just find a website that answers a factual question.

The recent focus on these more exploratory forms of search, known as Exploratory Search [172, 174], has identified some search scenarios that require much more diverse searching strategies, including when the

¹<http://www.bing.com>

users are (a) unfamiliar with a domain and its terminology, (b) unfamiliar with a system and its capabilities, or (c) unfamiliar with the full detail of their task or goal. Experts may also conduct demanding searches such as those needing to do: (a) comprehensive searches to find every relevant record (Legal, patent, medical), (b) negation searches to prove the absence of relevant work (e.g., patent, pharmaceuticals), (c) exception searches to find outlier documents (that take a different or contradictory point of view than commonly held), or (d) bridging searches that connect two disparate fields of study [164]. Exploratory search scenarios are characterized by needs that are “open-ended, persistent, and multifaceted, and information-seeking processes that are opportunistic, iterative, and multitactical” [174]. In each case advances in search need to do more than simply improve the matching of results to terms entered into a single box. Even in the late 1980s, Motro [126] designed the VAGUE interface based on the notion that often, when limited to simple keyword interfaces, users submit numerous evolutions of an original vague query in order to figure out which terms are going to produce all, or even just part, of their desired information.

In many cases, searching involves a range of tactics and techniques, rather than simply submitting a query and seeing a list of matching results. As part of the special issue on Exploratory Search, Marchionini [116] identified, although not exhaustively, a series of strategies that users may often need to employ to achieve their goals, such as comparing, synthesizing, and evaluating. MacKay and Watters [114] present a diary study documenting examples of search tasks that span multiple search sessions, where users return days later to continue on tasks such as job seeking or house hunting. Similarly, Morris [125] has documented the breadth of occasions where users clearly collaborate with family and colleagues on tasks such as holiday planning and team work projects. It is plain to see that a search interface needs to provide more than a simple keyword search form to support users in applying such strategies.

The recognition that there is more to search than basic Information Retrieval has led to many extensions and alternatives to the keyword search paradigm. An example extension is to cluster the results into groups that share attributes [190]. Alternatively, faceted browsing

[75, 168] provides meaningful or significant attributes of web pages to users at the beginning, so that they do not even have to think of words to put in the keyword search box. This can be especially useful in the occasions where people are unfamiliar with a domain and its terminology, for example. These, and many other advances, have much in common and while they each have specific advantages, it is not clear that including them all would provide a stronger search interface. Instead, designers now have the challenge of deciding: (a) what types of search strategies should be supported, if not all of them, and (b) which new features to include in order to support them. This challenge is particularly difficult, when so many advances have been proposed, each with different benefits, and when the benefits of each advance have often been shown independent of others.

1.2 A Taxonomy to Overcome the Challenge

The goal of this monograph is to support designers with this challenge by building a taxonomy of advances in the field of search that can be used as common ground when choosing which features to include within future search interfaces. To build the taxonomy we:

- (1) identify a model, produced by theory, which covers the full breadth of search from context and tasks, down to specific actions and results (Section 2);
- (2) summarize the specific advances in interactive search (Section 3);
- (3) discuss the way that these interactive search advances have been evaluated (Section 4), in accordance with the model of search presented in Section 2; and
- (4) present a taxonomy (in Section 5) of the search advances (from Section 3) that takes into account the type of search supported (from Section 2), how the advances have been evaluated (Section 4) and how prevalent they are on the Web.

Producing a structured and consistent taxonomy allows us to compare advances in interactive search, with two benefits. First, the taxonomy can be a reference for designers. The latter half of Section 5

describes a detailed process that designers can apply to systematically decide which features to include in their search designs. As the taxonomy includes advances in the field, their benefits, and how they have been evaluated, a designer can quickly compare options and choose appropriate features for their new interface. Second, the taxonomy can be used by academics to identify areas that require further study and contextualize future advances in the field.

1.3 The Scope of Our Approach

When defining a means of categorizing and communicating existing and on-going research, it is important to define the scope of our approach, so that it is correctly used. Here we specifically bound the content of this monograph in two areas: what we mean by search and what we mean by the Web.

1.3.1 What We Mean by Search

So far, this monograph has used the terms: search, seeking, and Information Retrieval interchangeably. For the rest of the monograph, however, we intend to follow a specific set of terminology, which is defined carefully in the model discussed in Section 2. Information Retrieval is perhaps the most well studied, and so most well-defined term used to describe searching. Typically, Information Retrieval refers to the paradigm where users enter a keyword into a system, which responds by returning the results that are most relevant to the keywords used. This monograph covers a much broader view of search than simply Information Retrieval. Information Seeking is another common term used to describe people's searching behavior, including activities such as searching, browsing, and navigating. Again, however, in this monograph we use the word search in a broader sense than Information Seeking. The model described in Section 2 defines search as the set of activities that take users from identifying a problem all the way to achieving their goals, which will, at times, involve Information Seeking, which in turn, may include Information Retrieval.

1.3.2 What We Mean by the Web

The “indexable web”, that is, the portion of the web indexed by major search engines, was estimated at 11.5-billion pages in 2005 [67], with Google reporting that they surpassed the 1-trillion mark in July 2008.² One major characteristic of the Web, therefore, is scale. In this monograph, however, search systems are discussed that search both the whole web, and certain domains within the Web and so scale is not always a primary concern for design. A more indicative and remarkable characteristic is the heterogeneity of the contents. The Web contains a variety of data and documents. Documents may be in plain text, HTML, XML, PDF, Rich Text Format (RTF), Microsoft Word (.DOC), spreadsheets, and a multitude of specialized or proprietary formats, including microformats [6]. Multiple forms of media, including still images, audio, and video, are widely available and indexed by general purpose as well as specialized search engines. These documents vary from highly structured databases, to semi-structured web pages, to unstructured text. Again, however, while some web search engines focus on the entire heterogeneous content of the Web, others focus on specific and bounded domains within the Web. In these known and bounded conditions, the format of documents is often known and fixed, and so is not always a concern for all web-based search systems.

In summary, as web search systems are discussed, some are limited by the diversity of online material, and the design of others is motivated by unique features of web collections. Both are important areas that sometimes share concerns but often differ significantly in the challenges they present during the design of search interfaces. An e-commerce site might, for example, try to support searchers with the categorization, price, and availability of their products. Such product-related attributes are not a concern for general Web search, which may include product results, reviews, specifications, and informational documents. Finally, it is important to remember that as users engage in search, they may be moving between the entire web and known collections within it, in order to achieve their goals.

²<http://googleblog.blogspot.com/2008/07/we-knew-web-was-big.html>.

2

A Model of Search

In order to survey search interface advances, we need to establish a framework and a vocabulary that allows us to organize, describe, and analyze the value of their contributions. This section describes a model of search [84] which grounds the remainder of the monograph by conveying both a structure that defines search and a set of relevant terminology. While many models of search exist, as discussed further below, the selected model was chosen because of its holistic perspective, considering different granularities of context determined by “a person’s task, its phase, and situation”, while still including important finite concepts such as topical relevance. The key benefit of this holistic view is that it maintains an understanding of the situational nature of information and that search activities are performed in the context of a larger work task, and even cultural influences. Therefore, the discussion of advances in search interaction, visualization, and Human Computer Interaction, must be set in the broader context of a human’s task. The model described in this section will be referenced throughout Sections 3 and 4, and will be used to classify search interfaces within the taxonomy described in Section 5.

2.1 A Context Model of Search

Figure 2.1 shows the model of searching contexts, produced by Jarvelin and Ingwersen [84], that provides the framework we will be referring to for the remainder of this monograph. Search can be modeled as a hierarchy of goals and tasks. Each task provides the goal and serves as the context for its subsidiary task(s). Figure 2.1 makes multiple levels of context explicit: socio-organizational and cultural, work task, seeking, and retrieval. Information Retrieval, as the smallest granule in the model, represents the action most often embodied by Keyword Search, where users are trying to find an often known nugget of information, such as the price of a particular server. Searchers may often find themselves performing a series of Information Retrieval actions as part of a broader Information-seeking task, such as trying to find the best possible server given a specific budget. Any Information-seeking task, however, is set within a larger Work Task, such as being asked to procure

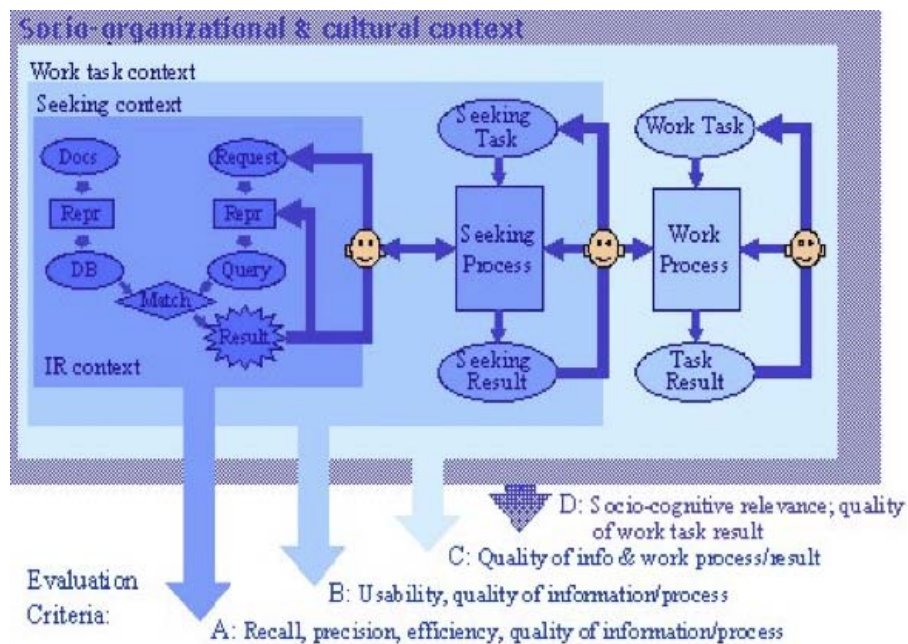


Fig. 2.1 Searching behavior is made up of multiple layered contexts, where simple information retrieval is the narrowly focused; figure from Järvelin and Ingwersen [84].

resources for a forthcoming project. The result of a Work Task will contain the product of one or more Information-seeking tasks. Finally, every Work Task sits within a much broader Socio-Organizational and Cultural context. The requirements and the importance of meeting them, for example, will be different when buying a server for a large organization, a safety critical organization (such as a hospital) and a small start-up business. Clearly, a home/personal context will have different demands than a corporate professional environment, and success in each will be under different criteria.

2.2 Information Retrieval

Within the information retrieval (IR) context, the searcher's goal is focused on finding documents, document sub-elements, summaries, or surrogates that are relevant to a query. This may be an iterative process, with human feedback, but it usually is limited to a single session. Typical IR tasks involve finding documents with terms that match terms presented by the searcher, or finding relevant facts or resources related to a query. Typically, within each IR task, the searcher formulates queries, examines results, and selects individual documents to view. As a result of examining search results and viewing documents, searchers gather information to help satisfy their immediate information-seeking problem and eventually the higher-level information need. The common element of all IR tasks as defined in this web-focused monograph is the query-result-evaluation cycle conducted over a collection with the "unit" of information being the document, a sub-element of the document, a summary, or a document surrogate. In the server purchasing scenario used above, the document being sought may be the information page on an e-commerce website, but may also be a picture, or downloadable PDF of its specification.

Web search engines like Google support one of the more common IR tasks on the Web: searching for web pages that match a given set of query terms. Services such as Ask.com, however, also attempt to provide fact-oriented answers to natural language queries, while Google will also answer specific conversions of mathematical calculations if the queries are constructed in certain formats. Keyword search is not the

only example of Information Retrieval on the Web, however, with tools like Flamenco and Clusty, each discussed in more detail in Section 3, provide hyperlinked terms that allow searchers to narrow their results and browse relationships between documents.

Evaluation measures traditionally include precision, recall, and their variations, which simply assess how relevant a document is to the given query. Consequently an IR system can be tested simply by whether it returns the most appropriate documents to a given query. Although the concept of relevance has multiple aspects, at the IR level, topical relevance is the typical aspect considered, that is, how relevant the documents are to the topic expressed by the query, and most systems focus on returning the most relevant documents rather than all of the relevant documents.

2.3 Information Seeking

The objective of the information-seeking (IS) task is to satisfy a perceived information need or problem [115, 117, 151]. Often, searchers undertake one or more IR tasks as part of a larger information-seeking (IS) task, although it is possible for a simple need and IS task to be achieved with a single IR task. At the IS level, searchers make strategic decisions about where, how, and even whether to find information related to their information needs. They may adopt an analytical strategy to decide whether to use an IR system (a specific website or searching service). They may also adopt a browsing strategy, where, for example, they start from a known point (perhaps the results of an IR query) and follow successive links to locate other related documents. The documents found while browsing or returned by an IR task will, as part of the IS task, be examined to extract information and synthesize it into a solution to the information need.

While the Web provides an environment for many IR tasks, and, consequently, many IS subtasks, searchers may also consult non-IR systems as well as other resources such as printed material, colleagues or friends, in order to achieve their goal. In fact, people do not always choose to seek information. They may prefer to exhibit alternative Information Behaviors and avoid information that is discomforting or

troublesome or which is judged not worth the effort [124]. They may also be biased in what information they seek, looking for information that supports preconceived ideas or opinions [119].

Information-seeking tasks have previously been structured as, for example, linear sequence of stages [101] or hierarchical decompositions of tasks [23]. Each of these tasks requires selecting a source, and then engaging in one or more information retrieval tasks which satisfy a portion of the overall need. From the perspective of an organization, Choo et al., [39] developed a behavioral model of organizational information seeking on the Web by integrating Ellis' [57] six stages of information seeking (starting, chaining, browsing, differentiating, monitoring, and extracting) with Aguilar's [1] four modes of scanning (undirected viewing, conditioned viewing, informal search, and formal search). Each of these tasks helps to satisfy part of an organization's information needs.

The hyperlink capability of the Web provides support for a browsing strategy. When browsing, each new piece of information that is gathered can provide new ideas, suggest new directions, and change the nature of the information need [13]. This leads to behavior that Bates refers to as berrypicking, reflecting the incremental collection of pieces of information that, as a whole, help to satisfy the information need. The choices made at each step are guided by the seeker's assessment of what is most likely to produce useful answers as they forage for information [134]. In an environment like the Web, this assessment is based on cues such as hyperlink text that provide an information scent which help users make cost/benefit trade-offs in their choices. In line with research suggesting that search is made up of sequences that are affected by the discovered information, Belkin et al. [18] created a set of "scripts" that describe the typical paths taken by 16 different types of users, including switch points between them for when discoveries change their user type.

Systems that attempt to provide support of IS tasks typically provide functionality beyond the query-result-evaluation cycle supported by IR systems. This may include search history mechanisms to support information seeking over multiple sessions, or mechanisms such as tagging to collect a set of documents that are each relevant to the larger information need. They may also provide overviews of a collection using textual or graphical methods. Individually, such UI features

may be tailored to support specific elements of the IS stages (e.g., topic exploration when writing a paper or monitoring of web resources).

The evaluation of interfaces that support information-seeking tasks typically involves assessing the quality of information acquired by users relative to the information need provided. IS tasks, such as purchasing a server, are usually assessed by the support the interface provided to users while carrying out the task, and the judged accuracy of the final decision, given the initial requirements. Hearst [72] provides a thorough review of information-seeking user interfaces and their evaluation.

2.4 Work Context

The perceived information need that motivates an IS task is itself motivated and initiated by a higher-level work task or context [29, 30, 84] or personally motivated goal [92]. The process is initiated when a searcher recognizes an information need (or is instructed to investigate one) based on an organizational or personal need [29, 30, 115].

Work tasks are situated in the context of work organization and reflect organizational culture and social norms, as well as organizational resources and constraints. As such, they constrain or guide the IS tasks. For example, the work context provides important information about the domain of information relevant to an information need. It defines and constrains the resources available to satisfy the need. Immediate access to reference books, libraries, electronic resources (e.g., the general web or specialized online databases), and human experts affect the strategic decisions searchers make at the work-context level.

Systems that support work-context tasks may provide specialized functions or support strategies for the work domain. Certain working domains, such as medicine, law, and even academia, have well-established work-context tasks and procedures for information seeking in support of those tasks. For example, in the medical field, studies have examined how physicians and healthcare searchers search for information [22, 61]. Consequently, services can provide mechanisms, for example, that identify prior art or key documents on a certain topic to support the work-context task of writing a paper. To support the learning of work-context processes, search systems may provide access

to tutorials or communities of practice related to the work task. They may also provide domain-specific frameworks for making sense of, integrating, and synthesizing information. For example, a system to support the development of teaching materials based on an oral history collection may provide a lesson plan tool that supports organizing and annotating retrieved information in a lesson plan format with links from items in the lesson plan back to retrieved elements of the collection.

Evaluation of systems in a work context usually focuses on the achievement of users in realistic scenarios, such as the grading of school or college essays, or, in the extreme, the success of organizations at procuring academic funding.

2.5 Summary

In the previous subsections, we have described the three key levels of searching context: information retrieval, information seeking, and work contexts. The last of these is also surrounded and influenced by the environment and social context that the work task is being carried out within. Each of these contexts provides a useful lens for understanding the benefits and contributions of novel user interfaces for search. When a new technique is designed, built, and tested, the work is usually motivated by a problematic scenario and a hypothesis that it can be overcome. This scenario might be carefully defined and IR focused, such that users cannot easily express a particular query or find the correct result. Similarly, the scenario and hypothesis may be much broader, as to support a work task like planning a vacation or researching for a report. In the following sections, we discuss the design and evaluation of recent interface contributions, by considering how they have been designed and tested to support searchers at different levels of context: IR, IS, and WC.

3

Survey of Search Systems

The aim of this section is to review many of the search visualization and exploration techniques that have developed and studied. Further, the aim is to demonstrate the diversity and range of available techniques, rather than produce a complete catalog like some much larger surveys [72]. The techniques described in this section, which are classified in the taxonomy in Section 5, are presented according to the following structure. First, techniques that have made use of the advantages of enriched metadata are discussed in Section 3.1, which has afforded a change in the way that searchers can interact and control the presentation of results. Second, Section 3.2 describes the varying approaches to directly organizing and presenting results. Finally, Section 3.3 addresses some alternative functionality that has enhanced visualizations, such as animation and use of alternative senses. Each technique described below is briefly discussed with the contexts of search described in Section 2. Further, while here many techniques are listed, the taxonomy described in Section 5 provides the means to compare their aims, strengths, and weaknesses, in terms of the context of information seeking supported and the way they have been evaluated, according to Section 4.

3.1 Adding Classifications

One of the main streams of research for enhancing search environments has been to use annotations, or classifications, to the documents or collections. Two challenges in adding classifications are the increasing scale of collections and the associated cost of annotating each document. We review four common approaches to adding classification to collections and overcoming the challenges: hierarchical classifications, faceted classifications, automated clustering, and social classifications.

3.1.1 Hierarchical Classifications

One early research project, the SuperBook interface, showed the benefits of using categories in a document collection, by organizing search results within a book according to the text's table of contents; here the book is the collection and the pages of the book are the documents. An evaluation found that it expedited simple IR tasks and improved accuracy, both by 25% [55]. This categorization approach has been used in many cases and has shown success in both fixed and uncontrolled collections, however, the usual approach for the latter is to model the first and allow document owners to assign their own documents into the hierarchy.

An example of a fixed and managed data set may be the genre-based classification of music in Amazon's¹ online music store, where every CD is assigned to one or more categories of music, including Pop, Rock, and Indie. Here there is incentive for Amazon, as the collection owner, to annotate the collection in this way to make it easier for their clients to find the music they want and, subsequently, encourage sales. Allen [5] investigated two such digital library interfaces, the Dewey Decimal System and the ACM Computer Reviews system, and showed that both used hierarchical classification effectively for organizing the collections.

In the same paper, Allen discusses the potential for use of Internet-wide information samples. Examples of hierarchically classified structures for the Web, as an ever increasing and unmanaged set

¹<http://www.amazon.com>.

of documents, are Google Directory² and Yahoo Directory.³ In both of these directories, the individuals that own or manage documents on the Web can submit their websites for inclusion under different parts of the directory. This approach has been popular in the past but has suffered where document owners do not submit their websites for inclusion.

In an attempt to remove the annotation cost and limitation of classification systems, Kules et al., [102] took a simple result set of federal agency/department reports, by mapping a set of URL prefixes to a known finite list. This approach took a known pattern from the data and used it to effectively categorize search results by government agency and department. In a more advanced approach, Chen and Dumais [36] showed that machine learning techniques can be applied to known categorized data items to automatically assign new web documents into categories. Using the documents that already exist in the LookSmart⁴ directory as a training-set, the machine-learning algorithm successfully and automatically categorized the results of a Yahoo search. The Support Vector Machine (SVM) algorithm [87] achieved 70% accuracy with the categorization provided by human participants, where the remaining 30% included partially matching annotations. The user study showed a strong preference for the added categorization provided by the process. Other approaches to automation exist and are mainly described in the automatic clustering section below.

The benefits of applying hierarchical categorizations have been proven numerous times in research. For question answering tasks, Drori and Alon have shown that search results augmented with category labels produced faster performance and were preferred over results without category labels [49]. Dumais et al. [50] also studied the effect of grouping search results by a two-level category hierarchy and found that grouping by a well-defined classification speeds user retrieval of documents.

In a similar approach to the book-search mentioned above, the Cha-Cha system organizes intranet search results by an automatically generated website overview (Figure 3.1). It reflects the underlying

²<http://directory.google.com>.

³<http://dir.yahoo.com>.

⁴<http://www.looksmart.com>.

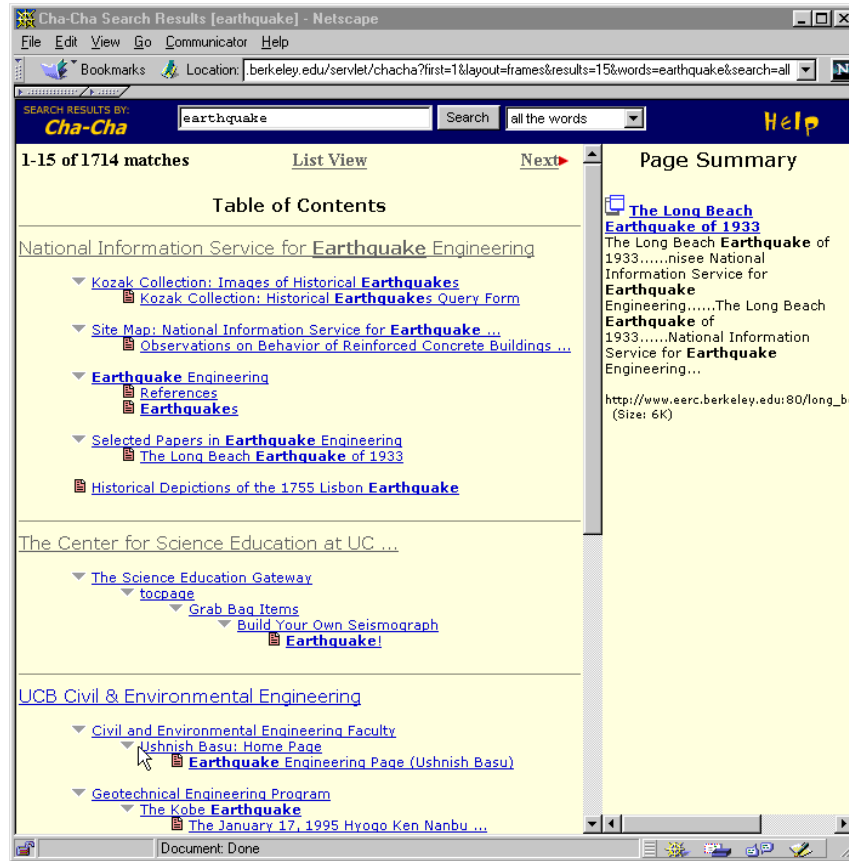


Fig. 3.1 The Cha-Cha system organized intranet search results by an automatically generated website overview.

structure of the Website, using the shortest path from the root to each document to dynamically generate a hierarchy for search results. Preliminary evaluations were mixed, but promising, particularly for what users considered “hard-to-find information” [38]. The WebTOC system (Figure 3.2) provides a table of contents visualization that supports search within a website, although no evaluation of its search capability has been reported [128]. WebTOC displays an expandable/collapsible outliner (similar to a tree widget), with embedded colored histograms showing quantitative variables such as size or number of documents under the branch.

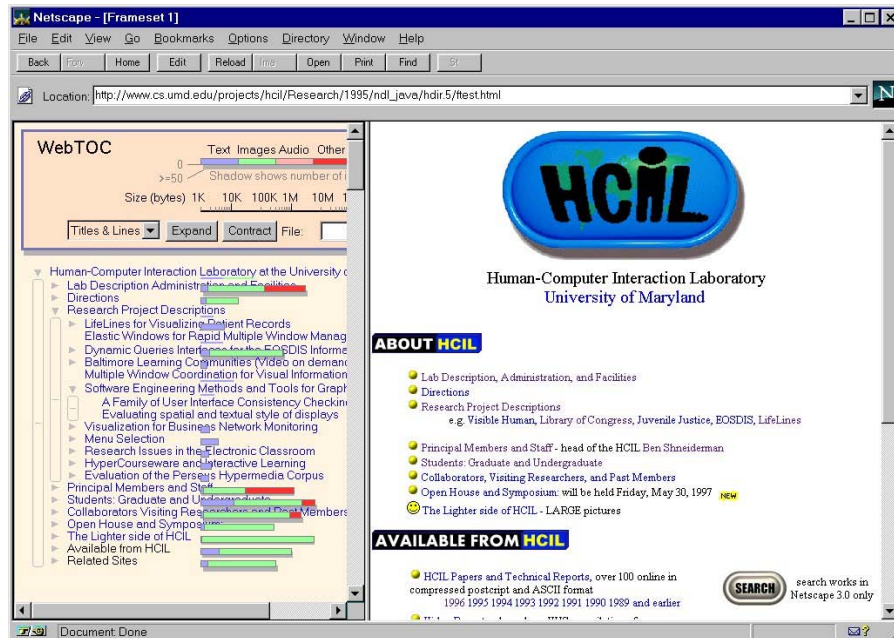


Fig. 3.2 The WebTOC system provides a table of contents visualization that supports search within a website.

Hierarchical classifications, as demonstrated by the examples discussed above, are largely designed to organize results into groups. While this allows users to perform some additional browsing tactics [12], such as going to a parent or child areas of the classification, the main benefits have been studied at the IR level. The studies performed by Allen [5] and Drori and Alon [49] show users performing faster in basic information retrieval tasks, without loss of accuracy. Faceted classifications discussed below, however, extend the notion of hierarchical classifications to support additional information-seeking behavior. Consequently, studies of faceted classification systems have more often focused on information-seeking behavior.

3.1.2 Faceted Classifications

Another approach that helps users find documents based on multiple orthogonal categorizations, such as thematic, temporal, and

geographic, builds on the traditional library science method called faceted classification. Restaurants, for example, can be classified by facets such as by location, price, food type, customer ratings, and size, with around 2–20 attribute values per facet (although some facets have many more values). Early designs, one of which was called query previews, also showed the number of documents having each attribute value. Query previews were updated with new values so as to prevent users from submitting searches that would produce zero results [48, 135, 165]. Faceted classification allows users to apply relevant constraints on their search in the context of the current problem and existing knowledge. For example, if users know their own budget and a minimum specification required for a computer, then they can apply constraints in these facets separately, and vary them individually (to check, for example, the effect of increasing the budget slightly). After applying their constraints they can then see all the relevant computers and possibly choose between them based upon the facets that remain unused.

Many systems have applied this approach in some way, where a single classification system has not been expressive enough for the documents in a collection. For example, before prototyping a faceted system,⁵ eBay already allowed users to apply constraints such as price, color, and size; the available constraints depended, of course, on the type of object being sought. Whereas eBay indexes its own auctions, websites such as shopping.com and Google Product Search⁶ provide a faceted search over many shopping sites from the whole web.

Flamenco⁷ (Figure 3.3) is a clear example of the features provided by faceted search using multiple hierarchical facets. Providing interfaces to fixed collections, including art, architecture, and tobacco documents, Flamenco presents faceted hierarchies to produce menus of choices for navigational searching [188]. A selection made in any facet is added to a list of constraints that make it clear to users what is forming the list of results that are being shown.

⁵ <http://express.ebay.com>.

⁶ <http://google.com/products>.

⁷ <http://flamenco.berkeley.edu/>.

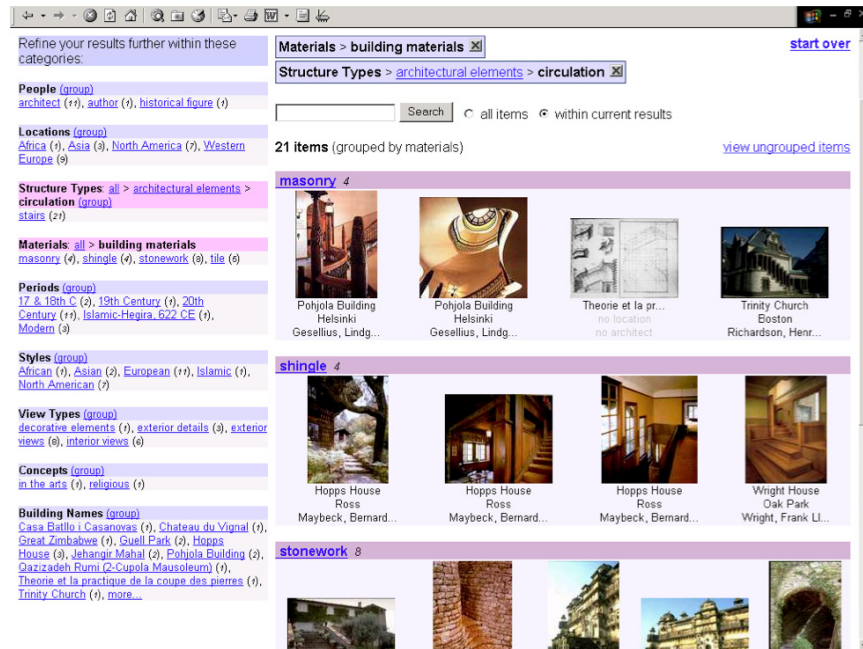


Fig. 3.3 The Flamenco interface permits users to navigate by selecting from multiple facets. In this example, the displayed images have been filtered by specifying values in two facets (Materials and Structure Types). The matching images are grouped by subcategories of the Materials facet's selected Building Materials category.

A usability study compared the Flamenco interface to a keyword search interface for an art and architecture collection for structured and open-ended exploratory tasks [188]. With Flamenco, users were more successful at finding relevant images (for the structured tasks) and reported higher subjective measures (for both the structured and exploratory tasks). The exploratory tasks were evaluated using subjective measures, because there was no (single) correct answer and the goal was not necessarily to optimize a quantitative measure such as task duration.

The success seen by Flamenco, having provided faceted classifications to assist search, has been used in many commercial and academic projects. Endeca⁸ is a commercial company that provides faceted

⁸<http://www.endeca.com>.

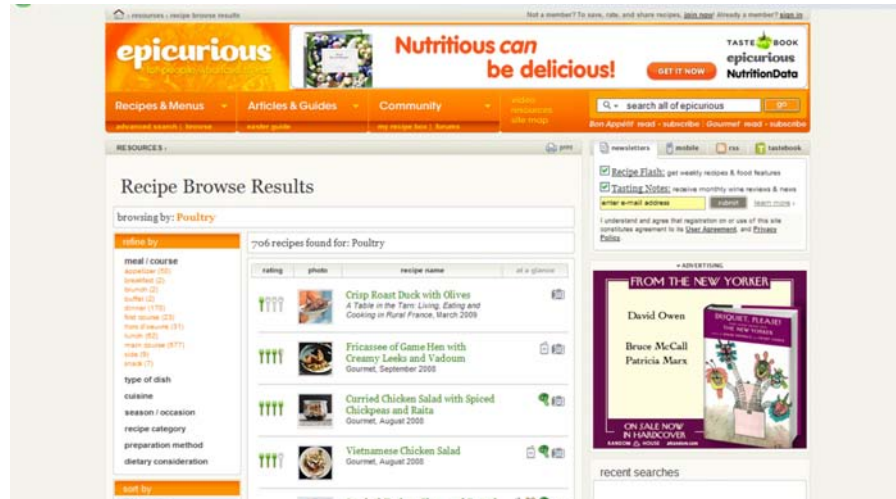


Fig. 3.4 epicurious is a commercial recipe website that uses Flamenco style faceted search in conjunction with keyword search.

search, branded as Guided Navigation, to large businesses including Wal-Mart and IBM. epicurious, shown in Figure 3.4, provides both faceted and keyword search over food recipes in a style that is similar to Flamenco’s experience. The lead researcher of Flamenco, Marti Hearst, refers to epicurious as a good example of faceted browsing in commercial conditions [73].

Huynh et al. [81] has developed Exhibit⁹ (Figure 3.5), a faceted search system that is similar to flamenco in many ways, but has some significant developments. One key advance is that, where a selection in Flamenco filters all the facets, a selection in Exhibit filters all the other facets and leaves the facet with the selection unchanged. This provides two benefits: first, users can easily change their selection and second, users can make multiple selections in one facet. This allows users to see, for example, the union of all the red and blue clothes, rather than just red or just blue. This support for multiple selection within a single facet has been recently added to ebay.com, but remains unavailable in services such as Google Product Search.

⁹<http://simile.mit.edu/exhibit/>.

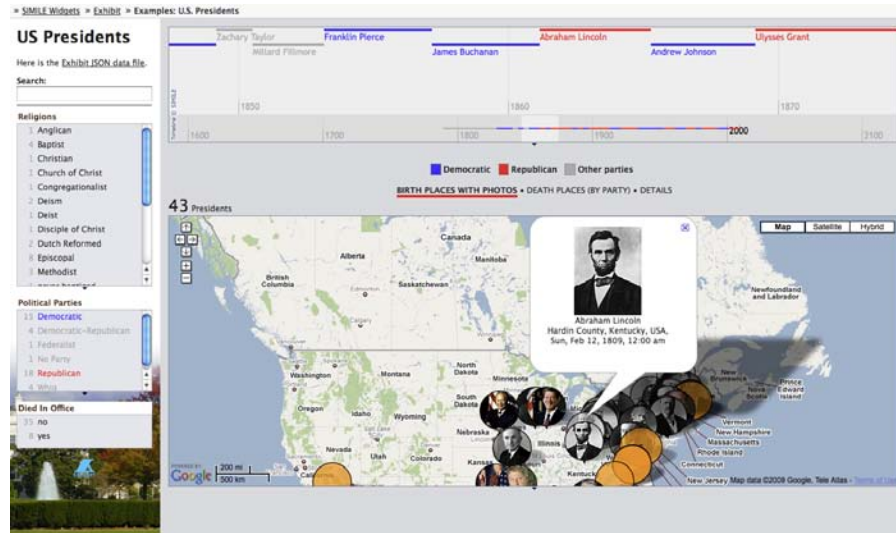


Fig. 3.5 The Exhibit faceted search interface takes a slightly different approach, only filtering facets without a selection, so that previous selections can be seen in the context of the options at the time.

The Relation Browser¹⁰ [192], shown in Figure 3.6, takes another approach to faceted search. One notable difference is that multiple selections lead to their intersection of results being displayed. Another feature that the Relation Browser provides is a preview of the affect of clicking has on other facets. Graphical representations behind each item in each facet show how many documents can be found by selecting it. When users hover over any item in any facet, the bar in the graphical representations is shortened to indicate how many documents will remain under each annotation should the users make the selection. This technique revives the query preview strategy, which is a helpful alternative to the simple numeric volume indicators [181] that are included in most classification-based systems. Aside from the graphical representation, the preview of the affect by simply hovering (or “brushing”) over the item is a technique that is being included in many new projects. A new version of the Relation Browser is now in development [32].

¹⁰<http://idl.ils.unc.edu/rave/>.



Fig. 3.6 The Relation Browser interface provides consistent facets, where the list of values is not filtered by selections. Instead, users can see (and preview by simply hovering) the reduction in files associated with each facet-value with the bar-chart style visualizations.

mSpace¹¹ [148], shown in Figure 3.7, represents yet another type of faceted search: a column-faceted interface. Like iTunes, mSpace presents facets in a left-to-right set of columns. Each column is fully populated so that users can still make a selection in any facet, but then only the columns to the right filter. This allows the facets to represent additional information that would otherwise be lost in faceted search. In their classical music example, where the facets are Era, Composer, Arrangement, and Piece (from left to right), if users select a Composer, they see a filtered list of the arrangements (s)he used and then a list of the pieces (s)he composed. When users make a second selection in the arrangement, other forms of faceted search would remove all the composers that did not use that arrangement. In mSpace, users are still

¹¹ <http://mspace.fm>.

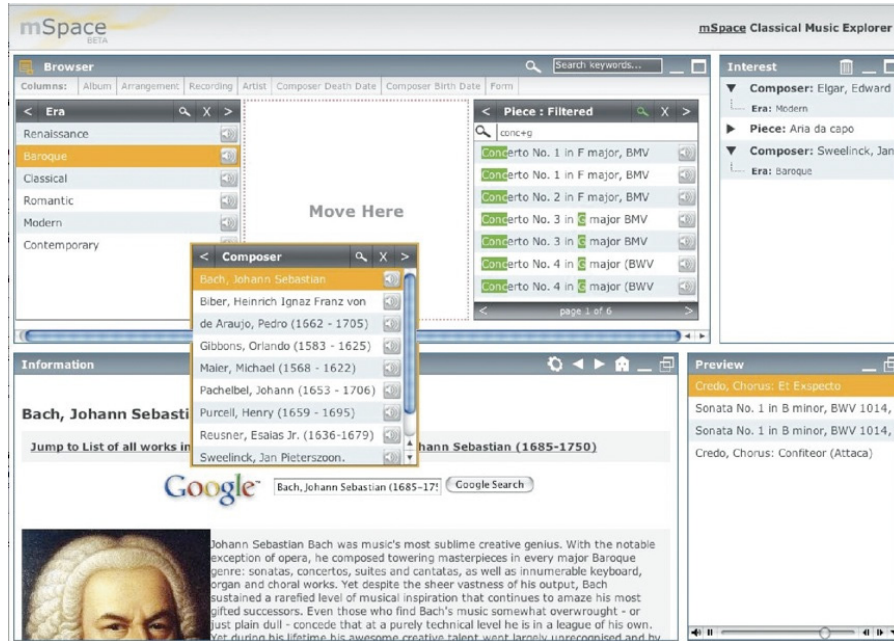


Fig. 3.7 The mSpace faceted column browser provides facets as columns that filter from left to right, as in iTunes. The facets, however, can be rearranged so that different parent-child relationships can be visualized spatially.

able to see all the arrangements that the selected composer used, and all the pieces of the selected arrangement.

The functionality of mSpace means that there is a type of information that is not conveyed by mSpace but seen in other forms of faceted search: the Era of the selected composer. This missing information is not a problem in mSpace, unlike other column-faceted browsers like iTunes, because the related items in facets to the left of a selection are highlighted. This allows users to see which era the selected composer is in, whilst still allowing them to get the added facts provided. The effect of this leftward highlighting, named Backward Highlighting [179] is that users incidentally discover more facts about the structure of a collection, which can make search easier for future searches in the same domain [183].

Finally, given the importance placed on the direction and order of facets in columns, mSpace allows users to reorder, remove, and

supplement the facets shown, using any of the facets that are available from the collection. This allows users to say that they would rather know all the composers that used a given arrangement than all of the arrangements used by a given composer. Another unique aspect of mSpace is that it assumes that a selection that is different to a previous selection in a facet is a change of selection and not a multiple selection. Consequently, users can quickly compare the difference between two items in a facet. The default assumption that users are changing their selection supports the concept of answering subjunctive questions about the collection [113], which means simply to compare the outcomes of multiple actions.

A recent longitudinal study of mSpace [182] has indicated that this more complex form of faceted search is easy to learn and is thereafter perceived as a powerful system, receiving positive subjective views. A log analysis showed that 50% of participants used facets in their first visit to the site and 90% in their second visit. Over the whole month-long period, there were more interactions with facets than individual keyword searches. Further, given that facets allow users to produce complicated queries than basic keyword searches, faceted searches represented two times the number of Boolean searches and three times the number of advanced searches.

Each of the faceted classification examples so far has been on fixed collections, but some research into faceted browsing has been looking at unknown and un-bounded document sets like the Web. The SERVICE¹² (Search Result Visualization and Interactive Categorized Exploration) search system couples the typical ranked list of web search results list with automatically generated facets [106] (Figure 3.8). Clicking on a category filters (or narrows) the displayed results to just the pages within that category. Moving the pointer over a category highlights the visible search results in that category in yellow. Moving the pointer over a result highlights all the categories in the overview that contain the result.

The facets are automatically generated by applying fast-feature classifiers [105] over the top 100 results of a Google query, and organized

¹²<http://www.cs.umd.edu/hcil/categorizedoverview>.

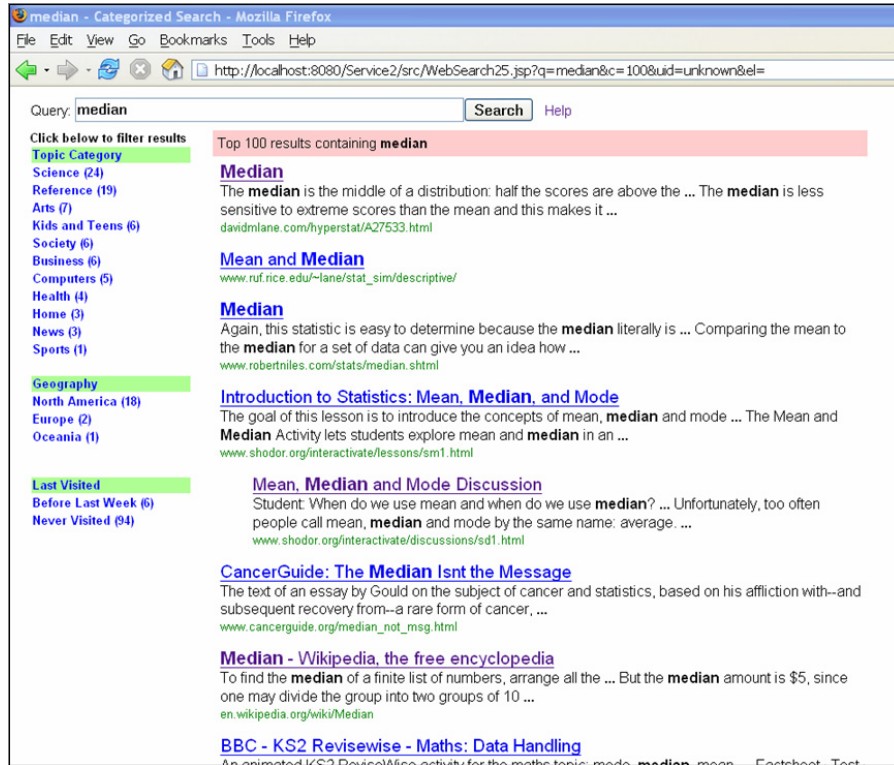


Fig. 3.8 Unlike most faceted systems, which work on known collections of documents, the SERVICE web search interface classifies un-bounded web collections in multiple facets [106].

them into known possible categories drawn from the Open Directory Project (ODP) and a database of US Government websites¹³: Topic, Geography, and US Government. A similar project: Dyna-Cat [136], shown in Figure 3.9, also automatically produced facets for sets of search results, showed that not only was there improvements in objective and subjective measures, that users were 50% faster in fact-finding tasks using Dyna-cat over typical ranked list keyword search interfaces.

Northern Light,¹⁴ a commercial search service, provides a similar capability by grouping results in their Custom Search Folders.

¹³ <http://www.lib.lsu.edu/gov/tree>.

¹⁴ <http://www.northernlight.com>.

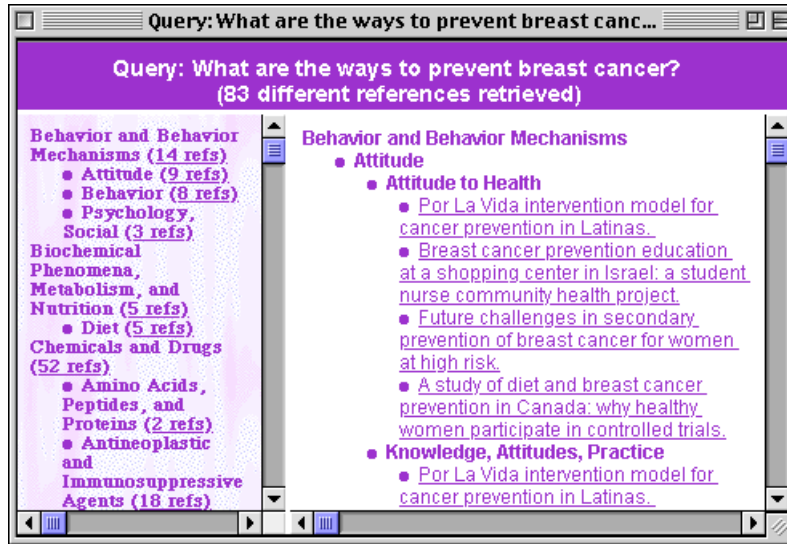


Fig. 3.9 The Dyna-Cat search interface automatically classifies search results within un-bounded document collations.

Exalead¹⁵ is another project that successfully organizes search results according to categories drawn from the Open Directory Project, and presents them alongside search results in a publicly available web search engine. PunchStock search, shown in Figure 3.10, presents a faceted view of personal image collections. The NCSU library, shown in Figure 3.11 also uses facets to enhance the search of their collection.

The faceted search interfaces described so far each have a space allocated to the presentation of facets, such as down one of the sides, across the top, or even along the bottom as in Google's product search.¹⁶ A recently proposed faceted search interface, called FacetPatch, embeds faceted options into the existing layouts of existing website [120]. Hovering over an attribute of a camera, for example, converts the attribute's value into a drop-down list from which users can select an alternative. Users can move, therefore, directly from result to result, by changing the brand or altering the pixel-count that they desire.

¹⁵ <http://exalead.com>.

¹⁶ <http://www.google.com/products/>.

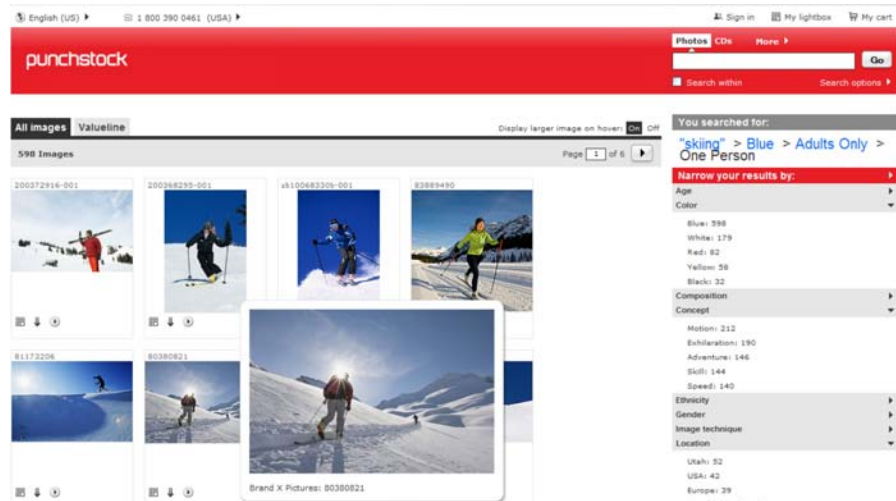


Fig. 3.10 The PunchStock photo search interface provides categorized overviews of photo search results.

While some of the studies discussed above focus on typically IR level metrics, like speed and accuracy of simple search tasks, there are notable exceptions. Flamenco, the Relation Browser, and mSpace have each been studied in more exploratory contexts, where the task is often more than to simply find a search result. The study by Yee et al. [188], sets users exploratory tasks that were not measured by speed or accuracy, as depth of exploration and amount of content covered could be considered more important than how fast and perhaps under-researched an answer is. This type of IS level analysis was also performed by Capra et al. [33]. Wilson et al [179] also studied the amount of incidental information (information not part of the core task) that could be remembered by participants in a study. The same team took this further by studying, over time, users of an mSpace interface providing access to a news footage archive [182]. Participants used mSpace, which was being logged, for their own work-goals over a four-week period, and engaged in periodic communication with the evaluating team. We can see from these studies that the types of exploring and browsing facilitated by faceted browsers have been studied for their support of IS and even WC levels of searching behavior.

The screenshot shows the NCSU Libraries search interface. At the top, there are navigation links: "Search the Collection | Browse Subjects | Services | Library Information | Community | News & Events". Below this is a header with "NCSU LIBRARIES" and "MY LIBRARY: Library Account | My Courses". The search bar contains the query "dog" and the search type is set to "Keyword Anywhere". There are "Go" and "Start" buttons, and a "Send search to:" dropdown menu.

The search results are displayed under the heading "Search 'dog':". A message states: "We found 4025 matching items. Limit results to currently available items." Below this is a "Browse By:" section with a grid of subject categories and their counts:

A - General Works (6)	M - Music (43)
B - Philosophy, Psychology, Religion (56)	N - Fine Arts (69)
C - Auxiliary Sciences of History (6)	P - Language and literature (969)
D - History (General) and History of Europe (52)	Q - Science (176)
E - History: America (64)	R - Medicine (71)
F - America: local history (27)	S - Agriculture (1871)
G - Geography, Anthropology, Recreation (53)	T - Technology (44)
H - Social sciences (106)	U - Military science (General) (4)
J - Political Science (24)	V - Naval science (7)
K - Law in general, Comparative and uniform law, Jurispruden ... (21)	Z - Bibliography, Library Science, Information resources (ge ... (15)
L - Education (44)	

Below the "Browse By:" section is a "Narrow Results By:" section with three categories: "Subject: Topic", "Subject: Genre", and "Format". Each category has a list of sub-categories with counts and a "Show More ..." link.

The main results area shows a list of search results. At the top, there are links for "Brief View | Full View" and a "Sort By:" dropdown menu set to "Relevance". The results are numbered 1 through 5:

- Dog world (Westchester, Ill.)**
Format: Journal or Magazine; Serial
Online: Search for electronic holdings
Print: Display bound volumes
- Dog world (Ashford, Kent)**
Format: Journal or Magazine; Serial
Online: Search for electronic holdings
Print: Display bound volumes
- Dog and the sheep. English.**
Published: 1681.
Format: eBook
Online: View resource online
- Advances in reproduction in dogs, cats and exotic carnivores : proceedings of the fourth International Symposium on Canine and Feline Reproduction, Oslo, Norway, 29 June-1 July 2000**
Author: International Symposium on Canine and Feline Reproduction (4th : 2000 : Oslo, Norway)
Published: 2001.
Format: Book
Veterinary Medical Library
SF427.2 .I575 2000 Stacks Available
- The AKC's world of the pure-bred dog**
Published: c1983.
Format: Book
D.H. Hill Library
SF426 .A43 1983 Stacks (8th floor) Available

Fig. 3.11 The NCSU library catalog provides categorized overviews of search results using subject headings, format, and library location, provided by a commercial enterprise search vendor: Endeca.

3.1.3 Automatic Clustering

Where collections have been large or unmanaged, faceted classifications have been less successfully applied. The examples where it has been

used above provide generic web-oriented facets, and often the facets that would be important depend on the query or subject of the users' varying goals. Another approach is to automatically identify attributes of a collection or result set that are important, rather than explicitly producing labeled annotations. This approach is called clustering, and has shown success under popular information retrieval metrics such as precision and recall [74, 118, 189, 191] or task completion time [169]

A one-level clustered overview was found helpful when the search engine failed to place desirable web pages high in the ranked results, possibly due to imprecise queries [89]. Clusty¹⁷ uses the clustering technique to produce an expandable overview of labeled clusters (Figure 3.12). The benefits of clustering include domain independence, scalability, and the potential to capture meaningful themes within a set of documents, although results can be highly variable [71].

Generating meaningful groups and effective labels, however, is a recognized problem [139] and where possible (usually, but not exclusively in fixed or managed collections) having an information architect design optimal annotations or labels, will provide a better search interface [162]. Stocia and Hearst extracted a category hierarchy from

The screenshot shows the Clusty search engine interface. At the top, there is a search bar with the query 'median' and a search button. Below the search bar, there is a navigation menu with 'clusters', 'sources', and 'sites'. The main content area displays 'Top 209 results of at least 5,950,000 retrieved for the query median (definition) (details)'. The results are listed as follows:

- Median - Wikipedia, the free encyclopedia**: In probability theory and statistics, a **median** is described as the number separating the higher half of a sample, a population, or a probability distribution, from the lower half. The **median** of a finite list of numbers can be found by arranging all the observations from lowest value to highest value and picking the middle one. Notation: Measures of ... **Medians of ... - Medians in ...** en.wikipedia.org/wiki/Median - [cache] - Live, Ask, Gigablast
- Median**: This is part of HyperStat Online, a free online statistics book. dawonline.com/hyperstat/A27533.html - [cache] - Live, Ask
- Median**: A simple explanation of **median**, for journalists and other writers who might not know math. www.robertshiles.com/stats/median.shtml - [cache] - Live, Ask, Gigablast
- How to Calculate the Median Value**: Math explained in easy language, plus puzzles, games, quizzes, worksheets and a forum. For K-12 kids, teachers and parents. www.mathsisfun.com/median.html - [cache] - Live, Ask
- Median Line Trading Techniques**: Technical analysis using Andrews **Median** lines, Babson Action Reaction lines, and Fibonacci relationships. www.medianline.com - [cache] - Live, Open Directory
- Median**: Here are the sample test scores you have seen so often: 100, 100, 99, 98, 92, 91, 91, 90, 88, 87, 87, 85, 85, 85, 80, 79, 76, 72, 67, 66, 45. www.mste.uiuc.edu/hill/stat/median.html - [cache] - Live, Ask
- median Definition**: One type of average, found by arranging the values in order and then selecting the one in the middle. If the total number of values in the sample is

Fig. 3.12 The Clusty metasearch engine uses automated clustering to produce an expandable overview of labeled clusters.

¹⁷<http://www.clusty.com>.

WordNet [122] using keywords from the document collection. They propose that this can be manually adjusted by an information architect. Similarly, Efron et al. [54] investigated the use of semi-automated methods, combining k-means clustering, and statistical classification techniques to generate a set of categories that span the concepts of the Bureau of Labor Statistics web pages and assign all to these categories. They found that concept learning based on human-supplied keywords performed better than methods using the title or full-text.

Automated classifications have, in large, been studied separated from the three levels of search, but have often compared the accuracy of the automatic classification with human classifiers. The studies by Stocia and Hearst [162] and Rivadeneria and Bederson [139] are examples, however, which show that automatic classifications can still lead to improved user performance in IR level searching activities.

3.1.4 Social Classifications

Recent developments in web technologies, such as web2.0, have led to a rise in social classifications such as bookmarks, tagging, and collaborative feedback ratings. For example, the Yoogle search engine allows users to explicitly move the search results up or down to provide relevance feedback into the weightings used in future searches. Allowing users to reorder results has recently been supported for Google users with accounts, but is designed to improve personalization of search results. It remains unclear as to whether Google will use this relevance feedback to reorder results. Little research, however, has explicitly proven the benefits of social classification schemes, even though the notion has become increasingly popular on the Web. Some research, however, is emerging: Millen et al. [121] have investigated the use of social bookmarking in enterprise search software and an eight-week field trial has shown positive results.

In Yoogle¹⁸ (Figure 3.13) mentioned above, the classification produced is only numeric (document ranking order) and affects the background search algorithms. Google has recently introduced a similar

¹⁸<http://www.yoogle.net/>.

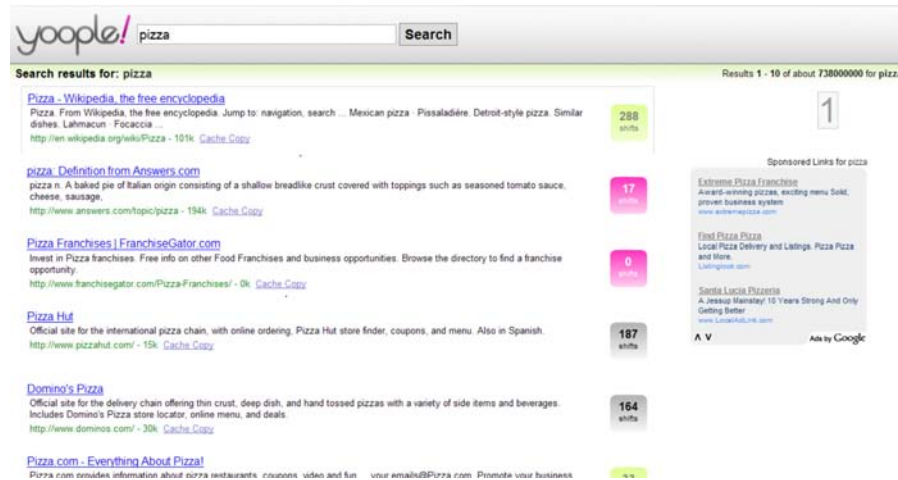


Fig. 3.13 Yoogle Search, where users can drag items in the results list to indicate where they think a result should appear.

personalization scheme called Search Wiki.¹⁹ While Google's Search Wiki allows users to simply promote or remove results, Wikia Search²⁰ (Figure 3.14) allows users to be more specific and rate results out of five. Both Google's Search Wiki and Wikia Search allow users to directly influence search rankings in a similar way to Yoogle. Further, however, they both also allow users to make annotations on results.

The notion of social tagging has allowed communities of people to develop flat classification schemes for collections. Flickr,²¹ a photo archiving website, depends almost entirely, although photo names and descriptions can be included, on user tagging to return images that relate to a keyword search. A negative side to such tagging classification schemes is that they are hard to present to users. Therefore, they are usually only used to aid keyword search rather than to help users interactively browse through documents, like in some of the faceted and category-based systems listed above. A popular example of presenting a flat set of tags to users has been to foreground popular tags in a tag cloud (Figure 3.15). Although little research has been performed on

¹⁹ <http://googleblog.blogspot.com/2008/11/searchwiki-make-search-your-own.html>.

²⁰ <http://search.wikia.com/>.

²¹ <http://www.flickr.com>.

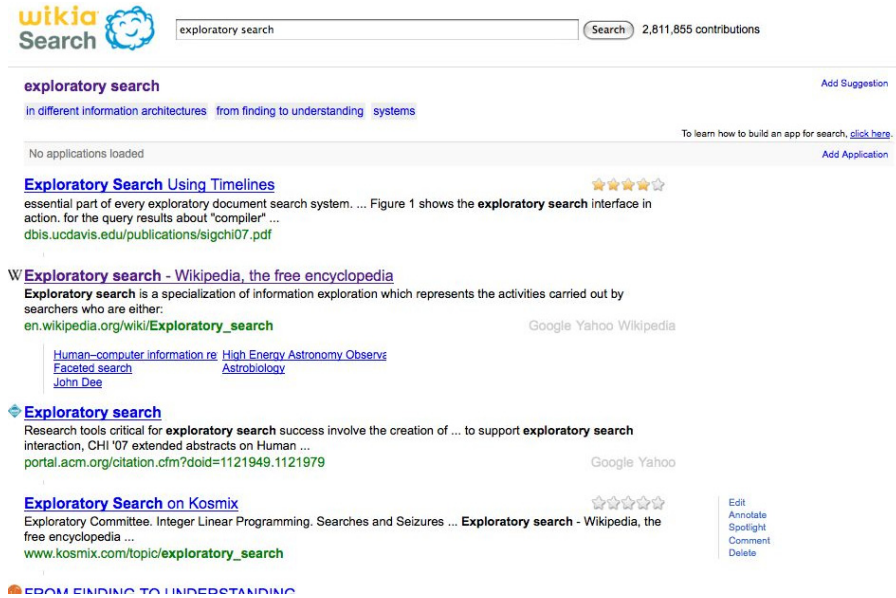


Fig. 3.14 Wikia Search allows users to explicitly rate results out of five to directly influence search ranking.

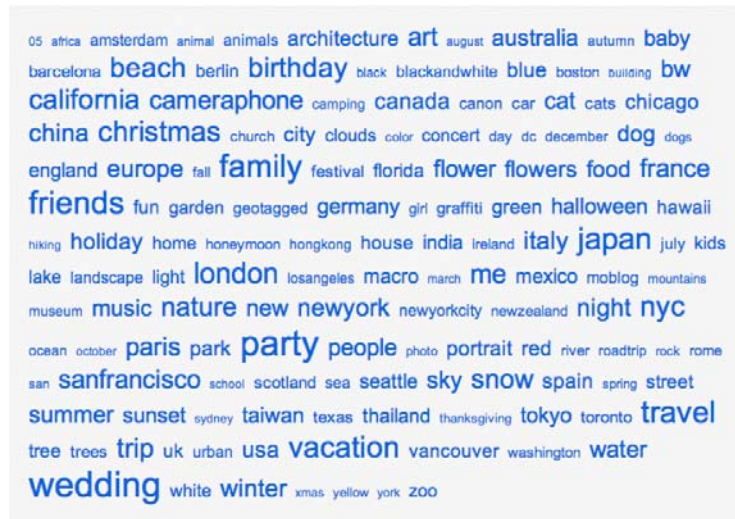


Fig. 3.15 A tag cloud taken from Flickr website, included in [65].

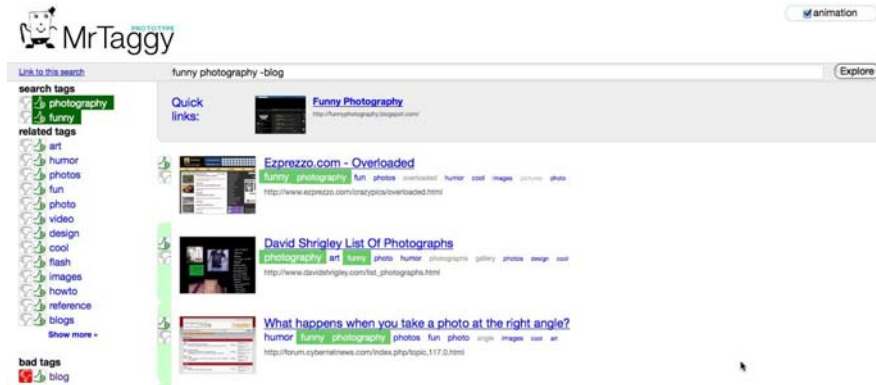


Fig. 3.16 The MrTaggy interface allows users to include or exclude tags from their search instead of using keywords.

tag clouds, there is a growing consensus that they are more valuable for users who are making sense of information, rather than for finding specific information [157].

Some systems are trying to take the benefits of tagging to produce what are known as folksonomies [80] or community-driven classification systems. Wu et al. [186] present some design prototypes that overcome some of the challenges in converting flat tags into a structure classification system. One approach, used by the MrTaggy interface shown in Figure 3.16, allows users to perform searches based entirely on community-generated tags by including or excluding tags from a list of related tags [90]. Searches are initiated with a pair of selections from two tag clouds: one containing adjectives and the other containing nouns and other objects.

3.2 Result Organization

At the opposite end to supporting users in better defining their own search query, which has been discussed above, is the presentation of the result sets and items that will help users to identify the specific result item(s) that help them achieve their goals. There is ample research that specifically looks at the best way to present a single result in a linear set, which is perhaps best represented by the view currently provided by Google: name, text snippet, and web address.

In the following subsections, however, we discuss the research that looks beyond a linear results list and visualizes a result set to help users find the specific results they are looking for.

3.2.1 Result Lists

The most common method to show a result set, seen in most web search engines, is to provide a simple list of results. The order that the result items are listed is determined by some metric, usually based on how relevant the result is to the search terms used. The importance of constructing this algorithm well has been motivated many times in research where users regularly enter single-term ambiguous queries [171], and view only a few results [82] and rarely stray past the first page of results [17]. The accuracy and efficiency of such algorithms, however, has been research for many years in the information retrieval space, and is not covered in the scope of this report; Becks et al. [16] present a good discussion of result ordering, including occasions where ranked relevance may not be the most appropriate ordering.

The representation of each result has also received much research and is especially important when research has shown that the acceptance of search systems has been significantly reduced by unclear or confusing representations.²² In Google, each search result has name (also a link to the finished document), a sample of text from the document, a URL for the document, and the size of the document. Additional information, such as previous visit dates and counts, can be added if a user account is available. Work by Chen et al. [37] investigated the structure of result representations and provides a framework for considering the presentation of search results. In one of the more significant publications in this area, White et al. [175] showed that best objective results occurred when the text sample in a representation included the query terms used in the original search. This allows users to see the context of their query in each result item, so that they can best judge its value in viewing the whole document. This was backed up by Drori and Alon [49], who showed objective and subjective benefits

²²<http://www.antarctica.net>.

for both query-relevant snippets and associated categories. Microsoft's Bing, however, includes an exception to this keyword-in-context rule for text snippets, providing the first whole paragraph of a Wikipedia page, which usually provides the best short overview for the topic being covered. This exception to the basic search result layout has been furthered by the now customizable search results provided by Yahoo's Search Monkey,²³ which allows users to apply templates for different popular result types, such as Wikipedia results, IMDB results, facebook results, and so on.

Other search engines, including Ask Jeeves²⁴ and Exalead²⁵ (Figure 3.17), also include thumbnails of the document represented by each result. Research into the benefits of page thumbnails [185] has shown most advantage when users are returning to a previous search to find a result from the previous session. With these conditional objective benefits, however, subjective results have been positive. Further, Teevan et al. [166] have shown that more abstract images, which integrate

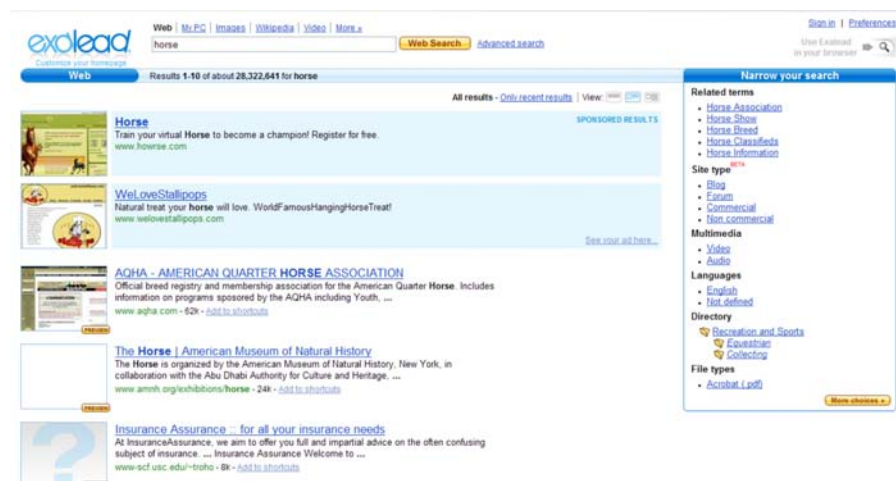


Fig. 3.17 Exalead search provides images and categories with each search result, and facets to help users narrow the results.

²³ <http://www.yahoo.com/searchmonkey>.

²⁴ <http://www.ask.com>.

²⁵ <http://www.exalead.com>.

color, titles, and key pictures from web pages, can support both search and re-visitation more evenly.

The advances made in the basic style of Results Lists have had little effect on the interaction that people have with search interfaces. Advances such as thumbnails have changed the representation, and have encouraged users to make better-educated judgments on results they see. Consequently, the contribution and subsequent study has mainly focused on IR level improvements of speed and accuracy.

3.2.2 2D Result Representations

With the aim of providing useful overview visualizations of result sets to users, much research has aimed at taking information visualization techniques and applying them to search results. One example is self-organizing maps (SOMs), originally produced by Kohonen [97]. SOMs automatically produce a two-dimensional (2D) visualization of data items, according to the attributes that they share, using an unsupervised machine-learning algorithm. Consequently, the SOM approach applies well when visualizing automatically clustered documents.

Au et al. [10] used SOMs that have been used to support exploration of a document space to search for patterns and gain overviews of available documents and relationships between documents. Chen et al., [37] compared an SOM with the Yahoo! Entertainment category, for browsing and searching tasks (Figure 3.18). They found that recall improved when searchers were allowed to augment their queries with terms from a thesaurus generated via a clustering-based algorithm. Similar work [108, 110] has shown positive results in using SOMs to support search.

An alternative 2D organization on result sets is known as a treemap [60]. In a treemap, the clusters of documents are grouped using a space-filling algorithm. As shown in Figure 3.19, the 2D space is divided into the top-level items in the hierarchy produced by the clustering algorithm, where the number of search results found within the category dictates the size of each part. Each of these sections is divided into their children within the hierarchy, where the number of search results again dictates the size of the subparts. This process is repeated as necessary according to the hierarchy and space available. With this,

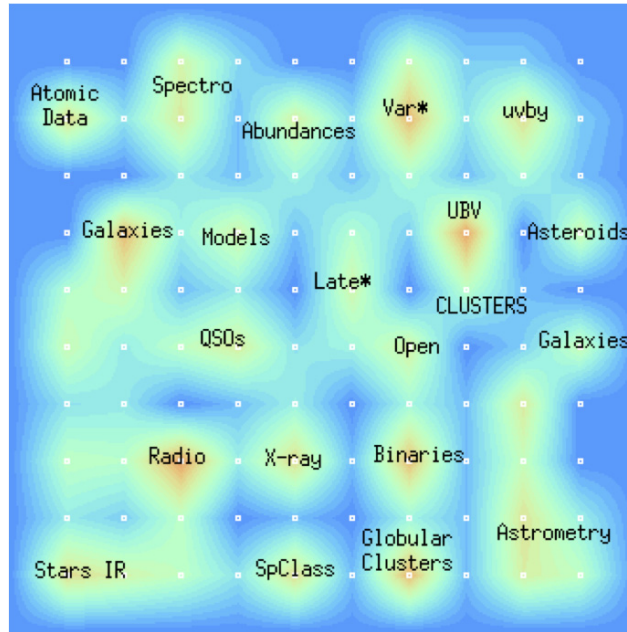


Fig. 3.18 A SOM of search results, image taken from <http://vizier.u-strasbg.fr>. The results are shown as clusters of key topics relating to a query.

users can access categories at any level of the hierarchy and see the results associated with that cluster. Color coding, or use of increasing color density, is often used to indicate a secondary dimension over the data, so that popularity and price, for example, are both conveyed (Figure 3.20).

Using another information visualization alternative, Citiviz (Figure 3.21) displays the clusters in search results using a hyperbolic tree [109] and a scatterplot [133]. The Technical Report Visualizer prototype [64] allows users to browse a digital library by one of two user-selectable hierarchical classifications, also displayed as hyperbolic trees and coordinated with a detailed document list. Chen [35] implemented clustered and time sliced views of a research literature to display the evolving research fronts over time. Chen's work concludes that there are a number of benefits in visualizing results like this, but that the metrics used to build and display the visualizations are important to the acceptance by users.

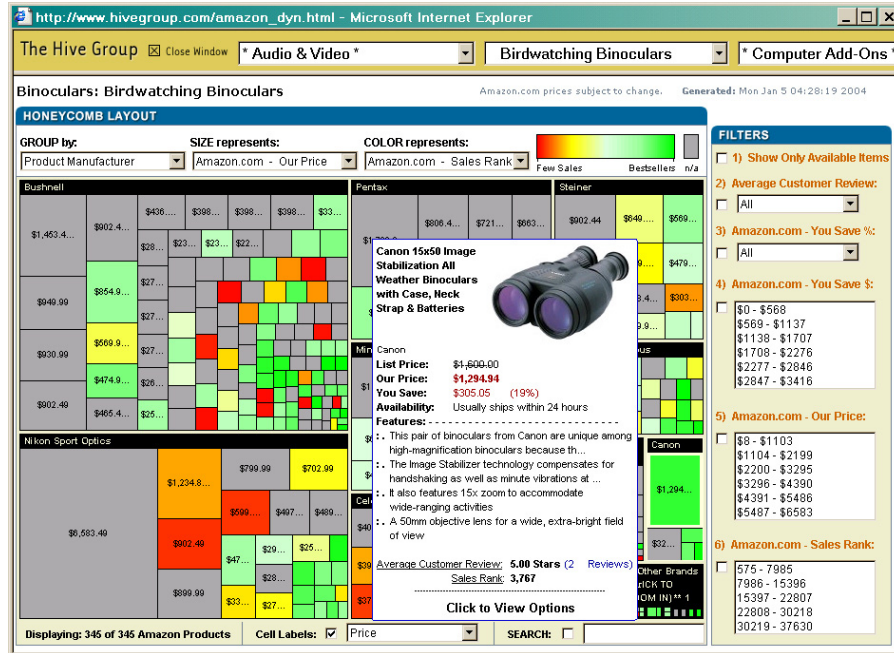


Fig. 3.19 The HiveGroup built a demonstration shopping application that allows users to explore Amazon.com products. (image from <http://www.cs.umd.edu/hcil/treemap-history/hive-birdwatching.jpg>).

Several web search (or metasearch) engines, including Grokker,²⁶ Kartoo,²⁷ and FirstStop WebSearch²⁸ incorporate visualizations similar to treemaps and hyperbolic trees. Kartoo (Figure 3.23) produces a map of clusters that can be interactively explored using Flash animations. Grokker clusters documents into a hierarchy and produces an Euler diagram, a colored circle for each top-level cluster with sub-clusters nested recursively (Figure 3.22). Users explore the results by “drilling down” into clusters using a 2D zooming metaphor. It also provides several dynamic query controls for filtering results. Unlike treemaps, however, the circular form often leads to wasted space in the visualization. Further, this early version of the Grokker interface has been found

²⁶ <http://www.grokker.com>.

²⁷ <http://www.kartoo.com>.

²⁸ <http://www.firststopwebsearch.com>



Fig. 3.20 This overview of web search results uses a treemap. Nesting is used to show top and second-level categories simultaneously. The top 200 results for the query “urban sprawl” have been categorized into a two-level government hierarchy, which is used to present a categorized overview on the left. The National Park Service has been selected to filter the results. The effect on the right side is to show just the three results from the Park Service [106].

to compare poorly with textual alternatives [139]. The authors found that the textual interfaces were significantly preferred. The conclusions were that web search results lack “1)... a natural spatial layout of the data; and 2)... good small representations,” which makes designing effective visual representations of search results challenging. Grokker’s website is no longer available. Refined visual structures making better use of space and built around meaningful classifications and coloring may ameliorate this problem, as illustrated by promising interfaces like WebTOC, which uses a familiar hierarchy of classification and a coded–coded visualization of each site’s content. Many users appreciate the visual presentation and animated transitions, so designing them to be more effective could lead to increased user acceptance.

In support of the conclusions about cluster map visualizations from Rivadeneira and Bederson [139], early information visualization

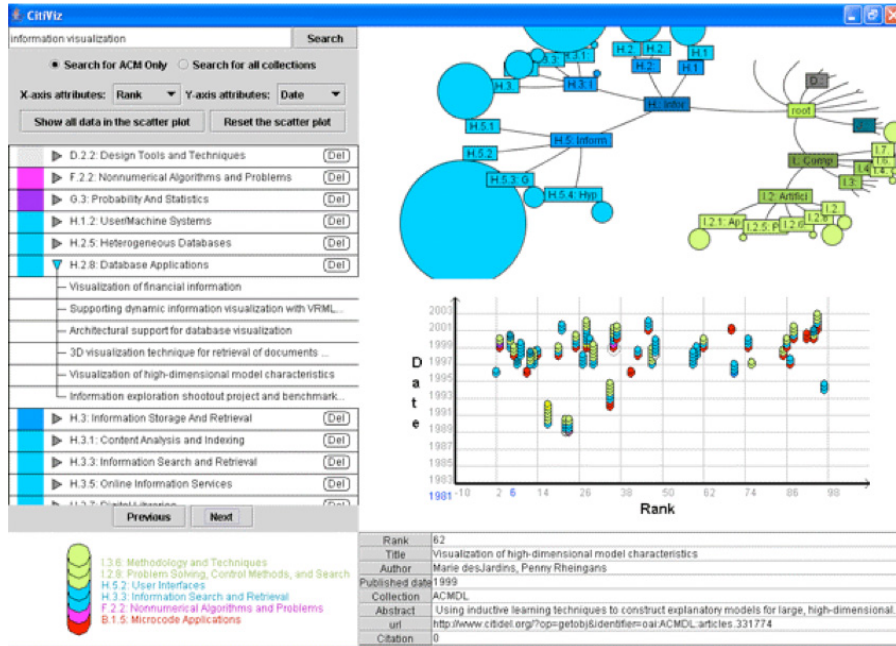


Fig. 3.21 The CitiViz search interface visualizes search results using scatterplots, hyperbolic trees, and stacked discs. The hyperbolic tree, stacked disks, and textual list on the left are all based on the ACM Computing Classification System. Although CitiViz is offline, its techniques have been replicated both online and offline many times.

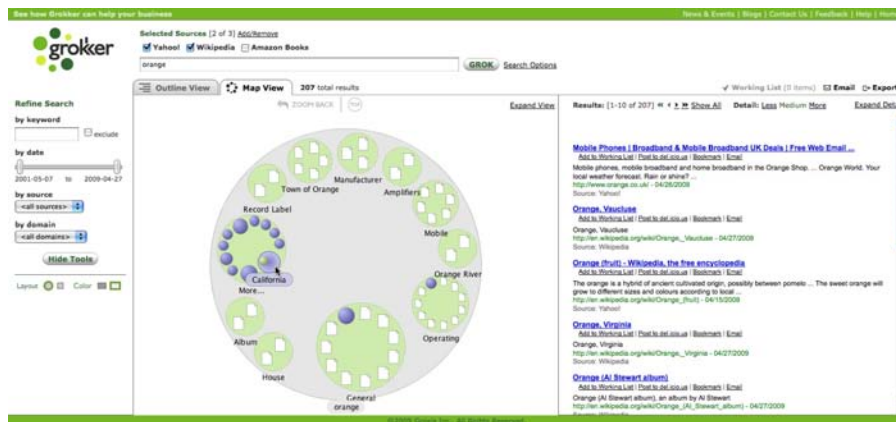


Fig. 3.22 Grokker clusters documents into a hierarchy and produces an Euler diagram, a colored circle for each top-level cluster with sub-clusters nested recursively. Later versions removed the random coloring in favor of a more muted interface.



Fig. 3.23 Kartoo generates a thematic map from the top dozen search results for a query, laying out small icons representing results onto the map.

research by Tufte [167] states that graphical visualizations need to have clear and meaningful axes to be effective for users. Other approaches to 2D graphical representations have focused on representing results in meaningful and configurable grids, in the same vein as scatterplots, where each axis is a specific metric, such as time, location, theme, size, or format, etc. The GRiDL prototype (Figure 3.24) displays search result overviews in a matrix using two hierarchical categories [154]. The users can easily identify interesting results by cross-referencing the two dimensions. The List and Matrix Browsers provide similar functionality [107]. The Scatterplot browser [53], also allows the users to see results in a visualization that is closer to a graph, where results are plotted on two configurable axes. Without specific regions plotted in the scatterplot visualization, it is harder for users to view subsets of results, but easier to identify single result items. Informal evaluations of these interfaces have been promising, although no extensive studies of the techniques have been published.

Another stream of research has focused on attributes of documents that can form dimensions that leverage common knowledge. For example, GeoVIBE (Figure 3.25) tightly coupled a geographic layout with an abstract layout of documents relative to defined points of interest

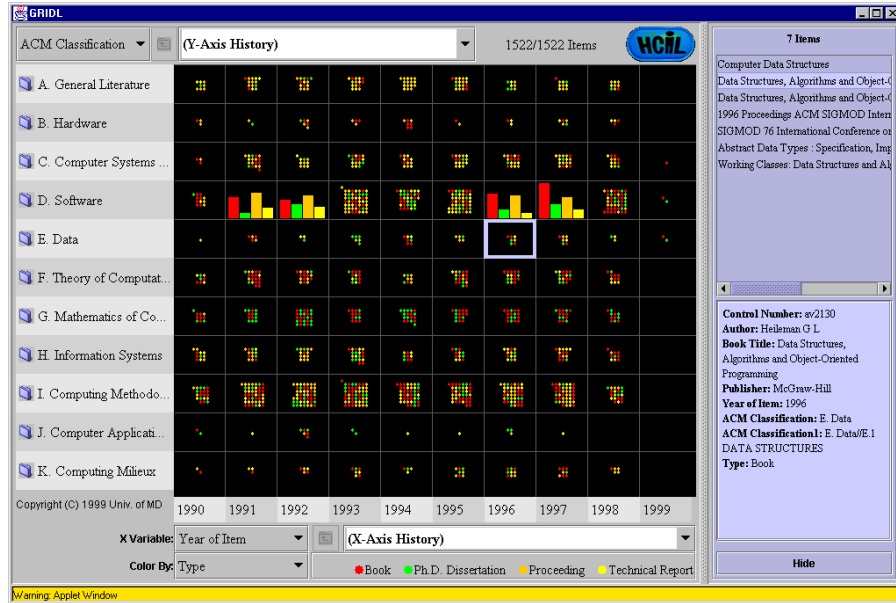


Fig. 3.24 The GRIDL prototype displays search results within the ACM Digital Library along two axes. In this screenshot, documents are organized by ACM classification and publication year. Individual dots are colored by type of document.

[31]. A similar, and familiar dimension is time, and research has also looked at visualizing documents and classifications on a timeline. An example timeline browser, Continuum [9], is shown in Figure 3.26.

Each of the visualizations above has focused on visualizing a result set, so that users can identify areas of results that may be relevant to them. A different approach is to provide an overview visualization to help direct people to relevant items in linear result sets. Query term similarity allows searchers to explore the contribution that each query term makes to the relevance of each result by displaying the terms and results in a 2D or 3D space [27]. Similar work on Hotmaps (Figure 3.27), by Hoeber and Yang [77], provides an overview of 100 search results, displaying a grid, with query terms as the horizontal axis and the 100 results as the vertical axis. Users are able to click on parts of the grid with high color intensity to find more relevant documents, at the term level, that would not have been in the top 10 results returned by Google. The approach of query and result set visualizations has

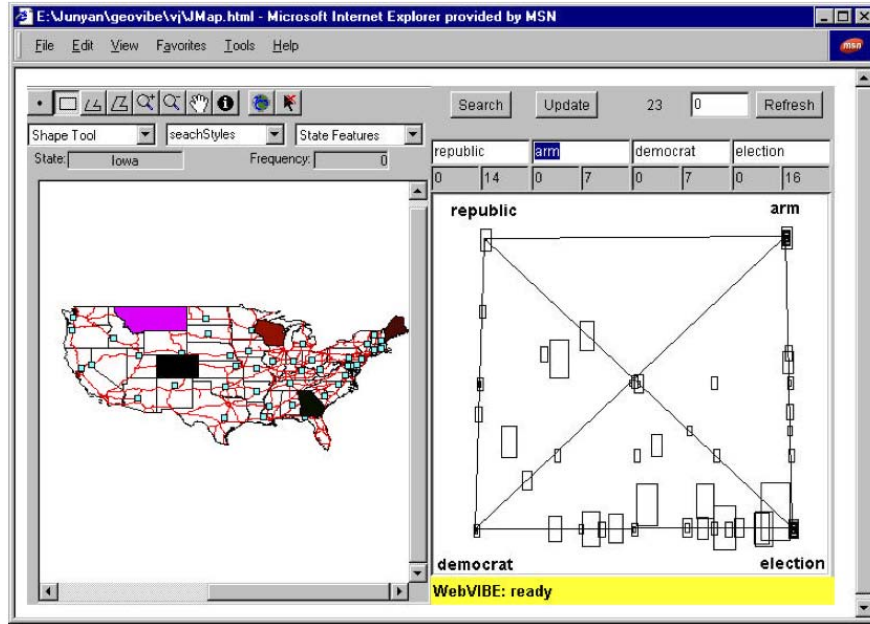


Fig. 3.25 GeoVIBE, a search interface using geography visualization for results.

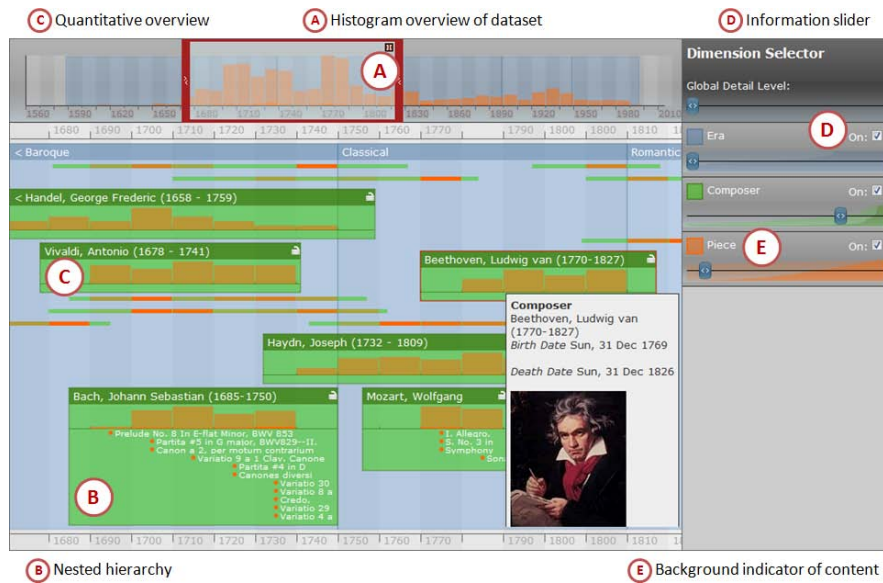


Fig. 3.26 Continuum represents documents in nested categorizations on a timeline.

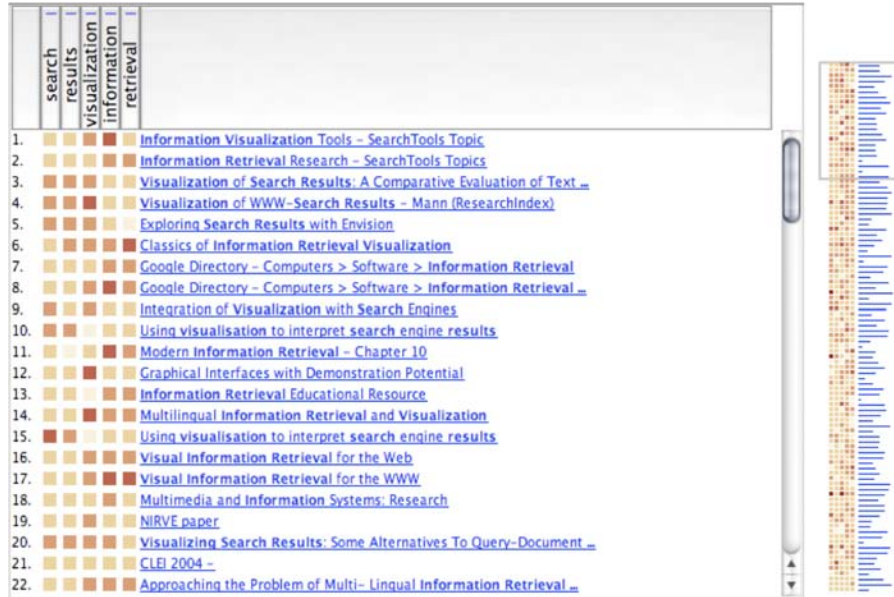


Fig. 3.27 Hotmaps, a 2D visualization of how query terms relate to search results.

shown significant search benefits for users, and augments, rather than visualizes, result sets.

The strength of the more successful 2D visualizations has been when the use of space is clear and meaningful to users. An alternative approach, to computer-generated visualizations or clusters of documents, is to allow users to arrange documents under their own parameters. The TopicShop interface [8], shown in Figure 3.28, permits users to meaningfully arrange sites on a canvas, where the clusters are meaningful to the individual and, when returning to a document, spatial consistency makes it easier for users to remember where they placed the result. More recent research in this area has explored 3D spaces, and is described in the section below.

Similar to the ideas behind TopicShop, Furnas and Rauch [62] present a set of design principles for information spaces and instantiates them in the NaviQue workspace. In NaviQue, a single information surface is used to display queries, results, information structures, and ad hoc collections of documents. Users can initiate a query simply

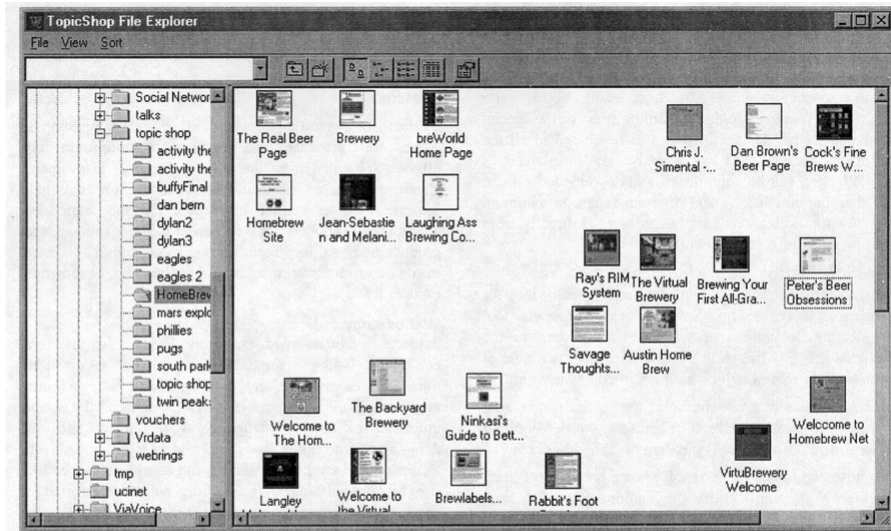


Figure 2: TopicShop Explorer, icons view. Each site is represented by a large thumbnail image and the site title. Users can organize sites by arranging them spatially, a technique especially useful in the early stages of exploration.

Fig. 3.28 The TopicShop Explorer interfaces combines a hierarchical set of topics with a user-controlled spatial layout of sites within each topic (shown here) or a detailed list of titles and attributes [8].

by typing onto a blank section of the workspace. Results are clustered around the query text. Kerne et al. [95] furthered this work in software called *combinFormation* by considering the spatial layout of different multimedia types for constructing personal collections (Figure 3.29).

The contributions made in different 2D visualizations have largely supported users in IS level activities, allowing them to see and often manipulate clusters of documents by interacting with the two axes. In *Hotmaps* [77], for example, the 2D space allows users to choose alternative approaches to parsing the result lists, other than linearly processing every result. The *matrix browser* allows users to change the dimensions used for each axis, and alter their scale and granularity. These have been studied at the IR and IS levels, showing improved speed or accuracy. These methods have also been shown to help users find results that would not necessarily come up in the top 10 of a keyword search interface. Some work has focused on the users' work context, for example. the *TopicShop* interface [8], has given control of

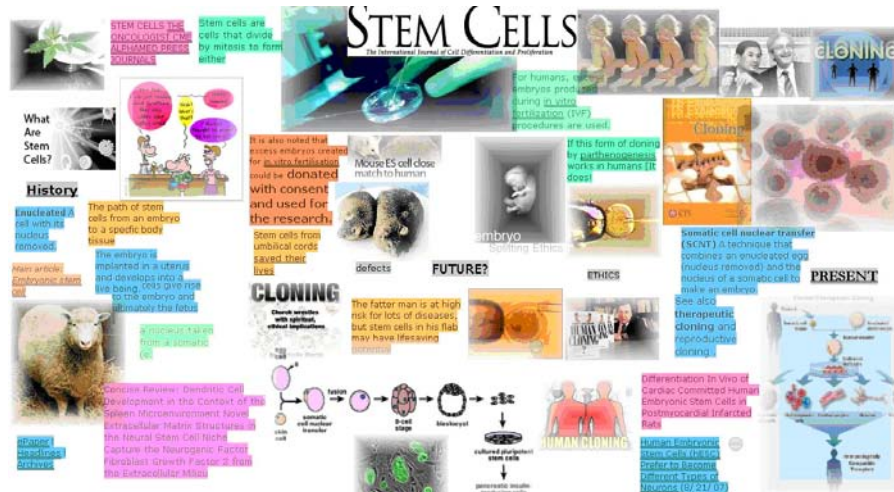


Fig. 3.29 combinFormation automatically searches the Web for multimedia relating to a query term and displays them on a 2D plane. Users can drag and drop the items to make arrange their own subcollections.

the 2D space to users as an open canvas, and they are able to search and organize resources for their work tasks.

3.2.3 3D Result Representations

When presenting research into 3D visualizations, it is first important to consider research into the benefits and weakness of 3D over 2D. Although research has shown success for some situations, such as simulations and 3D CAD/CAM design [152], research into the visualization of information has shown that a third dimension can inhibit users and make interfaces more confusing [138, 163]. Aspects of 3D representations, such as occlusion and cost of visualizing depth, must be considered [44]. Further, research by Modjeska [123] has shown that 25% of the population struggle with 3D visualizations displayed on a 2D device, such as a computer screen. Investigation by Sebrechts et al. [150] also showed that participants were significantly slower at using a 3D interface, unless they had significant computer skills. Considering these challenges, however, the research described below highlights some of the ideas that have been proposed for 3D visualizations.



Fig. 3.30 Data mountain provides an inclined plane upon which web pages are organized (image from PDF).

The Data Mountain browser (Figure 3.30), like the TopicShop explorer described above, allows users to arrange documents on a plane, creating a 2 1/2 dimension view with perspective, reduced size for distant items, and occlusion. In the Data Mountain, however, users have a 3D plane to arrange the documents [140]. Subsequent studies of the Data Mountain have shown to be unproductive [42]. Another approach that has been expanded to 3D environments is the hyperbolic tree [127].

As part of a discussion on applying 3D visualizations to web search results, Benford et al. [20] describe the VR-VIBE system (Figure 3.31). Each query is manually or automatically positioned in a 3D space. Documents are positioned near the queries for which they are relevant. If a document is relevant to multiple queries, it is positioned between them. Each document's overall relevance is shown by the size and shade of its representative icon.

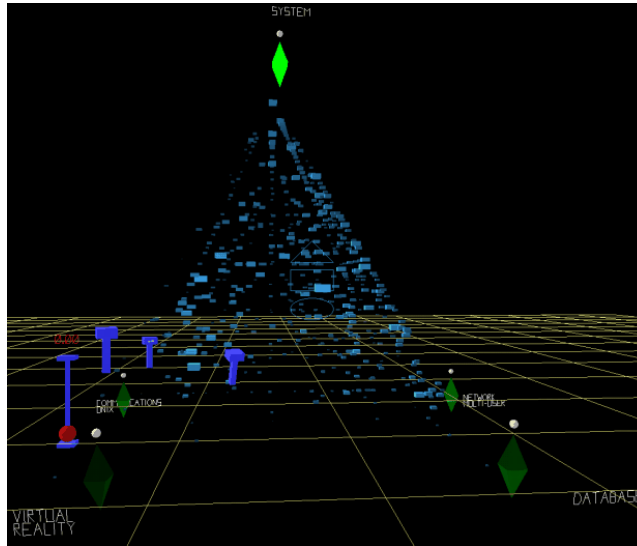


Fig. 3.31 VR-VIBE represents search results in a 3D environment, where results are displayed in relative proximity to the keywords they match.

Cone Trees [142] were designed to display a hierarchy in a 3D view, using depth to increase the amount of the tree that is visible. Users are then able to rotate the tree using a smooth animation, although this was found confusing by a number of participants in an evaluation and user studies have not shown advantages [41]; supporting the concerns noted by Modjeska.

Finally, to fully embrace 3D environments Börner [27] used Latent Semantic Analysis and clustering to organize and display a set of documents extracted from a digital library in a 3D space. A multi-modal, virtual reality interface, called the CAVE (Figure 3.32) enables users to explore the document collection; users control their movement through the CAVE with a special input device called a Wand. This requirement for a special input device, however, makes it an unrealistic option for web environments where most users will have typical mouse and keyboard input.

The effect of adding a third dimension has not made any specific advances in the three levels of search compared to the 2D advances. The majority of the interfaces discussed above are, in fact, a 3D version

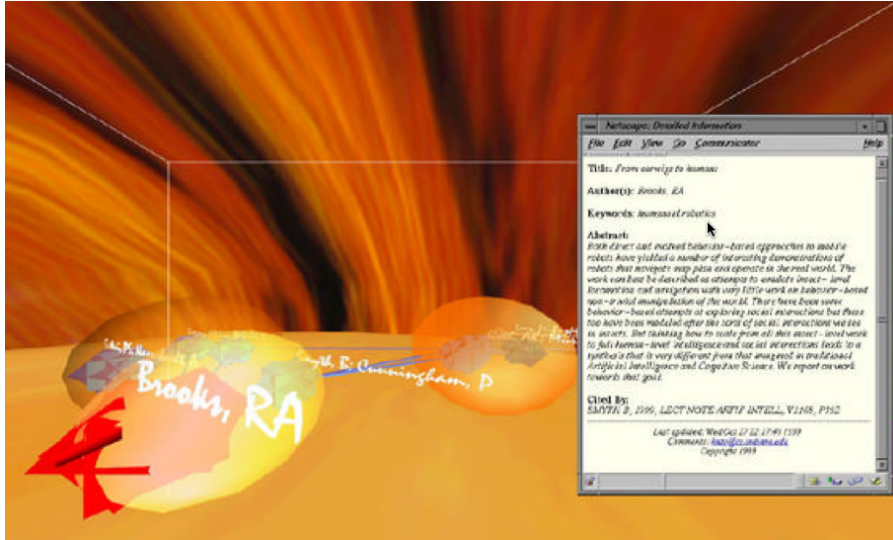


Fig. 3.32 Search results being visualized using the CAVE explorer (image from PDF).

of a 2D visualization, including the hyperbolic tree and Data Mountain interfaces. The implementation and evaluation of 3D visualizations have been scarce due to the limited capability of most web browsers, and when tested at the three levels of search (IR, IS, and WC), 3D visualizations have often hindered, rather than supported, participants in their searching activities.

3.3 Additional Functions

The previous two subsections have focused on: (1) coupling results with additional metadata and classifications, and (2) on providing alternative or complementary representations of results. This sub-section, however, focuses on specific individual features that can work together with both classifications and alternative visualizations to enhance the usability of search interfaces.

3.3.1 Previews

One of the challenges users face with classification systems is deciding which categories to select in order to find the documents they require.

This heavily depends on the label chosen to represent a category, where designers have to find a careful balance between clarity and simplicity; where Nielsen's heuristics [129] recommend clear layman language, it may not always be possible in some domains, at least with single terms or short phrases. If users do not recognize or understand an option, then they may not know how to proceed in their search, except through trial and error. Some users, however, simply abandon their search, at least with the search interface in question [147].

The Relation Browser previews the affect that a selection will have on the remaining items and facets, by temporarily reducing individual bar graph representations; selecting the item will make this change permanent. This notion has been used in research into the mSpace browser [147] to help users choose between the terms presented in a facet. When users hover over an item in an mSpace column facet, they are presented with a multimedia preview cue, which provides an example of the documents included within the category. A user evaluation showed that this preview cue significantly supported users in finding documents. Further research by Endeca [100] is investigating the ability to automatically summarize the result items and present users with a text description that is typical of the cluster.

For the three levels of search, presented in Section 2, previews contribute mainly toward IR and IS, by supporting users in making informed browsing and exploring decisions. This supports a number of additional search tactics [12], including the ability to weigh up a set of options, and tracing the metadata of the multimedia examples being previewed. schraefel et al. [147], however, evaluated mSpace's multimedia cues in a work-context scenario of purchasing music within a given budget.

3.3.2 More Like This

Search engines typically provide users with a link by each result to see more results that are similar according to the server's metrics. Although no specific research has been published, Endeca's search interface (Figure 3.33) allows users to express a similar desire to see related documents, but allows them to express the dimension by which they want

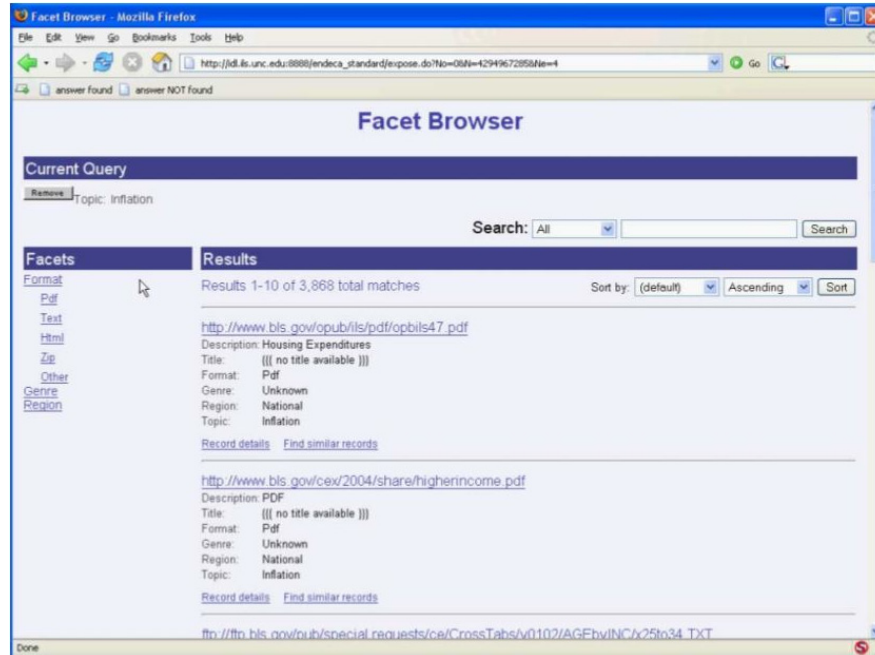


Fig. 3.33 Basic Endeca search interface, a commercial enterprise search vendor, studied by Capra et al. [33].

to see similarity. For example, they can choose to see more documents of a similar category, or size, or creation date.

By allowing users to follow un-anticipated paths during search, the simple notion of choosing to see more similar results to a particular result also supports a number of information tactics that are not supported by standard keyword search interfaces [12], and so contributes mainly to the information-seeking level of Jarvelin and Ingwersen's model of search. A study by Capra et al. [33] included the basic Endeca interface and, although this feature was not explicitly tested, the study involved both exploratory IS level and quick IR level tasks.

3.3.3 Connections

Networks have been used for knowledge discovery tasks, displaying connections between related documents and between related literatures. Stepping Stones (Figure 3.34) visualizes search results for a pair of

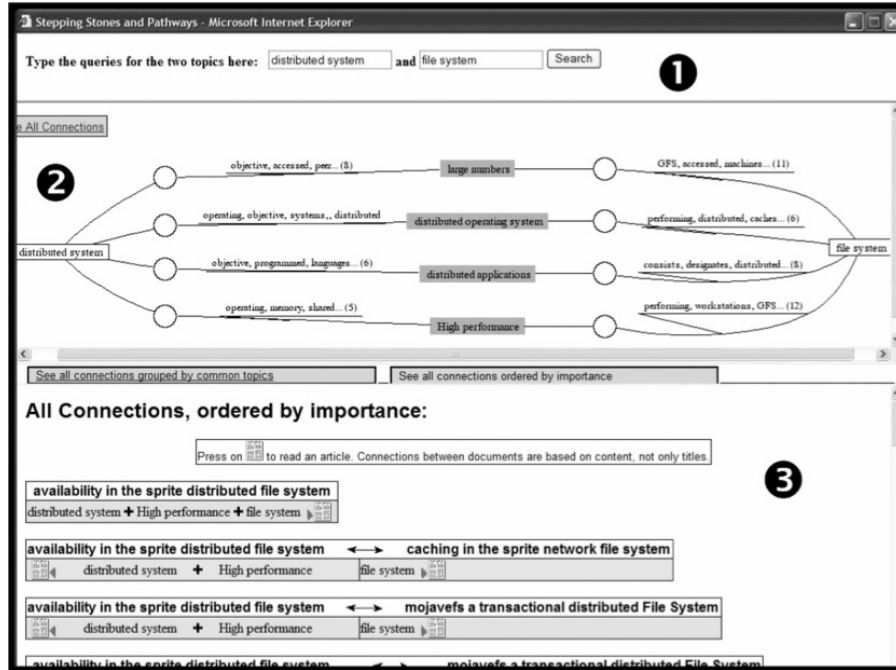


Fig. 3.34 The Stepping Stones interface visualizes the results of a pair of queries (1) as a graph (2) of topic nodes connected to the queries. Individual documents, shown in (3), cover portions of the path between the queries; image from [46].

queries, using a graph to show relationships between the two sets of results [46].

Similarly, Beale et al. [15] visualizes sequences of queries using a force-directed layout of node-link diagrams. Queries and their resulting documents are represented as nodes, with links between a query and its results. When a document is in the result set for multiple queries, it is linked to each query.

In some representations, single documents can appear in multiple collections or subcollections. When these different groups are arranged into regions that correspond to document attributes, often lines or colors are used to highlight related items. Such regions could be parts of a geographic map or, as in Figure 3.35 separates groups in a hierarchy of courts [91, 153]. Figure 3.35 shows a visualization where arcs that jump between the regions represent single items in different groups.

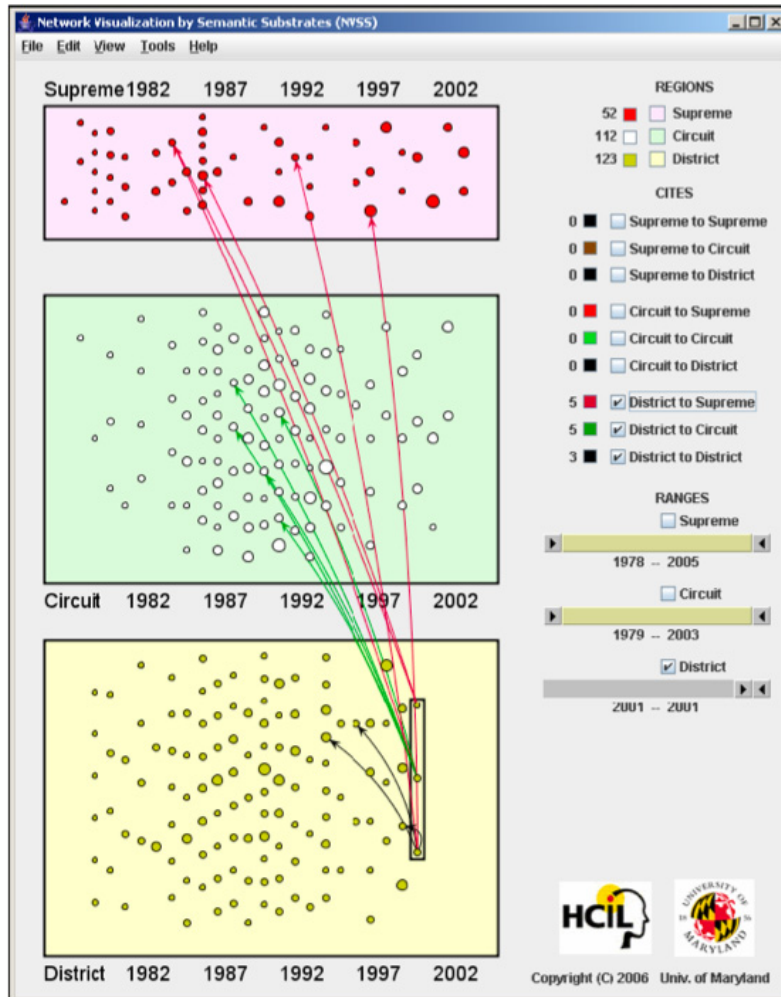


Fig. 3.35 This visualization arranges 287 court cases into three regions (Supreme Court, Circuit Court, and District Court). Within each region, cases are arranged chronologically. Of the 2032 citations, the selection box in the District region limits the display to just the citations from the three District court cases in 2001. This shows the greater importance of older Supreme Court opinions [153].

One challenge of representing links across multiple sets or dimensions is the limited number of methods to make connections. In Figure 3.35, the arcs are in various colors. Such use of colors and arrows can be limited. For example, it may be hard to display how a single item in

three different groups may be connected, or how an item relates to two parent levels of a hierarchy; such challenges become more important when hierarchical and faceted classification schemes are involved. In Continuum, relationships across multiple levels of hierarchy are shown by clustering and nesting results rather than keeping them in separate panes, see Figure 3.26.

Expressing and exploring connections provides much higher-level search interaction than simple IR activities, as the users are going beyond parsing results to assess how results fit into the domain of information, and thus the other results around them.

3.3.4 Animation

Animation on the Web is a debated area of research. Although lots of research has aimed at developing animation tools for the Web [59, 68], other research has investigated the limitations of technologies such as Flash [79]. Research by Robertson, Cameron, Czerwinski and Robbins, however, has shown that there are advantages of using animation to support users through transitions in an interface [141]. Further, recent research by Baudisch et al. [14] has presented a tool called Phosphor that highlights, with a fading period, any changes that occur in an interface. We can see examples of this sort of animated support on the Web through technologies such as AJAX [132]. Facebook, for example, allows users to post comments with AJAX, where the page does not refresh and both color and smooth spatial changes indicate to users that the interface is changing to accommodate their actions. Research into mSpace, which uses a lot of AJAX technology, also identified a similar emphasis on smooth transitions during discussions with participants [182].

The main emphasis on animation in web search visualization is on supporting users in comprehending changes that occur as a result of their actions. The Polyarchy browser [141], for example, helps users explore by animating the transition to alternative dimensions as the users browse across different attributes in the data set. Consequently, the advance provided by animation is mainly in the Information-seeking level of search, and has been mostly evaluated in kind.

3.3.5 Semantic Zooming

Semantic zooming is characterized by the progressive inclusion of additional information, as opposed to simply enlarging information. In the design of NaviQue [62], users can zoom into clusters and into information structures. The semantic zooming replaces items with representative icons as users zoom out, and replaces icons with the specific items when zooming in to show more detail. A history of all ad hoc collections is automatically maintained in one corner of the display. Users can drag and drop collections into a “pocket” in another corner so that they are immediately accessible, without need to pan or zoom the workspace.

As a form of animation, effective zooming on the Web is challenging and must be done to enhance a website and not affect its usability. In maintaining a typical search result list style, another option presented by research is a technique called WaveLens, which allows users to zoom in more on a single result, such that the information snippet expands to reveal more lines of text from the source document [131].

The contribution of semantic zooming supports users at IR and IS levels of search in that the technique allows users to investigate deeper into a topic as they “zoom” into it. The semantic zoom is considered a request for more detailed information on the item of focus, and is thus connected to more exploratory methods. Semantic zooming, however, can also simply reveal the answer to a quick keyword query, in which case it is related to selecting a web search result (as in the study of WaveLens) but without leaving the search engine.

3.3.6 Alternative Inputs

Multimedia can often be hard to describe with words, and querying over such collections requires that the documents be annotated. While some research is aimed at automatically increasing the amount of annotation of multimedia [25, 86], other approaches have examined the querying medium. Query-by-example [93, 144] is a strand of research, where users can submit audio to find related audio, images to find similar images, video to find similar video, etc. For example, Shazam²⁹ is a

²⁹<http://www.shazam.com>.

mobile phone service that users can ring while music is playing and responds by sending an SMS message with the name and artist of a song. Similarly, Retrievr³⁰ is a service that allows users to construct a simple image with paint tools, and finds images from Flickr³¹ that match the approximate shapes and colors. Although some research has discussed practical applications for systems such as query-by-humming, to search for music [99], the challenge for the Web is providing a means of input for users. Where Google provides a keyword box and Retrievr provides a sketch box with paint tools, it can become difficult to allow audio input, without requiring users to make recordings and upload them. Similar problems arise for video-querying for video.

3.4 Summary

In this Section, a range of diverse strategies to visualizing search results has been presented. Covering the use of enriched metadata (Section 3.1), multi-dimensional representations (Section 3.2) and visualization-enhancing techniques (e.g., animation in Section 3.3), a diverse range of approaches have been presented that try to support different contexts within the model of information-seeking described in Section 2. In the next section, we discuss the range of evaluation techniques that have been applied to these visualizations, in order to demonstrate their benefits for the different contexts of information seeking.

³⁰<http://labs.systemone.at/retrievr/>.

³¹<http://www.flickr.com>.

4

Evaluations of Search Systems

This section examines evaluation methods that have been used to assess different parts of information systems. Extensive reviews of information retrieval evaluations are provided by Hearst [72], which includes a discussion of formative evaluation and usability testing, and Yang [187]. In this section, we provide a synthesized view of techniques as appropriate to the three-level model. We break this evaluation discussion down into the three levels included in the framework from Section 2. The different levels of information-seeking context each require multiple, different, and complementary modes of evaluation. A single study may incorporate multiple levels of evaluation. In Section 5, as we classify the search interfaces described above by visualization approach and context of information seeking support, we also include the styles of evaluation within the taxonomy.

4.1 Information Retrieval Evaluations

The main concerns for much of the information retrieval community has been to assess the quality of indexing methods and the algorithms that match documents to the queries provided by users. The TREC conferences [70] have been arranged to evaluate document retrieval

systems for domains like the Web, spam, video, legal documents, and genomics; for collections with interesting or important characteristics such as very large collections; and for specialized tasks such as cross-language IR, filtering, finding novel documents, fact finding, and question answering.¹ Important concepts for evaluations at this level include document relevance, precision/recall and related measures, and batch versus interactive evaluation.

To support the TREC evaluations, predefined collections of documents, relevant queries and human relevance assessments, were produced and used as benchmark standards across any studies. The shared platform of evaluation provided the opportunity to not only evaluate, but also compete for most improved retrieval times or retrieval accuracy within the community. The most commonly used measurements in TREC were precision and recall. Precision is concerned with returning only relevant results, whereas recall measures the number of relevant documents being returned. Typically returning more documents means potentially sacrificing precision, and guaranteeing precision means reducing recall. Common approaches examine precision at certain levels (precision @ N), average precision, or by using precision–recall (P–R) curves [85]. Yang [187] and Kobayashi and Takeda [96] provide extensive reviews of evaluations focusing on the techniques and evaluations used for information retrieval.

When not testing the accuracy of results returned against a predefined corpus of documents, most IR studies, including many of the systems evaluated in the sections above, focus on simple measures of task performance such as speed for defining how well an interface supports users in finding information. A simple task, of finding a specific document, or a fact within a document, is provided to users along with a basic scenario for context. If users are able to find that answer quickly with one system than another, then the novel system has provided better support for search. This is a fairly accurate test for many basic search contexts. A basic fact-finding task represents many of the Web searches performed on search engines [156, 173], and so many studies are either based on this model or includes tasks of this nature.

¹<http://trec.nist.gov/tracks.html>.

Table 4.1. A sample of measures used at the IR evaluation level.

Measure	Reference
Precision, recall	[24, 34, 40]
Precision @ N	[3]
Mean average precision (MAP)	[3]
Expected search length	[52]
Average search length (ASL)	[112]
Cumulated gain, discounted cumulated gain	[85]
Relative Relevance, Ranked Half-Life	[26]
Query length	[83, 156]
Number of query refinements	[78, 83, 156]
Number of unique queries	[156, 173]
Number of operators in query	[156]
Number of results viewed	[78, 83, 156]
Search time	[173]

A summary of commonly used IR-level measures, along with references to example evaluations that have used them, is shown in Table 4.1.

It is clear from the model presented in Section 2, however, that information retrieval tasks are part of a much larger view of searching behavior. While research, such as that provided by White and Drucker indicate that many web searches are simple fact-finding tasks, the natural conclusion is that there are some search sessions that are not simple lookup tasks [28, 173]. The TREC conferences have incorporated the Interactive Track [51], the High Accuracy Retrieval of Documents (HARD) Track [3], and the Video Retrieval Evaluation (TRECVID) Track [158], which move beyond the batch-oriented evaluation of retrieval tasks by directly involving users conducting interactive searches in the evaluation process. However, the information needs in these tracks are still narrowly expressed in terms of documents to be retrieved, without reference to a higher-level information need. It is for this reason the methods of performing broader information-seeking evaluations is discussed in the next subsection.

4.2 Information-Seeking Evaluations

The evaluations of information retrieval typically have different aims than the evaluations of information seeking [94]. Information retrieval

evaluations have focused on system-oriented metrics, or the speed of simple, out of context, fact-finding tasks. Consequently, information retrieval studies have been criticized for a narrow conceptualization of the information need, relevance, and interaction [26]. Information-seeking research, therefore, focuses on evaluating systems for how they meet the needs of users. Important concepts at this level are the information need, the information-seeking context, information-seeking actions, tactics and strategies, and longitudinal changes, as are quantitative and qualitative measures.

To understand how search interfaces support a wide range of information-seeking tactics, recent work by Wilson and schraefel [183] and Wilson et al. [180, 184] has proposed an evaluation framework to systematically assess support for a range of known search tactics and types of users, called the Search Interface Inspector.² The framework first measures the functionality of the search system by the way it supports known tactics and moves employed with information [11, 12]. The framework then uses a novel mapping to summarize the measured support for the different user types [19], whose conditions vary on dimensions such as previous knowledge and intended use. This evaluation framework was later refined and validated to show that it could accurately predict the results of user studies [178]. With the confidence provided by the validation, the framework can be used to identify weaknesses in a system design or new function, so that it can be improved before user studies are carried out. Further, it can be used to inform the design of user studies, so that they accurately test the desired features.

One of the exploratory studies used to validate the framework produced by Wilson et al. was an information-seeking evaluation of faceted browsers [33]. The study used three types of tasks to evaluate the system, and performed a within participants study and a between participants study to get qualitative and quantitative results, respectively. The first type of task was a simple lookup task, which could be answered by using only one facet of the annotation. The second task type was a complex lookup, which involved multiple facets. The final task was exploratory, where users were asked to learn and produced a

²<http://mspace.fm/sii/>

summarized report about a certain topic. This third type of task is a good example of something that has not been included in Information Retrieval research, which, as mentioned above, has focused on matching documents to queries. Instead, by asking users to carry out learning tasks, we can assess the system for other types of information-seeking activities, such as comparison, synthesis, and summarization. As part of their discussion of higher-level problems encountered by users of information visualization systems, Amar and Stasko [7] discuss tasks that may be part of or require exploratory search.

Another contribution to information-seeking evaluations is on the discussion of time as a measurement by Capra et al. Although their tasks were timed, they suggest that time may not be a useful metric for exploratory tasks, as extended system use could mean that users have discovered increasing amounts of relevant information. In contrast to the information retrieval view that finding the answer quicker is more important, finishing an exploratory task early may indicate that a search system does not provide effective support for browsing. With this notion in mind, Kammerer et al. [90] conclude that the participants who used the MrTaggy interface (a) spent longer with their system, (b) had a higher cognitive load during search, and (c) produced better reports at the end of the tasks. In studies that aim to reduce cognitive load and reduce search time, however, there is usually only one correct answer. Conversely, during the MrTaggy experiment, the incentive was to produce a better report, and so a positive measure for the system was that it allows users to work harder.

The suitability of relevance in exploratory search conditions may also be in question for some information-seeking evaluations. In systems that use faceted classifications, for example, each document with a particular annotation has an equal weighting and thus every document suggested as a result of selecting a particular part of the classification will be equally relevant. Instead, Spink et al. [160, 161] have been designing a metric that tries to measure the progress of users in achieving their goal. Although designed for feedback to users, the rate of progress for similar tasks on different systems could be used to assess their support. Further, controlling for the amount of progress made by users in exploratory tasks would allow evaluators to once

again consider reduced time and cognitive load as positive measures. For example, evaluators could measure the time spent, and cognitive load as they make a pre-determined amount of progress.

Koshman [98] evaluated the VIBE (Visual Information Browsing Environment) prototype system, which graphically represents search results as geometric icons within one screen display. As part of understanding approaches to information seeking, the researchers sought to differentiate expert and novice performance through the use of a quasi-experimental within-participants design (see Borlund [26] or Shneiderman and Plaisant [155] for discussions of information-seeking evaluation).

Having discussed search tasks and measurements that may be unique to or important for information seeking, carefully controlled user studies can still be performed to evaluate systems in terms of information seeking. Käksi [89] provides a good example of a study that employs a within-subjects design, balanced task sets, time limitations, pre-formulated queries, cached result pages and limiting access to result documents.

Eye-tracking techniques have been used to study information-seeking behaviors within web search interfaces and library catalogs [45, 66, 104, 111]. These studies examined specific elements that searchers look at, how often they looked at them, for how long, and in what order. They provide insight into the cognitive activities that searchers undertake and tactics used when examining search results. For example, Kules et al. [104] studied gaze behavior in a faceted library catalog and determined that for exploratory search tasks, participants spent about half as much time looking at facets as they spent looking at the individual search results. The study also suggested that participants used the facets differently between the beginning and later stages of their searches. They conclude that the facets played an important role in the search process.

Finally, the study approach that investigates user interaction with software over a long period of time, such as longitudinal studies or studies that are repeated periodically with the same participants, provides a unique type of insight into realistic human behavior. One of the main arguments against the short task-oriented studies is the lack of

Table 4.2. A sample of measures used at the IS evaluation level. Some measures may be similar to IR level measures, but are used here to evaluate the process of IS tasks.

Measure	Reference
Number of moves/actions to support a search tactic	[177, 184]
Subjective measures (e.g., confidence, usefulness, usability, satisfaction)	[33, 98, 104, 106, 161]
Time on task	[33, 90, 98]
Time to select document	[66, 89]
Accuracy (correctness) and/or errors in task results	[33, 98, 111]
Number of results collected by user	[89, 90, 111]
Number of search terms	[89, 159, 170]
Number of queries	[106, 111, 160]
Number of results viewed	[160, 66, 111]
Time to learn	[98]
System feature retention	[98]
Rank of selected document in results	[89, 111, 106, 66]
Usage counts of selected interface features	[89, 170]
Gaze-related measures, e.g., number of fixations, location of fixations, rank of results fixated on, fixation duration; scanpaths, scanpath length	[45, 111, 104]
Number of search result pages viewed	[111]

realism. Further, the results of such user studies are often based on the participants' first or early responses to new designs, compared to their familiarity with existing software. By studying interaction over time, we can begin to evaluate changes in search tactics and subjective views as users adopt and adapt to new interfaces [102, 170, 182]. A summary of commonly used IS-level measures, along with references to example evaluations that have used them, is shown in Table 4.2.

4.3 Work-Context Evaluations

Evaluating search systems and the work task level requires measuring the success of work-context style problems. For example, Allen [4] investigated the interaction of spatial abilities with 2-D data representations. The work-context task given to participants was to read an article and then use the system to find a few good articles. At this level, many different information-seeking activities can be used, but ultimately the

system is assessed on its support for achieving the overall goal. As with evaluations at the information-seeking level, important concepts at this level include the information need, the information-seeking context, information-seeking actions, tactics, and strategies. At this level there is a stronger emphasis on domain-specific concepts. The quality of the work product also may be evaluated.

One of the challenges of evaluation, especially at the information seeking and work-context level, is that human participants interpret the tasks based on their own experiences and knowledge. There is a tension between the need to make results reliable and replicable and the need to make the task realistic for the participants. Borlund [26] advocated addressing this by incorporating participant-provided information needs into the experimental session along with a researcher-provided need. If the results for both tasks are consistent, the researcher can conclude that the researcher-provided task is realistic and thus reap the benefit of a realistic but tightly controlled task. Other research has used the previous actions of users to inform the realism of search tasks in system evaluations [56, 58]. Kules and Capra [103] propose a procedure for creating and validating tasks that exhibit exploratory characteristics, such as indicating uncertainty of the information need, the need for discovery, or being an unfamiliar domain for the searcher. The goal is to develop work tasks that are appropriate for the domain and system being evaluated and which are also comparable within and across studies.

Aside from the individuality of work task understanding, work-context tasks vary dramatically depending on the domain. Järvelin and Ingwersen [84] argue that the main challenge ahead for evaluating systems at the work-context level is that research needs to explicitly integrate context in terms of the high-level work task, the specific information-seeking task and the systems context. In particular, they note that the Web “is not a single coherent unit but appears quite different for different actors, tasks, and domains.” Some research has addressed domain-specific work-context evaluations. One comparative study attempted to rigorously evaluate work tasks by defining domain-specific measures of work product quality (assessing motivation, completeness, correctness, coherence, redundancy, and argument structure)

in addition to measuring efficiency, effectiveness, and precision of the search [88]. Subjects used three information retrieval systems to develop lesson materials about gorillas using a biological database. The goal was highly structured to create a lesson using the provided lesson template that prescribed topics to cover. Results showed significant system differences in efficiency and effectiveness but results for quality measures were mixed. They were not significant overall; however, significant differences were noted between individual sections of the lesson. Kules and Shneiderman [106] similarly evaluated the quality of the work product (ideas for newspaper articles) but found no significant differences. These two studies highlight the challenges of evaluating the contribution of a system to a high-level work task. Evaluators face the dilemma of trying to assess a system under conditions that are more realistic than at the IR or even IS level, while still effectively understanding the contribution that the system makes when it is only one of many factors. The use of longitudinal studies may help to overcome this challenge.

Qu and Furnas [137] used two variations of a sensemaking work task to motivate information seeking and topic organization tasks to study how people structure developing knowledge during an exploratory search. Study participants were given 50 minutes to collect and organize information about an unfamiliar topic and produce an outline of an oral presentation. The researchers examined sequences of actions that participants took, including issuing queries, bookmarking pages, and creating folders. They interpreted different sequences as indications

Table 4.3. A sample of measures used in studies focused work-context evaluation level. Some measures may also be used in the IS but provide insight into the process of achieving work-context tasks.

Measure	Reference
Number of records printed	[4]
Number of items viewed or printed	[4]
Number of items selected	[88]
Number of relevant items selected	[88]
Quality of results	[88, 106]
Rank of relevant documents	[56]
Participant assessed relevance	[56]
Sequences of actions	[137]

of different aspects of the sensemaking task. A summary of commonly used WC-level measures, along with references to example evaluations that have used them, is shown in Table 4.3.

4.4 Summary

Evaluations are most mature and rigorous at the IR level. At the IS level, evaluations are developing, with several good examples of studies that balance rigor and realism. The work level is where the biggest challenges remain. It is also where improvements in evaluation can contribute the most. Certainly, evaluations at this level need to reflect the nature of the domain and the tasks and information needs characteristic of that domain. Methodologies that enhance rigor and comparability without sacrificing validity are necessary. Longitudinal methods are likely to be useful. More generally, studies are likely to benefit from an emphasis on fewer, but deeper, tasks, reflecting the more complex nature of work tasks.

5

Taxonomy of Search Visualizations

So far in the monograph, we have discussed (1) a model of search that captures multiple levels of context, (2) many novel advances in search result visualization, and (3) the way in which such advances have been evaluated. Below we present and discuss a taxonomy of these search advances, according to their support for search, amount of evaluation, and prevalence on the Web.

5.1 The Taxonomy

We present a taxonomy of the search visualization advances discussed in this monograph (Table 5.1). The purpose of the taxonomy is to capture (1) how these advances are designed to support search, (2) to what extent they have been evaluated, and (3) how prevalent they are on the Web. The next subsection discusses the value of this information to academics and designers of future web search interfaces.

The first facet, shown in the first set of three columns, is the level of search context, as according to the model used throughout the monograph, that is being supported by the interface feature. Although each

Table 5.1. Interface advances discussed within the monograph above, categorized by three facets: Level of Search Support, Depth of Evaluation, and Prevalence on the Web.

Interface Advances	Level of Search Support			Depth of Evaluation			Prevalence Online		
	Information Retrieval	Information Seeking	Work Context	Initial Testing	Multiple User Studies	Diverse Forms of Study	Demonstrator Only	Many Instances Available	Widely Used
3.1.1 Hierarchical Classifications									
1. Standard Web Directories	O			O					O
2. Cha-Cha	O			O			O		
3. WebTOC		O		O			O		
3.1.2 Faceted Classifications									
4. Flamenco		O				O	O		
5. mSpace			O			O	O		
6. Exhibit		O		O					
7. Relation Browser RB+++		O			O		O		
8. SERVICE		O		O			O		
9. Dyna-Cat		O		O			O		
10. FacetPatch (Contextual facets)		O		O			O		
3.1.3 Automatic Clustering									
11. Clusty		O			O			O	
3.1.4 Social Classifications									
12. Social Rankings		O						O	
13. Tag Clouds		O		O					O
14. MrTaggy		O		O			O		
3.2.1 Result Lists									
15. Page Thumbnails		O			O			O	
16. Keywords in Context		O			O				O
17. Larger Result Snippets		O			O		O		

(Continued)

Table 5.1. (Continued)

Interface Advances	Level of Search Support			Depth of Evaluation			Prevalence Online		
	Information Retrieval	Information Seeking	Work Context	Initial Testing	Multiple User Studies	Diverse Forms of Study	Demonstrator Only	Many Instances Available	Widely Used
3.2.2 2D Result Representations									
18. SOMS		○				○		○	
19. Treemaps		○				○		○	
20. Hyperbolic Trees		○			○		○		
21. Scatterplot/Matrix		○				○		○	
22. Euler Circles		○			○		○		
23. Geographical Plots		○			○		○		
24. Timelines		○			○		○		○
25. Per-Query-Term Relevance		○			○		○		
26. User-Organized Canvases			○		○		○		
3.2.3 3D Result Representations									
27. User-Organized 3D Spaces			○				○		
28. 3D Hyperbolic Trees		○					○		
29. Cone Trees		○			○		○		
30. 3D VR Exploration		○					○		
3.3 Additional Functions									
31. Multimedia Previews		○					○		
32. Content Summarization		○					○		
33. Similar Results		○					○		○
34. Results Connections		○					○		
35. Animated Transitions		○					○		○
36. Semantic Zooming		○					○		○
37. Query-by-Example			○				○		

interface technique may support, to some extent, each of the three levels, they have been allocated to the level at which they are primarily designed to support. In fact, a lengthier alternative analysis of these interface advances could discuss the way in which the advances all contribute to each level of search context.

The second facet, shown in the second set of three columns, represents the amount of study that the techniques have received. While these could be categorized into groups such as formative studies, empirical user studies, longitudinal log analyses, etc. we are more interested, here, to learn the diversity of evidence that has been produced. They could have also been categorized by whether they have been studied in the same dimensions as they have been designed to support. Many studies, however, include information retrieval and information-seeking tasks, and so they may have received relatively little study, but in both contexts. Some techniques have received little published study and evaluation, like allowing users to view similar results to any one result returned [#33 in taxonomy], and so our understanding of their benefits is through experience and intuition. Other techniques, such as treemaps and SOMs [#18, #19], have received much evaluation and optimization since they were first proposed, including initial testing, carefully constructed lab studies, and analyses of working online deployments.

The third facet, contained in the final set of three columns, captures how prevalent visualizations have become on the Web. The items familiar in most search engines, such as “similar results” and “keyword-in-context result snippets”, are dominant, regardless of how well they have been studied, or to what level of context they support searchers. Other techniques, however, have been well studied but are not yet in widespread use, such as faceted browsing on general Web search engines [#8, #9]. It might be noted that the most common model of faceted browsing, represented by Flamenco [#4], is more prevalent on the Web than some of the alternatives. Google product search,¹ for example, provides this kind of typical faceted interaction to filter the results by price ranges and even specific vendors.

¹<http://www.google.com/products>.

These three facets, in combination, provide insights into the advances that have been made in web search visualizations. The diagram shows, for example, advances that are highly prevalent on the Web, but have received little published evaluation, and only support a low information retrieval context of search, such as web directories [#1]. Similarly, and perhaps more important for some readers, we can see heavily studied advances that are not yet prevalent on the Web, such as the faceted model of search provided by mSpace [#5].

5.2 Using the Taxonomy

There are two main audiences for the taxonomy above: (1) search interface designers and (2) academics working on future advances in this area. Similarly, the content of the taxonomy and the gaps in the diagram provide useful information to designers and academics. These uses are discussed below.

One challenge in building web search interfaces is in choosing the best combination of tools and features that will best support the target audience, if known. Clearly, there are many advances that can be used, but using lots, or even all of them, would provide a cluttered and perhaps unusable interface. For designers working with such a challenge, the taxonomy acts as a guide to identifying potential appropriate options. Further, the gaps in the diagram highlight areas where new and novel ideas might make a distinct interface and provide a business edge.

For academics, the taxonomy provides two types of information. First, much prior art is discussed in this article and included in the diagram, which can be used to help identify related work. Second, and more importantly, however, the diagram quite clearly marks under-researched ideas and research gaps. One conclusion that can be drawn from the diagram is that there has been a lot of work on information-seeking level advances, but comparatively little work that has focused on work-contexts. Section 4.3 discusses some studies that have focused on work context scenarios, and some recent work has focused on tasks such as booking holidays, writing papers, and even collaborating with other people [125]. For the many existing advances that are captured by

the taxonomy, however, seeing how different techniques relate to each other across the captured dimensions raises some un-answered research questions. Academics may want to compare well studied, but less popular advances, with those that are popular despite having received little published evaluation.

6

Conclusions

This monograph presented four steps in discussing the advances in visualizing web search results. First, we presented previous and related work, which was followed by a model of search that has been used throughout the document. This model of search includes multiple levels of context from basic Information Retrieval needs, to more exploratory information-seeking behavior, and finally the context of the tasks being completed by users, known as the work context. Second, we offered a wide range of innovations in web search interfaces, while discussing how each makes contributions to the three levels of search context. Third, Section 4 discussed the techniques that have been used to evaluate these advances at the three different levels of search context. We discussed the increasing difficulty that evaluators experience while studying the contributions of new designs at higher-level work contexts, as regular simple measures such as time and accuracy of specific low-level tasks do not necessarily apply. Finally, we presented a taxonomy in Section 5 that captures the interface advances. The taxonomy captures (1) which level of search context the advance primarily support, (2) how much study the advances have received, and (3) how prevalent the advances are on the Web.

As well as supporting readers in finding much of the prior art that exists in web search result visualization, this monograph and taxonomy helps the designers of future search systems to make informed choices about which advances may be appropriate for their system and audience. Further, the taxonomy helps identify under-researched ideas and research gaps in web search result visualization. Notably, there have been far fewer advances that consider the higher-level work contexts of searchers. Consequently we advocate that, in order to better understand novel ideas, future evaluations focus on the work contexts of users, and in supporting them to achieve their higher-level goals.

We began this monograph by celebrating the success of keyword search, while clarifying when alternative modes of search are needed. First, as users' demands continue to grow and their needs evolve, opportunities are emerging for exploratory search strategies. Often users want to get answers to their questions, not just web pages. Second, their questions are increasingly complex and may take hours or weeks to resolve. A third change is the move toward new Semantic Web technologies and related strategies, which offer improved possibilities for machine-assisted resolution for complex queries, especially with keywords which have multiple meanings. We have shown that there are emerging innovative user interfaces and visual presentations that may help users achieve their goals more rapidly and with greater confidence in the validity of the results. Through empirical studies and log analysis, researchers are coming to better understand the ways in which people search, the tasks they have, and how they collaborate with their colleagues. This survey monograph provides a resource to the designers and researchers who are developing search systems so that they can more make more informed and confident design decisions as they create novel and effective interfaces.

Acknowledgments

The authors thank Marti Hearst, those who provided images, and several others for providing feedback and comments.

References

- [1] F. J. Aguilar, *General Managers in Action*. New York, NY: Oxford University Press, 1988.
- [2] H. Alani and C. Brewster, "Ontology ranking based on the analysis of concept structures," in *Proceedings of the 3rd International Conference on Knowledge Capture*, pp. 51–58, New York, NY, USA: ACM Press, 2005.
- [3] J. Allan, "Hard track overview in trec 2003 high accuracy retrieval from documents," in *Proceedings of the Text Retrieval Conference*, pp. 24–37, 2003.
- [4] B. Allen, "Information space representation in interactive systems: Relationship to spatial abilities," in *Proceedings of the Third ACM Conference on Digital Libraries*, pp. 1–10, Pittsburgh, Pennsylvania, United States: ACM Press, 1998.
- [5] R. Allen, "Two digital library interfaces that exploit hierarchical structure," in *Proceedings of Electronic Publishing and the Information Superhighway*, pp. 134–141, Boston, MA, USA, 1995.
- [6] J. Allsop, *Microformats: Empowering Your Markup for Web 2.0*. friends of ED, 2007.
- [7] R. Amar and J. Stasko, "A knowledge task-based framework for design and evaluation of information visualizations," in *Proceedings of the IEEE Symposium on Information Visualization*, pp. 143–150, Austin, Texas, USA: IEEE Computer Society, 2004.
- [8] B. Amento, W. Hill, L. Terveen, D. Hix, and P. Ju, "An empirical evaluation of user interfaces for topic management of web sites," in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pp. 552–559, Pittsburgh, Pennsylvania, United States: ACM Press, 1999.

- [9] P. André, M. L. Wilson, A. Russell, D. A. Smith, A. Owens, and m. c. schraefel, "Continuum: Designing timelines for hierarchies, relationships and scale," in *Proceedings of the 20th Annual ACM Symposium on User Interface Software and Technology*, pp. 101–110, Newport, Rhode Island, USA: ACM Press, 2007.
- [10] P. Au, M. Carey, S. Sewraz, Y. Guo, and S. Rüger, "New paradigms in information visualization," in *Proceedings of the 23rd annual international ACM SIGIR conference on Research and Development in Information Retrieval*, pp. 307–309, Athens, Greece: ACM Press, 2000.
- [11] M. J. Bates, "Idea tactics," *Journal of the American Society for Information Science*, vol. 30, no. 5, pp. 280–289, 1979.
- [12] M. J. Bates, "Information search tactics," *Journal of the American Society for Information Science*, vol. 30, no. 4, pp. 205–214, 1979.
- [13] M. J. Bates, "The design of browsing and berrypicking techniques for the online search interface," *Online Review*, vol. 13, no. 5, pp. 407–424, 1989.
- [14] P. Baudisch, D. Tan, M. Collomb, D. Robbins, K. Hinckley, M. Agrawala, S. Zhao, and G. Ramos, "Phosphor: Explaining transitions in the user interface using afterglow effects," in *Proceedings of the 19th annual ACM symposium on User Interface Software and Technology*, pp. 169–178, Montreux, Switzerland: ACM Press, 2006.
- [15] R. Beale, R. J. McNab, and I. H. Witten, "Visualising sequences of queries: A new tool for information retrieval," in *Proceedings of the IEEE Conference on Information Visualisation*, pp. 57–63, Phoenix, AZ, USA: IEEE Computer Society, 1997.
- [16] A. Becks, C. Seeling, and R. Minkenbergh, "Benefits of document maps for text access in knowledge management: A comparative study," in *Proceedings of the ACM Symposium on Applied Computing*, pp. 621–626, Madrid, Spain: ACM Press, 2002.
- [17] S. M. Beitzel, E. C. Jensen, A. Chowdhury, D. Grossman, and O. Frieder, "Hourly analysis of a very large topically categorized web query log," in *Proceedings of the 27th annual international ACM SIGIR conference on Research and Development in Information Retrieval*, pp. 321–328, Sheffield, UK: ACM Press, 2004.
- [18] N. J. Belkin, C. Cool, A. Stein, and U. Thiel, "Cases, scripts, and information-seeking strategies: On the design of interactive information retrieval systems," *Expert Systems with Applications*, vol. 9, no. 3, pp. 379–395, 1995.
- [19] N. J. Belkin, P. G. Marchetti, and C. Cool, "Braque: Design of an interface to support user interaction in information retrieval," *Information Processing & Management*, vol. 29, no. 3, pp. 325–344, 1993.
- [20] S. Benford, I. Taylor, D. Brailsford, B. Koleva, M. Craven, M. Fraser, G. Reynard, and C. Greenhalgh, "Three dimensional visualization of the world wide web," *ACM Computing Surveys*, vol. 31, no. 4, p. 25, 1999.
- [21] T. Berners-Lee, Y. Chen, L. Chilton, D. Connolly, R. Dhanaraj, J. Hollenbach, A. Lerer, and D. Sheets, "Tabulator: Exploring and analyzing linked data on the semantic web," in *Proceedings of the 3rd International Semantic Web User Interaction Workshop*, 2006.

- [22] S. K. Bhavnani, "Important cognitive components of domain-specific search knowledge," in *NIST Special Publication 500-250: The Tenth Text Retrieval Conference*, (E. M. Vorhees and D. K. Harman, eds.), Washington, DC: NIST, 2001.
- [23] S. K. Bhavnani and M. J. Bates, "Separating the knowledge layers: Cognitive analysis of search knowledge through hierarchical goal decompositions," in *Proceedings of the American Society for Information Science and Technology Annual Meeting*, pp. 204–213, Medford, NJ: Information Today, 2002.
- [24] D. C. Blair and M. E. Maron, "An evaluation of retrieval effectiveness for a full-text document retrieval system," *Communications of the ACM*, vol. 28, no. 3, pp. 289–299, 1985.
- [25] S. Bloehdorn, K. Petridis, C. Saathoff, N. Simou, V. Tzouvaras, Y. Avrithis, S. Handschuh, I. Kompatsiaris, S. Staab, and M. G. Strintzis, "Semantic annotation of images and videos for multimedia analysis," in *Proceedings of the 2nd European Semantic Web Conference*, pp. 592–607, Heraklion, Crete, Greece: Springer, 2005.
- [26] P. Borlund, "The IIR evaluation model: A framework for evaluation of interactive information retrieval systems," *Information Research*, vol. 8, no. 3, p. Paper 152, 2003.
- [27] K. Börner, "Visible threads: A smart vr interface to digital libraries," in *Proceedings of IST/SPIE's 12th Annual International Symposium on Visual Data Exploration and Analysis*, pp. 228–237, San Jose, CA, USA: SPIE, 2000.
- [28] A. Broder, "A taxonomy of web search," *ACM SIGIR Forum*, vol. 36, no. 2, pp. 3–10, 2002.
- [29] K. Byström and P. Hansen, "Work tasks as units for analysis in information seeking and retrieval studies," in *Emerging Frameworks and Methods*, (H. Bruce, R. Fidel, P. Ingwersen, and P. Vakkari, eds.), Greenwood Village, CO: Libraries Unlimited, 2002.
- [30] K. Byström and P. Hansen, "Conceptual framework for tasks in information studies: Book reviews," *Journal American Society Information Science and Technology*, vol. 56, no. 10, pp. 1050–1061, 2005.
- [31] G. Cai, "Geovibe: A visual interface for geographic digital libraries," in *Proceedings of the First Visual Interfaces to Digital Libraries Workshop*, (K. Borner and C. Chen, eds.), pp. 171–187, Roanoke, VA, USA: Springer-Verlag, 2002.
- [32] R. Capra and G. Marchionini, "The Relation Browser tool for faceted exploratory search," in *Proceedings of the 2008 Conference on Digital Libraries (JCDL '08)*, Pittsburgh, Pennsylvania, June 2008.
- [33] R. Capra, G. Marchionini, J. S. Oh, F. Stutzman, and Y. Zhang, "Effects of structure and interaction style on distinct search tasks," in *Proceedings of the ACM/IEEE-CS Joint Conference on Digital Libraries*, pp. 442–451, Vancouver, British Columbia, Canada: ACM Press, 2007.
- [34] S. Chakrabarti, B. Dom, D. Gibson, S. Kumar, P. Raghavan, S. Rajagopalan, and A. Tomkins, "Experiments in topic distillation," in *ACM SIGIR workshop on Hypertext Information Retrieval on the Web*, pp. 13–21, Melbourne, Australia, 1998.

- [35] C. Chen, "Citespace II: Detecting and visualizing emerging trends and transient patterns in scientific literature," *Journal of the American Society for Information Science and Technology*, vol. 57, no. 3, pp. 359–377, 2006.
- [36] H. Chen and S. Dumais, "Bringing order to the web: Automatically categorizing search results," in *Proceedings of the SIGCHI conference on Human factors in computing systems*, pp. 145–152, The Hague, The Netherlands: ACM Press, 2000.
- [37] H. Chen, A. L. Houston, R. R. Sewell, and B. R. Schatz, "Internet browsing and searching: User evaluations of category map and concept space techniques," *Journal of the American Society for Information Science*, vol. 49, no. 7, pp. 582–608, 1998.
- [38] M. Chen, M. Hearst, J. Hong, and J. Lin, "Cha-cha: A system for organizing intranet search results," in *Proceedings of the 2nd USENIX Symposium on Internet Technologies and Systems*, Boulder, CO, 1999.
- [39] C. W. Choo, B. Detlor, and D. Turnbull, *Web Work: Information Seeking and Knowledge Work on the World Wide Web*. Dordrecht, The Netherlands: Kluwer Academic Publishers, 2000.
- [40] C. W. Cleverdon, J. Mills, and M. Keen, "Factors determining the performance of indexing systems," in *ASLIB Cranfield Project*, Cranfield, 1966.
- [41] A. Cockburn and B. Mckenzie, "An evaluation of cone trees," in *Proceedings of the British Computer Society Conference on Human-Computer Interaction*, Sheffield, UK: Springer-Verlag, 2000.
- [42] A. Cockburn and B. Mckenzie, "Evaluating the effectiveness of spatial memory in 2D and 3D physical and virtual environments," in *Proceedings of the SIGCHI conference on Human factors in computing systems: Changing our world, changing ourselves*, pp. 203–210, Minneapolis, Minnesota, USA, 2002.
- [43] W. S. Cooper, "On selecting a measure of retrieval effectiveness," *Journal of the American Society for Information Science*, vol. 24, no. 2, pp. 87–100, 1973.
- [44] J. Cugini, C. Piatko, and S. Laskowski, "Interactive 3D visualization for document retrieval," in *Proceedings of the Workshop on New Paradigms in Information Visualization and Manipulation, ACM Conference on Information and Knowledge Management*, Rockville, Maryland, USA, 1996.
- [45] E. Cutrell and Z. Guan, "What are you looking for?: An eye-tracking study of information usage in web search," in *Proceedings of the SIGCHI conference on Human factors in computing systems*, pp. 407–416, San Jose, California, USA: ACM press, 2007.
- [46] F. Das-Neves, E. A. Fox, and X. Yu, "Connecting topics in document collections with stepping stones and pathways," in *Proceedings of the 14th ACM International Conference on Information and Knowledge Management*, pp. 91–98, Bremen, Germany: ACM Press, 2005.
- [47] L. Ding, A. J. T. Finin, R. Pan, R. Cost, Y. Peng, P. Reddivari, V. Doshi, and J. Sachs, "Swoogle: A search and metadata engine for the semantic web," in *Proceedings of the thirteenth ACM international conference on Information and knowledge management*, pp. 652–659, Washington DC, USA, ACM New York, NY, USA, 2004.

- [48] K. Doan, C. Plaisant, and B. Shneiderman, "Query previews in networked information systems," in *Proceedings of the Third Forum on Research and Technology Advances in Digital Libraries*, pp. 120–129, Washington, DC, USA: IEEE Computer Society, 1996.
- [49] O. Drori and N. Alon, "Using documents classification for displaying search results list," *Journal of Information Science*, vol. 29, no. 2, pp. 97–106, 2003.
- [50] S. Dumais, E. Cutrell, and H. Chen, "Optimizing search by showing results in context," in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, Seattle, WA*, New York: ACM Press, 2001.
- [51] S. T. Dumais and N. J. Belkin, "The TREC interactive tracks: Putting the user into search," in *TREC: Experiment and Evaluation in Information Retrieval*, (E. Voorhees and D. Harman, eds.), MIT Press, 2005.
- [52] M. D. Dunlop, "Time, relevance and interaction modelling for information retrieval," *SIGIR Forum*, vol. 31, no. SI, pp. 206–213, 1997.
- [53] C. Eaton and H. Zhao, Visualizing Web Search Results. 2001.
- [54] M. Efron, J. Elsas, G. Marchionini, and J. Zhang, "Machine learning for information architecture in a large governmental website," in *Proceedings of the 4th ACM/IEEE-CS Joint Conference on Digital libraries*, pp. 151–159, Tuscon, AZ, USA: ACM Press, 2004.
- [55] D. E. Egan, J. R. Remde, L. M. Gomez, T. K. Landauer, J. Eberhardt, and C. C. Lochbaum, "Formative design evaluation of superbok," *ACM Transactions on Information Systems*, vol. 7, no. 1, pp. 30–57, 1989.
- [56] S. Elbassuoni, J. Luxenburger, and G. Weikum, "Adaptive personalization of web search," in *Workshop on Web Information Seeking and Interaction at SIGIR*, pp. 23–30, Amsterdam, Netherlands, 2007.
- [57] D. Ellis, "A behavioral model for information retrieval system design," *Journal of Information Science*, vol. 15, no. 4–5, pp. 237–247, 1989.
- [58] D. Elweiler and I. Ruthven, "Towards task-based personal information management evaluations," in *Proceedings of the 30th annual international ACM SIGIR conference on Research and Development in Information Retrieval*, pp. 23–30, Amsterdam, The Netherlands, 2007.
- [59] P. Faraday and A. Sutcliffe, "Authoring animated web pages using "contact points",," in *Proceedings of the SIGCHI conference on Human factors in computing systems: The CHI is the limit*, pp. 458–465, Pittsburgh, PA, USA: ACM Press, 1999.
- [60] Y. Feng and K. Borner, "Using semantic treemaps to categorize and visualize bookmark files," in *Proceedings of SPIE — Volume 4665, Visualization and Data Analysis 2002*, (R. F. Erbacher, P. C. Chen, M. Groehn, J. C. Roberts, and C. M. Wittenbrink, eds.), Bellingham, WA, USA: SPIE–The International Society for Optical Engineering, 2002.
- [61] V. Florance and G. Marchionini, "Information processing in the context of medical care," in *Proceedings of the 18th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval*, pp. 158–163, Seattle, WA, USA: ACM Press, 1995.

- [62] G. W. Furnas and S. J. Rauch, "Considerations for information environments and the navique workspace," in *Proceedings of the Third ACM Conference on Digital libraries*, pp. 79–88, Pittsburgh, PA, USA: ACM Press, 1998.
- [63] M. Gao, C. Liu, and F. Chen, "An ontology search engine based on semantic analysis," in *ICITA '05. Third International Conference on Information Technology and Applications, 2005*, pp. 256–259, IEEE, 2005.
- [64] M. Ginsburg, "Visualizing digital libraries with open standards," *Communications of the Association for Information Systems*, vol. 13, pp. 336–356, 2004.
- [65] R. Godwin-Jones, "Emerging technologies tag clouds in the blogosphere: Electronic literacy and social networking," *Language Learning and Technology*, vol. 10, no. 2, pp. 8–15, 2006.
- [66] L. A. Granka, T. Joachims, and G. Gay, "Eye-tracking analysis of user behavior in WWW search," in *Proceedings of the 27th annual international ACM SIGIR conference on Research and development in information retrieval*, Sheffield, United Kingdom: ACM Press, 2004.
- [67] A. Gulli and A. Signorini, "The indexable web is more than 11.5 billion pages," in *Proceedings of the 14th International Conference on the World Wide Web*, pp. 902–903, Chiba, Japan: ACM Press, 2005.
- [68] J. Haaajanen, M. Pesonius, E. Sutinen, J. Tarhio, T. Terasvirta, and P. Vanninen, "Animation of user algorithms on the web," in *Proceedings of IEEE Symposium on Visual Languages*, pp. 356–363, Isle of Capri, Italy: IEEE Computer Society, 1997.
- [69] W. Hall and K. O'Hara, "Semantic web," in *Encyclopedia of Complexity and Systems Science*, (R. A. Meyers, ed.), Springer, 2009.
- [70] D. K. Harman, "The TREC conferences," in *Readings on information retrieval*, San Francisco, CA, USA: Morgan Kaufmann Publishers Inc, 1997.
- [71] M. Hearst, "The use of categories and clusters for organizing retrieval results," in *Natural Language Information Retrieval*, (T. Strzalkowski, ed.), Boston: Kluwer Academic Publishers, 1999.
- [72] M. Hearst, *Search User Interfaces*. Cambridge University Press, 2009.
- [73] M. Hearst, A. Elliot, J. English, R. Sinha, K. Swearingen, and P. Yee, "Finding the flow in web site search," *Communications of the ACM*, vol. 45, no. 9, pp. 42–49, 2002.
- [74] M. Hearst and J. Pedersen, "Reexamining the cluster hypothesis: Scatter/gather on retrieval results," in *Proceedings of the 19th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval, Zurich, Switzerland*, New York: ACM Press, 1996.
- [75] M. A. Hearst, "Next generation web search: Setting our sites," *IEEE Data Engineering Bulletin: Special Issue on Next Generation Web Search*, vol. 23, no. 3, pp. 38–48, 2000.
- [76] M. Hildebrand, J. V. Ossenbruggen, and L. Hardman, "/facet: A browser for heterogeneous semantic web repositories," in *Proceedings of the 5th International Conference on the Semantic Web (ISWC'06)*, pp. 272–285, Athens, GA, USA, 2006.

- [77] O. Hoerber and X. D. Yang, “The visual exploration of web search results using hotmap,” in *Proceedings of the International Conference on Information Visualization*, pp. 272–285, London, UK: IEEE Computer Society, 2006.
- [78] C. Holscher and G. Strube, “Web search behavior of internet experts and newbies,” in *Proceedings of the 9th International WWW Conference*, pp. 157–165, 2000.
- [79] A. Holzinger and M. Ebner, “Interaction and usability of simulations and animations: A case study of the flash technology,” in *Proceedings of IFIP TC13 International Conference on Human Computer Interaction*, pp. 777–780, Zürich, Switzerland: IOS Press, 2003.
- [80] A. Hotho, R. Jaschke, C. Schmitz, and G. Stumme, “Information retrieval in folksonomies: Search and ranking,” *The Semantic Web: Research and Applications*, vol. 4011, pp. 411–426, 2006.
- [81] D. F. Huynh, R. Miller, and D. Karger, “Exhibit: Lightweight structured data publishing,” in *Proceedings of the World Wide Web Conference*, pp. 737–746, Banff, Alberta, Canada: ACM Press, 2007.
- [82] B. Jansen and A. Spink, “An analysis of web information seeking and use: Documents retrieved versus documents viewed,” in *Proceedings of the 4th international conference on Internet computing*, pp. 65–69, Las Vegas, NV, USA, 2003.
- [83] B. Jansen, A. Spink, J. Bateman, and T. Saracevic, “Real life information retrieval: A study of user queries on the web,” *ACM SIGIR Forum*, vol. 32, no. 1, pp. 5–17, 1998.
- [84] K. Järvelin and P. Ingwersen, “Information seeking research needs extension towards tasks and technology,” *Information Research*, vol. 10, no. 1, p. paper 212, 2004.
- [85] K. Jarvelin and J. Kekalainen, “IR evaluation methods for retrieving highly relevant documents,” in *Proceedings of the 23rd annual international ACM SIGIR conference on Research and development in information retrieval*, pp. 41–48, Athens, Greece, ACM, 2000.
- [86] J. Jeon, V. Lavrenko, and R. Manmatha, “Automatic image annotation and retrieval using cross-media relevance models,” in *Proceedings of the 26th annual international ACM SIGIR conference on Research and development in informaion retrieval*, pp. 119–126, Toronto, Canada: ACM Press, 2003.
- [87] T. Joachims, *Text Categorization with Support Vector Machines: Learning with Many Relevant Features*. Springer, 1997.
- [88] S. Kabel, R. De Hoog, B. J. Wielinga, and A. Anjewierden, “The added value of task and ontology-based markup for information retrieval,” *Journal of the American Society for Information Science and Technology*, vol. 55, no. 4, pp. 348–362, 2004.
- [89] M. Käkki, “Findex: Search result categories help users when document ranking fails,” in *Proceeding of the SIGCHI Conference on Human Factors in Computing Systems*, pp. 131–140, Portland, OR, USA: ACM Press, 2005.
- [90] Y. Kammerer, R. Narin, P. Pirolli, and E. Chi, “Signpost from the masses: Learning effects in an exploratory social tag search browser,” in *Proceedings*

- of the 27th international conference on Human factors in computing systems, pp. 625–634, Boston, MA: ACM Press, 2009.
- [91] H. Kang and B. Shneiderman, “Exploring personal media: A spatial interface supporting user-defined semantic regions,” *Journal of Visual Languages & Computing*, vol. 17, no. 3, pp. 254–283, 2006.
- [92] J. Kari, “Evolutionary information seeking: A case study of personal development and internet searching,” *First Monday*, vol. 11, no. 1, 2006.
- [93] T. Kato, T. Kurita, N. Otsu, and K. Hirata, “A sketch retrieval method for full color image database-query byvisual example,” in *Proceedings of 11th IAPR International Conference on Pattern Recognition*, pp. 530–533, The Hague, The Netherlands: IEEE Computer Society, 1992.
- [94] D. Kelly, S. Dumais, and J. Pedersen, “Evaluation challenges and directions for information-seeking support systems,” *Computer*, vol. 42, no. 3, pp. 60–66, 2009.
- [95] A. Kerne, E. Koh, B. Dworaczyk, J. M. Mistrot, H. Choi, M. S. Smith, R. Graeber, D. Caruso, A. Webb, R. Hill, and J. Albea, “Combinformation: A mixed-initiative system for representing collections as compositions of image and text surrogates,” in *Proceedings of the 6th ACM/IEEE-CS joint conference on Digital libraries*, pp. 11–20, Chapel Hill, NC, USA: ACM Press, 2006.
- [96] M. Kobayashi and K. Takeda, “Information retrieval on the web,” *ACM Computing Surveys*, vol. 32, no. 2, pp. 144–173, 2000.
- [97] T. Kohonen, *Self-Organizing Maps*. Springer, 2001.
- [98] S. Koshman, “Testing user interaction with a prototype visualization-based information retrieval system,” *Journal of the American Society for Information Science and Technology*, vol. 56, no. 8, pp. 824–833, 2005.
- [99] N. Kosugi, Y. Nishihara, T. Sakata, M. Yamamuro, and K. Kushima, “A practical query-by-humming system for a large music database,” in *Proceedings of the eighth ACM international conference on Multimedia*, pp. 333–342, New York, NY, USA: ACM Press, 2000.
- [100] B. Kotelly, “Resonance: Introducing the concept of penalty-free deep look ahead with dynamic summarization of arbitrary results sets,” in *Workshop on Human-Computer Interaction and Information Retrieval*, Cambridge Massachusetts, USA: MIT CSAIL, 2007.
- [101] C. C. Kuhlthau, “Inside the search process: Information seeking from the user’s perspective,” *Journal of the American Society for Information Science*, vol. 42, no. 5, pp. 361–371, 1991.
- [102] B. Kules, “Methods for evaluating changes in search tactics induced by exploratory search systems,” in *ACM SIGIR 2006 Workshop on Evaluating Exploratory Search Systems*, Seattle, WA, 2006.
- [103] B. Kules and R. Capra, “Creating exploratory tasks for a faceted search interface,” in *Second Workshop on Human-Computer Interaction and Information Retrieval (HCIR’08)*, Seattle, WA, USA, 2008.
- [104] B. Kules, R. Capra, M. Banta, and T. Sierra, “What do exploratory searchers look at in a faceted search interface?,” in *Proceedings of the 9th ACM/IEEE-CS joint conference on Digital libraries*, pp. 313–322, Austin, TX, USA: ACM Press, 2009.

- [105] B. Kules, J. Kustanowitz, and B. Shneiderman, "Categorizing web search results into meaningful and stable categories using fast-feature techniques," in *Proceedings of the Sixth ACM/IEEE-CS Joint Conference on Digital Libraries*, pp. 210–219, Chapel Hill, NC: ACM Press, 2006.
- [106] B. Kules and B. Shneiderman, "Users can change their web search tactics: Design guidelines for categorized overviews," *Information Processing & Management*, vol. 44, no. 2, pp. 463–484, 2008.
- [107] C. Kunz, "Sergio — an interface for context driven knowledge retrieval," in *Proceedings of eChallenges*, Bologna, Italy, 2003.
- [108] K. Lagus, S. Kaski, and T. Kohonen, "Mining massive document collections by the WEBSOM method," *Information Sciences: An International Journal*, vol. 163, no. 1–3, pp. 135–156, 2004.
- [109] J. Lamping and R. Rao, "The hyperbolic browser: A focus + context technique for visualizing large hierarchies," *Journal of Visual Languages and Computing*, vol. 7, no. 1, pp. 33–55, 1996.
- [110] X. Lin, D. Soergel, and G. Marchionini, "A self-organizing semantic map for information retrieval," in *Proceedings of the 14th annual international ACM SIGIR conference on Research and development in information retrieval, Chicago, Illinois, United States*, New York, NY, USA: ACM Press, 1991.
- [111] L. Lorigo, H. H. B. Pan, T. Joachims, L. Granka, and G. Gay, "The influence of task and gender on search and evaluation behavior using google," *Information Processing and Management*, vol. 42, no. 4, pp. 1123–1131, 2006.
- [112] R. Losee, *Text Retrieval and Filtering: Analytic Models of Performance*. Kluwer Academic Publishers, 1998.
- [113] A. Lunzer and K. Hornbæk, "Side-by-side display and control of multiple scenarios: Subjunctive interfaces for exploring multi-attribute data," in *Proceedings of Australian Special Interest Group Conference on Computer-Human Interaction*, Brisbane, Australia, 2003.
- [114] B. Mackay and C. Watters, "Exploring multi-session web tasks," in *CHI '08: Proceeding of the twenty-sixth annual SIGCHI conference on Human factors in computing systems*, Florence, Italy: ACM Press, 2008.
- [115] G. Marchionini, *Information Seeking in Electronic Environments*. Cambridge University Press, 1995.
- [116] G. Marchionini, "Exploratory search: From finding to understanding," *Communications of the ACM*, vol. 49, no. 4, pp. 41–46, 2006.
- [117] G. Marchionini and B. Shneiderman, "Finding facts vs. browsing knowledge in hypertext systems," *Computer*, vol. 21, no. 1, pp. 70–80, 1988.
- [118] B. Marshall, D. McDonald, H. Chen, and W. Chung, "Ebizport: Collecting and analyzing business intelligence information," *Journal of the American Society for Information Science and Technology*, no. 10, pp. 873–891, 2004.
- [119] K. Mcpherson, "Opinion-related information seeking: Personal and situational variables," *Personality and Social Psychology Bulletin*, vol. 9, no. 1, p. 116, 1983.
- [120] Y. Medynskiy, M. Dontcheva, and S. M. Drucker, "Exploring websites through contextual facets," in *Proceedings of the 27th International Conference on Human Factors in Computing Systems*, pp. 2013–2022, Boston, MA, USA: ACM Press, 2009.

- [121] D. R. Millen, J. Feinberg, and B. Kerr, "Dogear: Social bookmarking in the enterprise," in *Proceedings of the SIGCHI conference on Human Factors in computing systems*, pp. 111–120, Montréal, Québec Canada: ACM Press, 2006.
- [122] G. A. Miller, "Wordnet: A lexical database for english," *Communications of the ACM*, vol. 38, no. 11, p. 39, 1995.
- [123] D. K. Modjeska, *Hierarchical Data Visualization In Desktop Virtual Reality*. University of Toronto, 2000.
- [124] C. N. Mooers, "'mooers' law or why some retrieval systems are used and others are not," *American Documentation*, vol. 11, no. 3, 1960.
- [125] M. R. Morris, "A survey of collaborative web search practices," in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, Florence, Italy: ACM Press, 2008.
- [126] A. Motro, "Vague: A user interface to relational databases that permits vague queries," *ACM Transactions on Information Systems*, vol. 6, no. 3, pp. 187–214, 1988.
- [127] T. Munzner, "H3: Laying out large directed graphs in 3D hyperbolic space," in *Proceedings of IEEE Symposium on Information Visualization*, pp. 2–10, Phoenix, AZ, USA: IEEE Computer Society, 1997.
- [128] D. A. Nation, C. Plaisant, G. Marchionini, and A. Komlodi, "Visualizing web-sites using a hierarchical table of contents browser: WebTOC," in *Proceedings of the Third Conference on Human Factors and the Web*, Denver, CO, USA, 1997.
- [129] J. Nielsen and R. Molich, "Heuristic evaluation of user interfaces," in *CHI'90: Proceedings of the SIGCHI conference on Human factors in computing systems*, Seattle, WA, USA: ACM Press, 1990.
- [130] A. Oulasvirta, J. P. Hukkinen, and B. Schwartz, "When more is less: The paradox of choice in search engine use," in *Proceedings of the 32nd international ACM SIGIR conference on Research and development in information retrieval*, Boston, MA, USA: ACM Press, 2009.
- [131] T. Paek, S. Dumais, and R. Logan, "Wavelens: A new view onto internet search results," in *Proceedings of the 2004 conference on Human factors in computing systems*, pp. 727–734, Vienna, Austria: ACM Press, 2004.
- [132] L. D. Paulson, "Building rich web applications with AJAX," *Computer*, vol. 38, no. 10, pp. 14–17, 2005.
- [133] S. Perugini, K. Mcdevitt, R. Richardson, M. Perez-Quiones, R. Shen, N. Ramakrishnan, C. Williams, and E. A. Fox, "Enhancing usability in citidel: Multimodal, multilingual, and interactive visualization interfaces," in *Proceedings of the 4th ACM/IEEE-CS joint conference on Digital libraries*, pp. 315–324, Tuscon, AZ, USA: ACM Press, 2004.
- [134] P. Pirolli and S. Card, "Information foraging in information access environments," in *Proceedings of the SIGCHI conference on Human factors in computing systems*, pp. 51–58, Denver, CO, USA: ACM Press, 1995.
- [135] 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*, vol. 17, no. 3, pp. 320–341, 1999.

- [136] W. Pratt, "Dynamic organization of search results using the umls," *American Medical Informatics Association Fall Symposium*, vol. 480, no. 4, 1997.
- [137] Y. Qu and G. Furnas, "Model-driven formative evaluation of exploratory search: A study under a sensemaking framework," *Information Processing and Management*, vol. 44, no. 2, pp. 534–555, 2008.
- [138] K. Risdien, M. P. Czerwinski, T. Munzner, and D. B. Cook, "Initial examination of ease of use for 2D and 3D information visualizations of web content," *International Journal of Human-Computers Studies*, vol. 53, no. 5, pp. 695–714, 2000.
- [139] W. Rivadeneira and B. B. Bederson, *A Study of Search Result Clustering Interfaces: Comparing Textual and Zoomable User Interfaces*. University of Maryland: HCIL, 2003.
- [140] G. Robertson, M. Czerwinski, K. Larson, D. C. Robbins, D. Thiel, and M. V. Dantzich, "Data mountain: Using spatial memory for document management," in *Proceedings of the 11th annual ACM symposium on User interface software and technology*, pp. 53–162, San Francisco, CA, USA: ACM Press, 1998.
- [141] G. Robertson, M. C. K. Cameron, and D. Robbins, "Polyarchy visualization: Visualizing multiple intersecting hierarchies," in *Proceedings of the SIGCHI conference on Human factors in computing systems*, pp. 423–430, Minneapolis, Minnesota, USA: ACM Press, 2002.
- [142] G. G. Robertson, J. D. Mackinlay, and S. K. Card, "Cone trees: Animated 3D visualizations of hierarchical information," in *Proceedings of the SIGCHI conference on Human factors in computing systems: Reaching through technology*, pp. 189–194, New Orleans, Louisiana, USA: ACM Press, 1991.
- [143] S. E. Robertson and M. M. Hancock-Beaulieu, "On the evaluation of IR systems," *Information Processing and Management*, vol. 28, no. 4, pp. 457–466, 1992.
- [144] S. Santini and R. Jain, "Beyond query by example," in *Proceedings of the sixth ACM international conference on Multimedia*, pp. 345–350, Bristol, UK, 1998.
- [145] m. c. schraefel, "Building knowledge: What's beyond keyword search?," *Computer*, vol. 42, no. 3, pp. 52–59, 2009.
- [146] m. c. schraefel, D. A. Smith, A. Owens, A. Russell, C. Harris, and M. Wilson, "The evolving mSpace platform: Leveraging the semantic web on the trail of the memex," in *Proceedings of the Sixteenth ACM Conference on Hypertext and Hypermedia*, pp. 174–183, Salzburg, Austria: ACM Press, 2005.
- [147] m. c. schraefel, M. L. Wilson, and M. Karam, *Preview Cues: Enhancing Access to Multimedia Content*. School of Electronics and Computer Science, University of Southampton, 2004.
- [148] m. c. schraefel, M. L. Wilson, A. Russell, and D. A. Smith, "mSpace: Improving information access to multimedia domains with multimodal exploratory search," *Communications of the ACM*, vol. 49, no. 4, pp. 47–49, 2006.
- [149] B. Schwartz, *The Paradox of Choice: Why More Is Less*. Harper Perennial, 2005.
- [150] M. M. Sebrechts, J. V. Cugini, S. J. Laskowski, J. Vasilakis, and M. S. Miller, "Visualization of search results: A comparative evaluation of text, 2D, and

- 3D interfaces,” in *Proceedings of the 22nd annual international ACM SIGIR conference on Research and development in information retrieval*, pp. 3–10, Philadelphia, PA, USA: ACM Press, 1999.
- [151] B. Shneiderman, “Designing information-abundant web sites: Issues and recommendations,” *International Journal of Human-Computer Studies*, vol. 47, no. 1, pp. 5–29, 1997.
- [152] B. Shneiderman, “Why not make interfaces better than 3D reality?,” *Computer Graphics and Applications, IEEE*, vol. 23, no. 6, pp. 12–15, 2003.
- [153] B. Shneiderman and A. Aris, “Network visualization by semantic substrates,” *IEEE Transactions on Visualization and Computer Graphics*, vol. 12, no. 5, pp. 733–740, 2006.
- [154] B. Shneiderman, D. Feldman, A. Rose, and X. F. Grau, “Visualizing digital library search results with categorical and hierarchical axes,” in *Proceedings of the Fifth ACM International Conference on Digital Libraries*, pp. 57–66, San Antonio, TX: ACM Press, 2000.
- [155] B. Shneiderman and C. Plaisant, *Designing the User Interface: Strategies for Effective Human-Computer Interaction*. Boston: Pearson/Addison-Wesley, 2004.
- [156] C. Silverstein, H. Marais, M. Henzinger, and M. Moricz, “Analysis of a very large web search engine query log,” *ACM SIGIR Forum*, vol. 33, no. 1, pp. 6–12, 1999.
- [157] J. Sinclair and M. Cardew-Hall, “The folksonomy tag cloud: When is it useful?,” *Journal of Information Science*, vol. 34, no. 1, pp. 15–29, 2008.
- [158] A. F. Smeaton, P. Over, and W. Kraaij, “Evaluation campaigns and TRECVID,” in *Proceedings of the 8th ACM international workshop on Multimedia information retrieval*, pp. 321–330, Santa Barbara, California, USA: ACM Press, 2006.
- [159] A. Spink, “Study of interactive feedback during mediated information retrieval,” *Journal of the American Society for Information Science*, vol. 48, pp. 382–394, 1997.
- [160] A. Spink, “A user-centered approach to evaluating human interaction with web search engines: An exploratory study,” *Information Processing & Management*, vol. 38, no. 3, pp. 401–426, 2002.
- [161] A. Spink and T. D. Wilson, “Toward a theoretical framework for information retrieval (IR) evaluation in an information seeking context,” in *Proceedings of Multimedia Information Retrieval Applications*, p. paper 9, Glasgow, UK: BCS, 1999.
- [162] E. Stoica and M. A. Hearst, “Nearly-automated metadata hierarchy creation,” in *Proceedings of the Human Language Technology Conference of the North American Chapter of the Association for Computational Linguistics*, Boston, MA USA, 2004.
- [163] A. G. Sutcliffe and U. Patel, “3D or not 3D: Is it nobler in the mind?,” in *Proceedings of the BCS Human Computer Interaction Conference on People and Computers XI*, pp. 79–94, London, UK: Springer-Verlag, 1996.
- [164] D. Swanson, N. Smalheiser, and A. Bookstein, “Information discovery from complementary literatures: Categorizing viruses as potential weapons,”

- Journal American Society Information Science and Technology*, vol. 52, no. 10, pp. 797–812, 2001.
- [165] E. Tanin, C. Plaisant, and B. Shneiderman, “Browsing large online data with query previews,” in *Proceedings of the Symposium on New Paradigms in Information Visualization and Manipulation*, Washington D.C., USA: ACM Press, 2000.
- [166] J. Teevan, E. Cutrell, D. Fisher, S. M. Drucker, P. A. G. Ramos, and C. Hu, “Visual snippets: Summarizing web pages for search and revisitation,” in *Proceedings of the 27th international conference on Human factors in computing systems*, pp. 2023–2032, Boston, MA, USA: ACM Press, 2009.
- [167] E. R. Tufte, *Envisioning Information*. Graphics Press Cheshire, 1990.
- [168] D. Tunkelang, *Faceted Search*. Morgan and Claypool, 2009.
- [169] O. Turetken and R. Sharda, “Clustering-based visual interfaces for presentation of web search results: An empirical investigation,” *Information Systems Frontiers*, vol. 7, no. 3, pp. 273–297, 2005.
- [170] P. Vakkari, “Cognition and changes of search terms and tactics during task performance: A longitudinal case study,” in *Proceedings of the International Conference on Computer Assisted Information Retrieval (RAIO)*, Paris, France, 2000.
- [171] P. Wang, M. Berry, and Y. Yang, “Mining longitudinal web queries: Trends and patterns,” *Journal of the American Society for Information Science and Technology*, vol. 54, no. 8, pp. 742–758, 2003.
- [172] R. W. White, S. M. D. B. Kules, and m. c. schraefel, “Introduction,” *Communications of the ACM*, vol. 49, no. 8, pp. 36–39, 2006.
- [173] R. W. White and S. M. Drucker, “Investigating behavioral variability in web search,” in *Proceedings of the 16th International Conference on World Wide Web*, Banff, Canada: ACM Press, 2007.
- [174] R. W. White and R. Roth, *Exploratory Search: Beyond the Query-Response Paradigm*. Morgan & Claypool, 2009.
- [175] R. W. White, I. Ruthven, and J. M. Jose, “Finding relevant documents using top ranking sentences: An evaluation of two alternative schemes,” in *Proceedings of the 25th annual international ACM SIGIR conference on Research and development in information retrieval*, pp. 57–64, Tampere, Finland: ACM Press, 2002.
- [176] M. Wilson, “mspace: What do numbers and totals mean in a flexible semantic browser,” *SWUI’06*, 2006.
- [177] M. Wilson and m. c. schraefel, “Bridging the gap: Using ir models for evaluating exploratory search interfaces,” in *SIGCHI 2007 Workshop on Exploratory Search and HCI*, San Jose, CA, USA, 2007.
- [178] M. L. Wilson, “An analytical inspection framework for evaluating the search tactics and user profiles supported by information seeking interfaces,” PhD Thesis, University of Southampton, 2009.
- [179] M. L. Wilson, P. André, and m. c. schraefel, “Backward highlighting: Enhancing faceted search,” in *Proceedings of the 21st Annual ACM Symposium on User Interface Software and Technology*, Monterey, CA, USA: ACM Press, 2008.

- [180] M. L. Wilson, P. André, D. A. Smith, and m. c. schraefel, *Spatial Consistency and Contextual Cues for Incidental Learning in Browser Design*. School of Electronics and Computer Science, University of Southampton, 2007.
- [181] M. L. Wilson and m. c. schraefel, “mSpace: What do numbers and totals mean in a flexible semantic browser,” in *Proceedings of the 3rd International Semantic Web User Interaction Workshop*, Athens, GA, USA, 2006.
- [182] M. L. Wilson and m. c. schraefel, “A longitudinal study of both faceted and keyword search,” in *Proceedings of the ACM/IEEE-CS Joint Conference on Digital Libraries*, Pittsburgh, PA, USA: IEEE Computer Society, 2008.
- [183] M. L. Wilson and m. c. schraefel, “The importance of conveying inter-facet relationships for making sense of unfamiliar domains,” in *CHI’09 Workshop on Sensemaking*, Boston, MA, USA, 2009.
- [184] M. L. Wilson, m. c. schraefel, and R. W. White, “Evaluating advanced search interfaces using established information-seeking models,” *Journal of the American Society for Information Science and Technology*, vol. 60, no. 7, pp. 1407–1422, 2009.
- [185] A. Woodruff, R. R. A. Faulring, J. Morrision, and P. Pirolli, “Using thumbnails to search the web,” in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pp. 198–205, Seattle, Washington, USA: ACM Press, 2001.
- [186] H. Wu, M. Zubair, and K. Maly, “Harvesting social knowledge from folksonomies,” in *Proceedings of the seventeenth conference on Hypertext and hypermedia*, pp. 111–114, Odense, Denmark: ACM Press, 2006.
- [187] K. Yang, “Information retrieval on the web,” *Annual Review of Information Science and Technology*, vol. 39, pp. 33–80, 2005.
- [188] K.-P. Yee, K. Swearingen, K. Li, and M. Hearst, “Faceted metadata for image search and browsing,” in *Proceedings of the ACM Conference on Human factors in Computing Systems*, pp. 401–408, Fort Lauderdale, FL, USA: ACM Press, 2003.
- [189] O. Zamir and O. Etzioni, “Grouper: A dynamic clustering interface to web search results,” *Computer Networks*, vol. 31, pp. 1361–1374, 1999.
- [190] O. Zamir, O. Etzioni, O. Madani, and R. M. Karp, “Fast and intuitive clustering of web documents,” in *Proceedings of the 3rd International Conference on Knowledge Discovery and Data Mining (KDD’97)*, pp. 287–290, Newport, CA, USA, 1997.
- [191] H.-J. Zeng, Q.-C. HE, Z. Chen, W.-Y. Ma, and J. Ma, “Learning to cluster web search results,” in *Proceedings of the 27th Annual International Conference on Research and Dvelopment in Information Retrieval, Sheffield, United Kingdom*, New York: ACM Press, 2004.
- [192] J. Zhang and G. Marchionini, “Evaluation and evolution of a browse and search interface: Relation browser,” in *Proceedings of the National Conference on Digital Government Research*, pp. 179–188, Atlanta, GA, USA: Digital Government Research Center, 2005.