Agarwal, N.K. & Poo, D.C.C. (2007). Collaborating to search effectively in different searcher modes through cues & specialty search. In D.H. Goh & S. Foo (Eds.) *Social Information Retrieval Systems: Emerging Technologies & Applications for Searching the Web Effectively* (pp. 1-30), Idea Group Inc. https://doi.org/10.4018/978-1-59904-543-6.ch001

Chapter I
Collaborating to Search Effectively in Different Searcher Modes Through Cues and Specialty Search

Naresh Kumar Agarwal
National University of Singapore, Singapore

Danny C.C. Poo
National University of Singapore, Singapore

ABSTRACT

Searchers generally have difficulty searching into knowledge repositories because of the quantity of data involved and because search mechanisms are not tailored to the differing needs of the searcher at different points in time. Also, every searcher generally searches alone without taking into account other users with similar search needs or experience. While the Internet may have contributed to information overload, the connectivity it has provides the potential to different searchers to collaborate when looking for information. In this chapter, we: (1) review concepts related to social information retrieval and existing collaborative mechanisms, (2) discuss two collaborative mechanisms—cues and specialty search, and (3) see cues and specialty search in the context of the changing needs of a searcher in one of four modes. A case study of an online portal for the Singapore education community is used to show how collaboration could enhance learning and search efficacy.

INTRODUCTION

Knowledge repositories are increasingly a part of any enterprise. Masses of documents, e-mails, databases, images, and audio/video recordings form vast repositories of information assets to be tapped by employees, partners, customers, and other stakeholders (Papadopoullos, 2004). The content provided in such repositories is large, diverse, and huge in quantity. Searchers generally have difficulty searching into such kinds of repositories because of the quantity of data involved and because searcher mechanisms are not tailored to their differing needs at different points in time. Also, every searcher generally searches alone, without taking into account other users who would have conducted similar searches or have a similar work role as the searcher.

A searcher typically does not just access organization-level repositories, but has access to vast amounts of information from the Internet. The growth of the Internet has brought information access to individuals from all walks of life and has connected the world like never before. According to the Berkeley study "How Much Information" (Swearingen et al., 2003), print, film, magnetic, and optical storage media produced about 5 exabytes of new information in 2002 (1 gigabyte = 10^9 bytes; 1 terabyte = 10^{12} bytes; 1 exabyte = 10^{18} bytes; 5 exabytes are equivalent to all words ever spoken by human beings). The study estimated that the amount of new information stored in these media had doubled between 1999 and 2002, and grew about 30% each year. While there is no dearth of information, there is a long and meandering path before this information translates to knowledge and understanding. Sieving the important from the unimportant, the relevant from the non-relevant, getting answers to the questions, and making sense of all the data available are some of the challenges faced by searchers of information. The World Wide Web, while providing increased connectivity and accessibility to information, has also increased the amount of information a person must read and digest each day—a problem commonly referred to as *information overload*. Compared to the growth of the World Wide Web, "development of the human brain has been tardy: it has grown *only linearly* from 400 to 1400 cubic centimeters in the last 3.5 million years" (Chakrabarti et al., 1999).

To help retrieve information from this huge maze, search engines come in handy and serve as catalogs of the Web. They index the Web pages by using computer programs called 'spiders' or 'robots', which crawl from site to site and create a database that stores indices of Web pages on the Web. Users can enter search terms to query against the index database. The search engine processes the query and returns a list of Web pages, along with short descriptions of each page (Fang, Chen, & Chen, 2005). The search engines' critical role in helping people find information online makes them the gatekeepers to online information (Morahan-Martin & Anderson, 2000).

However, "search engines do not index sites equally, may not index new pages for months, and no engine indexes more than about 16% of the Web" (Lawrence & Giles, 2000, p. 33). This was in 2000, and the coverage of search engines has increased since then (but the size of the Web has also increased, along with the non-indexable 'deep Web'). Problems due to synonymy and polysemy plague the current information searches (Deerwester, Dumais, Furnas, Landauer, & Harshman, 1990; Morahan-Martin & Anderson, 2000). Synonymy is the semantic relation that holds between two words that can, in a given context, express the same meaning. Polysemy is the ambiguity of an individual word or phrase that can be used, in different contexts, to express two or more different meanings (WordNet 2.0, 2003). For example, the keywords "female sibling" and "sister" might mean the same thing but give different results on searching (the problem of synonymy). On the other hand, searching for the keyword 'apple' may give you a page full of links to 'Apple Computers', while you might be searching for information related to the fruit. Similarly, searching for 'Java' may give you top links about the Java programming language, while you might be interested in coffee or the Indonesian island of Java. This is the problem of polysemy. Search engines suffer from another major drawback—they make an underlying presumption that the user can formulate on-point queries to effectively narrow down the volume of information available (Narayanan, Koppaka, Edala, Loritz, & Daley, 2004).

Effective query formulation is possible only when the users are already familiar with the topic of research and they indeed can see the subtle differences in vocabulary of the search topic (Belkin, 2000). Yet another problem in using the search engines of today is that the interests of the users vary with time and cannot be represented by a fixed set (Narayanan et al., 2004) or a 'one-size-fits-all' model widely prevalent in the search engines of today. Thus, there is a lack of fit between the information systems available for search and the task needs of different searchers or of the same searcher at different times. Also, "most Web search engines in use fail to take advantage of the intentions, interests and preferences of their users" (Pujol, Sanguesa, & Bermudez, 2003). Every searcher is also expected to *reinvent the wheel* each time he or she searches, while there might be other searchers with similar needs or those who are experts in the area of the searcher's needs, whose expertise is not tapped in a useful manner.

The emphasis should hence be on addressing questions posed by users, through facilitating information search and knowledge discovery (Marchionini, 1997; Fayyad, Piatetsky-Shapiro, Smyth, & Uthurusamy, 1996). To facilitate this process of knowledge discovery, information providers should attempt to understand the context surrounding each search task rather than simply presenting searchers with a series of links.

People are very impressed with Web searches today but it's really quite poor compared to what it should be…a bunch of links that sort of start a treasure hunt that on average takes about 11 minutes. Bill Gates, Microsoft Chairman (Live! Forum, Singapore, July 1, 2005)

Knowledge discovery can happen if information systems are designed to store the search patterns of users and facilitate a new searcher by comparing his search behavior with records of prior searches. Once a pattern of similarity is found, tools and information may be extended to the new searcher that had served the needs of an existing searcher. If earlier searchers with similar needs had been satisfied with the information, there is the likelihood that the information is useful to a new searcher too. While the Internet may have contributed to information overload, the connectivity it has brought provides the potential to different searchers of information to *collaborate* and work together when looking for information. Collaborative or *social* approaches to searching harness voluntary efforts of several people that relate to each other through networked information systems. Social information discovery and filtering systems rely on the existence of other people who locate and evaluate relevant sources and are willing to share the discovered information (Karamuftuoglu, 1998; Hill, Stead, Rosenstein, & Furnas, 1995; Starr, Ackerman, & Pazzani, 1996).

For collaboration to be successful, the *similarity of information needs* between that of the searcher with those of previous searchers must be effectively established. The collaborator or collaborative mechanisms must be able to help the searcher either through expertise or experience, or similarity of needs. While providing collaborative mechanisms, an information provider must also take into account the different modes a searcher is in at different points in time (based on the characteristics of the task at the hand, or the qualities/expertise of the searcher), and provide technology features that match the task and searcher characteristics reflected from the searcher mode.

The objectives of this chapter are three-fold:

- To review concepts related to social information retrieval and some existing collaborative mechanisms.
- To discuss two collaborative mechanisms: *cues* and *specialty search.* These could be built in a system and will allow a searcher to retrieve information collaboratively with other like-minded searchers.

- To see cues and specialty search in the context of the changing needs of the information searcher at different points in time who could be in one of four modes—*novice, data gatherer, known-item searcher,* or *focused searcher.*

A case study of an online portal for the Singapore Education Community will be used for illustration. An outcome of the project is to demonstrate how students, teachers, and other users could collaborate among themselves to enhance learning and the efficacy of search.

In the next section, we seek to provide a basic background of concepts related to social information retrieval, and discuss collaboration and existing collaborative mechanisms in greater detail. We then look at two collaborative mechanisms, contextual cues and specialty search, followed by a discussion of the concept of searcher modes—the differing needs of a searcher at different points in time. We examine cues and specialty search in the context of these modes. Next, we provide an illustration using a case study on collaborating for education-related search. This is followed by a section on future trends. Finally, the conclusion highlights some key points and concludes the chapter.

BACKGROUND: COLLABORATION AND COLLABORATIVE MECHANISMS

Concepts Related to Social Information Retrieval/ Collaborative Search

Before delving deeper into collaboration and existing collaborative mechanisms, let us examine a few related concepts.

A commonly held view with sundry minor variants is that *data* is raw numbers and facts; *information* is processed data or "a construct on a continuum somewhere between data and knowledge" (North, North, & Benade, 2004; see Figure 1), while *knowledge* is authenticated information (Machlup, 1980; Dretske, 1981; Vance, 1997).

Yet the presumption of hierarchy from data to information to knowledge with each varying along some dimension, such as context, usefulness, or interpretability, rarely survives scrupulous evaluation (Alavi & Leidner, 2001). According to North et al. (2004), "information is determined or defined by its use…information has value when it is relevant to the task at hand, it is available in the right format at the right place, and is considered fairly accurate and recent." The goal of collaborative mechanisms espoused

Figure 1. Continuum of data, information, and knowledge

in this chapter is to ensure that the searcher gets access to the right information at the right time, using the help of other like-minded searchers or collaborative mechanisms.

Information need is the recognition that our knowledge is inadequate to satisfy a goal that we have (Case, 2002). "Need for information consists of the process of perceiving a difference between an ideal state of knowledge and the actual state of knowledge" (van de Wijngaert, 1999, p. 463). Search for information is based on some need, task, or problem at hand. Our propositions in this chapter are based on the premise that there are other people out there with needs similar to ours or those who have had similar needs in the past. We investigate ways and means to best collaborate with those with similar needs.

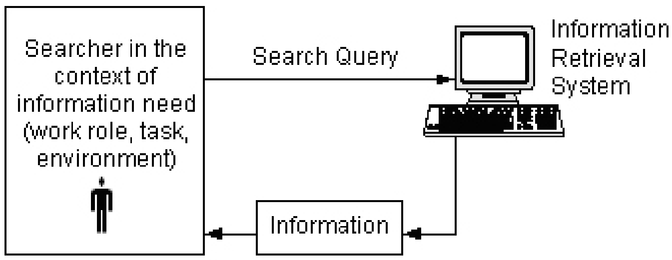
An *information retrieval* (IR) system has the goal of "leading the user to those documents that will best enable him/her to satisfy his/her need for information" (Robertson, 1981) or "for the user to obtain information from the knowledge resource which helps her/him in problem management" (Belkin, 1984). Information retrieval implies searching for information using a computer or information system.

Figure 2 shows the components of a basic information retrieval system. Based on the task at hand and the search context, the user tries to express his/her need in a few *keywords* and enters it into the information retrieval system. Depending on the information retrieval algorithm implemented, the system returns the information (typically a set of links pointing to the information) that has words matching with the search keywords. In the classical IR sense, "an information retrieval system does not inform (i.e., change the knowledge of) the user on the subject of his inquiry. It merely informs on the existence (or non-existence) and whereabouts of documents relating to his request" (Lancaster, 1968).

The classic information retrieval research tradition commenced with the Cranfield tests in the 1950s and 1960s, and continued with the MEDLARS evaluation, research on relevance judgment (1970s), automated systems, theoretical work by Van Rijsbergen and Robertson (late 1970s), empirical work on relevance feedback, and comparisons of Boolean and best match searching (1980s). Statistical as well as cognitive approaches have been researched over the years (Ellis, Allen, & Wilson, 1999). Apart from the classical 'system-oriented' approach (where an IR system is an integral part), studies have also been done from the perspective of the user and his needs (person-oriented studies), under the umbrella of 'information seeking'. Here, the process of searching may not necessarily involve searching from an information retrieval system. See Case (2002) for a complete review.

Moving from classical information retrieval to social or collaborative information retrieval is the onus of the builders of information retrieval systems, that is, the information providers, who can provide mechanisms to help searchers collaborate amongst each other. We posit that the

Figure 2. Information retrieval system



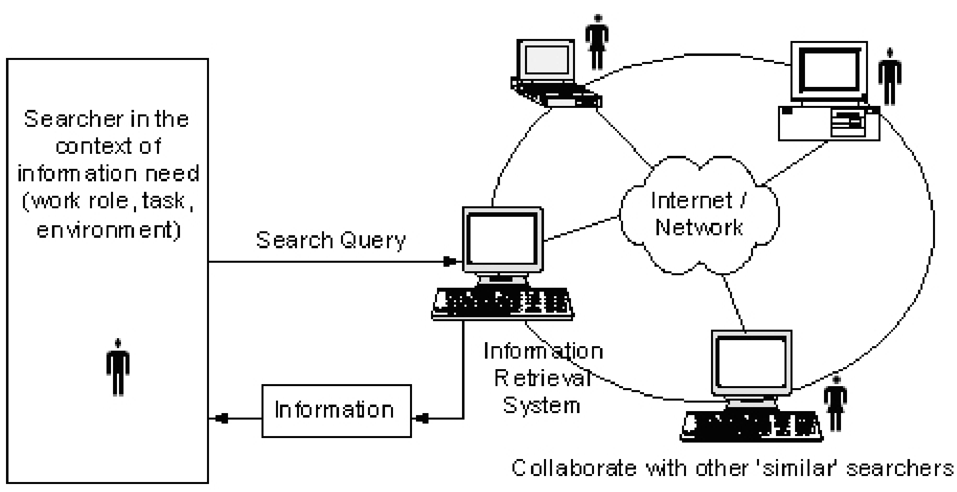
mechanisms provided must also take into account the differing modes of the searcher at different pints in time (see the subsection 'Changing Needs of the Searcher: Four Searcher Modes' later in the chapter).

Social information retrieval refers to a family of information retrieval techniques that assist users in obtaining information to meet their information needs by harnessing other users' expert knowledge or search experience. Users are linked through networked information systems such as the Internet. Such systems rely on other people who have found relevant information and are willing to share it (Karamuftuoglu, 1998; Hill et al., 1995; Starr et al., 1996). While classical IR deals with the interaction of an individual with an information system when looking for information (other users do not come into the picture), social or collaborative approaches to information retrieval makes use of the expertise of *other* users when searching for information. While the former can be likened to an *individual effort*, the latter is more of a *team effort* to search.

Collaborative search "exploits repetition and regularity within the query-space of a community of like-minded individuals in order to improve the quality of search results. In short, search results that have been judged to be relevant for past queries are promoted in response to similar queries that occur in the future" (Smyth et al., 2005, p. 1419). It relies on searchers willing to collaborate over network systems such as the Internet to contribute information to be used by other needy searchers with similar needs. Collaborative Web search combines "techniques for exploiting knowledge of the query-space with ideas from social networking to develop a Web search platform capable of adapting to the needs of (ad-hoc) communities of users. In brief, the queries submitted and the results selected by a community of users are recorded and reused in order to influence the results of future searches for similar queries. Results that have been reliably selected for similar queries in the past are promoted" (Smyth et al., 2005, p. 1419; Freyne, Smyth, Coyle, Balfe, & Briggs, 2004; Smyth, Balfe, Briggs, Coyle, & Freyne, 2003).

Figure 3 shows an IR system that facilitates collaboration. As shown in the figure, a searcher situated in a unique search context (a particular task at hand, work role, or surrounding environment) has a particular need for information. Based

Figure 3. Social information retrieval/collaborative search



on his understanding of the topic or domain under search, he keys in a search query in an information retrieval system, hoping for results. Instead of simply doing a straightforward keyword search and spitting out results (classic information retrieval), the system enables the searcher to collaborate with other searchers who have conducted similar searches before or are experts in the domain of search. This would require matching the searcher with other similar searchers who could be of help to him. Specific collaborative mechanisms may be provided by the system. The information provided may be more useful to the searcher than without collaboration.

Apart from simple collaboration, we also posit that the information retrieval system understands the unique mode the searcher is in (based on his search context) and provides mechanisms that take the searcher mode into account. Let us delve deeper into the idea behind collaborative search.

Why Collaborative Search?

The idea is simple. Behind every search for information, there is an information need. This need is always *instrumental* (Green, 1990) in that it involves reaching a desired goal. The searcher wants to accomplish a certain goal with the retrieved information. This goal might be solving a problem, answering a question, accomplishing a work task, satisfying one's curiosity, or even entertainment. Knowing the information will put the searcher at or closer to an end state he/she wants to achieve (Case, 2002). In our world of 6 billion people, it is very likely that someone, somewhere would have encountered similar situations or contexts of information need as we have—that is, this person could have solved or is solving a similar problem, have looked for answers to a similar question, is in a similar work role, or is as curious as we are in looking for celebrity gossip. In an earlier era, locating such people with similar needs or who have had similar needs in the past could have been limited by physical or geographical boundaries. The exercise would have been expensive and infeasible. The expansion of the World Wide Web, which continues to grow continuously and exponentially, has opened up opportunities like never before. Locating people with similar interests, experiences, work roles, and more importantly, with similar needs for information has been made as simple as a click of a mouse, making collaboration easy and natural.

For instance, vocabulary mismatch is a deep-rooted problem in information retrieval as users often use different or too few words to describe concepts in their queries as compared to the words used by authors to describe similar concepts. This leads to inadequate search results (Komarjaya, Poo, & Kan, 2004). Query expansion (or query reformulation), the process of expanding/disambiguating a user's query with additional related words and phrases, has been suggested to address the problem (Komarjaya et al., 2004). However, finding and using appropriate related words remains an open problem. Collaborative querying is an approach whereby related queries (the query clusters) may be calculated based on the similarities of the queries with past search experiences (as documented in the query logs) and either recommended to users or used as expansion terms (Fu, Goh, Foo, & Supangat, 2004). However, as pointed out by Fu et al. (2004), calculating the similarity between different queries and clustering them automatically (query clustering) are crucial steps here.

Not just query reformulation, search engines, in many ways and in their very implementation, make use of the similar experiences of past searchers to arrive at results. For instance, Google's PageRank "relies on the uniquely democratic nature of the Web by using its vast link structure as an indicator of an individual page's value. In essence, Google interprets a link from page A to page B as a vote, by page A, for page B." Google also analyzes the page that casts the vote. If it considers the page that casts the vote more important, the votes cast by it will weigh more heavily and help to make other pages important (Google Technology, 2004). Thus, if more and more pages link to a particular page, it becomes more and more important. This emphasis on *social role* in Web page ranking, compared to basic keyword/frequency matching, has found phenomenal success, with Google emerging as one of the most favorite search engines (Sherman, 2006).

There is huge potential for collaboration to extend beyond page linking/ranking to the very manner of looking for information in the World Wide Web. The motivation is compelling. The search engines of today, though good and ever improving, are not perfect. Users are still swamped with huge amounts of non-relevant data. Information gathering could easily extend from minutes to hours to days. In addition, the information needs of a searcher vary with time, the task at hand, and the ever-changing context or environment in which the information need arises.

In such a scenario, "…support for communication and collaboration is as important as support for information-seeking activities, and…indeed, support for the former is needed to support the latter" (Levy & Marshall, 1994). Virtual communities such as Weblogs (or blogs), online groups, and discussion forums are all aimed at supporting information sharing, and their success implies their effectiveness (Chi & Pirolli, 2006). "Cooperation may yield more benefits than simply making search more parallel and making it less prone to failure. Membership in a group provides actual or potential resources that can be utilized or mobilized to achieve individual goals. This is known as *social capital*" (Chi & Pirolli, 2006). Many informal group memberships are visible in the form of *communities of practice* or CoPs. While the term CoP is widely used (see Cox, 2005, for a review of different definitions), it comes from theories based on the idea of learning as social participation (Wenger, 1998). Wenger, McDermott, and Snyder (2002) define CoPs as "groups of people who share a concern, a set of problems, or a passion about a topic, and who deepen their knowledge and expertise in this area by interacting on an ongoing basis" (p. 7). Cox (2005, p. 531) lists Wenger's (1998, pp. 125-126) indicators of communities of practice:

1) sustained mutual relationships—harmonious or conflictual; 2) shared ways of engaging in doing things together; 3) the rapid flow of information and propagation of innovation; 4) absence of introductory preambles, as if conversations and interactions were merely the continuation of an ongoing process; 5) very quick setup of a problem to be discussed; 6) substantial overlap in participants' descriptions of who belongs; 7) knowing what others know, what they can do, and how they can contribute to an enterprise; 8) mutually defining identities; 9) the ability to assess the appropriateness of actions and products; 10) specific tools, representations, and other artifacts; 11) local lore, shared stories, inside jokes, knowing laughter; 12) jargon and shortcuts to communication as well as the ease of producing new ones; 13) certain styles recognized as displaying membership; 14) a shared discourse reflecting a certain perspective on the world.

Such indicators of CoPs and the way they work are important in understanding the efficacy of many existing social and collaborative approaches to searching such as social bookmarking, social networking, folksonomies, and so forth.

Existing Collaborative Mechanisms

By tapping into the resources and expertise of those more knowledgeable or experienced, there is huge potential to improve the efficacy of information search. Searchers could collaborate in a number of ways in order to retrieve information effectively. Techniques could include sharing of search queries, social bookmarking and tagging, folksonomies, social network analysis, subjective relevance judgments, and collaborative filtering.

Social Bookmarking and Tagging

Collaborative tagging "describes the process by which many users add metadata in the form of keywords to shared content" (Golder & Huberman, 2005). Users can collaboratively tag various content such as bookmarks, documents, photographs, blog entries, and so forth. Bookmarking is the activity when a Web user makes note of a favorite site or hyperlink on his browser. A user can manage, tag, comment upon, and publish his bookmarks on the Web, which represent a user's personal library being placed on the Web. When aggregated with other personal libraries, it allows for rich, social networking opportunities (Hammond, Hanny, Lund, & Scott, 2005). This is primarily the idea of social bookmarking and tagging. Hammond and his colleagues review various available tools to help achieve the same. Social bookmark services like del.icio.us (*http://del.icio.us/*) allow users to freely choose category names and tags without any a priori dictionary, taxonomy, or ontology to conform to. Such services may be seen as "social annotations" of the Web (Wu, Zhang, & Yu, 2006). However, "without a shared taxonomy or ontology, social annotations suffer the usual problem of ambiguity of semantics. The same annotation may mean different things for different people and two seemingly different annotations may bear the same meaning" (Wu et al., 2006, p. 418). Wu et al. suggest a method to group synonymous tags together and to identify and separate highly ambiguous tags. Social bookmarking is only as reliable as the people doing the tagging and provides a subjective, rather than an objective, opinion of the people tagging. You trust a stranger's recommendations for a topic. The positive part is that you may find better resources through somebody else's time and effort spent on research. Another downside is that there is no common language, so somebody else's bookmarked sites may not be related to what you are looking for. There is also the risk of spam being tagged, which could result in undesirable clutter (Asmus, Bonner, Esterhay, Lechner, & Rentfrow, 2005).

Folksonomies

A taxonomy is a structured way to categorize information and provides a subject-based classification that arranges the terms in a controlled vocabulary into a hierarchy. By relating word relationships (synonyms, broader terms, and narrower terms) and gathering the results in a common bucket, taxonomies can be used to bring common or similar material together. Humans can rapidly navigate taxonomies to find high concentrations of topic-specific, related information (Lederman, 2005; Papadopoullos, 2004). When such a taxonomy is generated by Internet users (instead of by professionals or content creators/authors) for their own individual use that is also shared throughout a community, using an open-ended labeling system to categorize various types of content, we get a novel combination of *folk* (not formal or professional) and *taxonomy*, that is, *folksonomy*. However, unlike a taxonomy, a folksonomy comprises terms in a flat namespace, where there is no hierarchy between terms. It is simply the set of terms that a group of users tagged content with, and not a predetermined set of classification terms or labels. Flickr (*www.flickr.com*) provides a collaborative way of tagging and categorizing photographs; del.icio.us (*http://del.icio.us/*) is a collection of bookmarks of various users; You Tube (*www.youtube.com*) allows tagging, sharing, and hosting of short video clips; CiteULike (*www.citeulike.org*) tags scientific publications; while 43Things (*www.43things.com*) allows users to annotate their goals and plans with keywords, and connects users with similar pursuits (Mathes, 2004; Golder & Huberman, 2005; Hammond et al., 2005; Mika, 2005). On the downside, there is absence of polysemy and synonymy management in folksonomies. For example, a goal to stop procrastinating has been tagged variously in 43Things as "stop procrastinating," "stop procrastination," "procrastinate less," "stop procrastinate," "stop procrastinating and do things asap," "do less of procrastination," and so forth. Thus, synonyms, ambiguity, and improper use of case sensitivity and punctuation marks is commonplace. However, the imperfections of tagging are nonetheless acceptable so far, and users can instantly link to other relevant, timely, socially ranked objects (Mika, 2005).

Social Network Analysis

"One of the most consistent findings in the social science literature is that who you know often has a great deal to do with what you come to know. Yet both practical experience and scholarly research indicate significant difficulty in getting people with different expertise, backgrounds and problem solving styles to effectively integrate their unique perspectives" (Cross, Borgatti, & Parker, 2002). From the view of social network analysis, the social environment can be expressed as patterns or regularities in relationships (referred to as 'structure') among interacting units, where structure is measured using quantities called structural variables (Wasserman & Faust, 1994, p. 3). According to Wasserman and Faust, a social network consists of a finite set or sets of actors and the relation or relations defined on them. Actors are discrete individual, corporate, or collective social units (Wasserman & Faust, 1994, p. 17). The presence of relational information is a critical and defining feature of a social network. According to Scott (2000), relational data are the contacts, ties and connections, the group attachments and meetings, which relate one agent [actor] to another and so cannot be reduced to the properties of the individual agents themselves. Relations are not the properties of agents, but of systems of agents; these relations connect pairs of agents into larger relational systems. The methods appropriate to relational data are those of network analysis, whereby the relations are treated as expressing the linkages that run between agents (p. 3).

In addition to the use of relational concepts, the central principles of the network perspective are (Wasserman & Faust, 1994, p. 4):

- Actors and their actions are viewed as interdependent rather than independent, autonomous units.
- Relational ties (linkages) between actors are channels for transfer or "flow" of resources (either material or nonmaterial).
- Network models focusing on individuals view the network structural environment as providing opportunities for or constraints on individual action.
- Network models conceptualize structure (social, economic, political, and so forth) as lasting patterns of relations among actors.

Wasserman and Faust also present Freeman's mathematical definition for a social network $Y = <S, G_d, X>$, where the triple consisting of the algebraic structure S, the directed graph or sociogram G_d, and the adjacency matrix or sociomatrix X is viewed as a social network. These three notations S, G_d, and X are usually viewed together as providing the three essential components of the simplest form of a social network (Wasserman & Faust, 1994, p. 40).

- The algebraic structure S is a set of nodes and a set of arcs (from graph theoretic notations).
- A sociogram (notation G_d above) is a graph produced from the sets of nodes and arcs. 'Invented' by Jacob L. Moreno in 1933, a sociogram is a picture in which people (or more generally, any social units) are represented as points in two-dimensional space, and relationships among pairs of people are represented by lines linking the corresponding pairs (Wasserman & Faust, 1994, pp. 11-12; Scott, 2000, pp. 9-10). For Moreno, social configurations had definite and discernible structures, and the mapping of these structures into a sociogram allowed a researcher

to visualize the channels through which, for example, information could flow from one person to another, and through which one individual could influence another (Wasserman and Faust, 1994, p. 10).

- A sociomatrix (notation X above) is a two-way matrix used to present the data for each relation, where the rows and columns refer to the actors making up the pairs.

Social network analysis is an effective tool for collaborative search and social information retrieval. This is highlighted by Morville (2002), who points to the reciprocal relationship between people and content (we use *people* to find content ↔ we use *content* to find people). Using people to find content requires knowing what/who other people know. Using content to find people demands good search, navigation, and content management systems. Morville (2002) points out that with the way document surrogates such as *abstracts* are often used in information retrieval to represent the knowledge contained within that content, documents themselves may be considered as "human surrogates" representing the knowledge and interests of authors, while humans also serve as surrogates for one another. This suggests a need for metadata schema, tools, people directories, and incentives to enable and encourage explicit connections between documents and authors (Morville, 2002). There are a number of Internet social networks such as Orkut (*www.orkut.com*), Hi5 (*www.hi5.com*), Yahoo 360° (*http://360.yahoo.com*), Classmates (*www.classmates.com*), Friendster (*www.friendster.com*), MySpace (*www.myspace.com*), and LinkedIn (*www.linkedin.com*) (links business contacts), which are highly popular. By looking at the profile of a person in Orkut or Hi5 and the communities/groups he/she is part of, one can get a pretty good idea about the personality of the person in question. However, privacy and safety may be a matter of concern here, including revealing information such as profiles clicked at. But since everybody is free to look at each other's networks, most users do not seem to mind revealing certain aspects of themselves in such social networks.

Collaborative Filtering and Recommender Systems

Recommender systems use the opinions of a community of users to help individuals in that community more effectively identify content of interest from a potentially overwhelming set of choices (Resnick & Varian, 1997; Herlocker, Konstan, Terveen, & Riedl, 2004). Collaborative filtering is a technology for recommender systems that includes a wide variety of algorithms for generating recommendations (Herlocker et al., 2004). While 'collaborative filtering' is a specific technique/algorithm for implementing recommender systems—a term widely used along with or synonymously with recommender systems—one should note that 'recommender systems' is the more general term. This is because recommenders may not explicitly collaborate with recipients who may be unknown to each other. Also, recommendations may suggest particularly interesting items, in addition to indicating those that should be filtered out (Resnick & Varian, 1997).

The central idea of collaborative or social filtering is to base personalized recommendations for users on information obtained from other, ideally like-minded users (Billsus & Pazzani, 1998), the underlying assumption being *'those who agreed in the past will agree again in the future'.* Collaborative filtering systems "propose a similarity measure that expresses the relevance between an item (the content) and the preference of a user. Current collaborative filtering analyzes a rating database of user profiles for similarities between users (user-based) or items (item-based)" (Wang, Pouwelse, Lagendijk, & Reinders, 2006). For example, Amazon.com has popularized item-based collaborative filtering by recommending other related books/items ("Users who bought this item also bought…”). A problem with collaborative filtering is that recommendations do not exactly correspond to how recommendations are made in social settings, where people like to refer to “experts” to look for recommendations in an area. For example, when looking for a cooking recipe of a specific community, you would want a recommendation from that community, and not from your own peers or the population as a whole (Tkatchenko, 2005). Tkatchenko also mentions the issue of privacy, the question of how to hide individual ratings and still obtain good recommendations. Recommender systems also suffer from the *cold-start problem,* that is, the problem that systems based purely on collaborative filtering cannot provide much value to their early users, and indeed cannot provide much value to new users until after they have populated their profiles (Konstan, 2004).

We have seen a number of existing collaborative mechanisms—social bookmarking and tagging, folksonomies, social network analysis, and collaborative filtering/recommender systems. Other techniques can include sharing of search queries/collaborative querying, subjective relevance judgments, collaborative digital reference services, cooperative software agents for information retrieval, and so forth. Most prior research on collaborative IR has looked at collaboration from the perspective of the user with an information need collaborating with an experienced searcher to address the former’s need (e.g., Fowell & Levy, 1995; Blake & Pratt, 2002). Systems have been developed (e.g., Procter, Goldenberg, Davenport, & McKinlay, 1998) that focus on collaboration among equally experienced members (as opposed to a novice collaborating with an expert). Work has also been done on collaborative browsing by allowing collaborators to see a trace of all the documents that users visited (e.g., Nichols et al., 2000; Twidale & Nichols, 1998; Blake & Pratt, 2002). Blake and Pratt (2002) propose a tool to support the collaborative information synthesis process used by public health and biomedical scientists.

USING CUES AND SPECIALTY SEARCH TO COLLABORATE EFFECTIVELY

Now that we have seen the reasons for collaborating and some of the existing collaborative mechanisms available, let us discuss two specific collaborative mechanisms—cues and specialty search. We will also see how these mechanisms must be placed in the context of the changing needs of the information searcher (different searcher modes) at different points in time.

Contextual Cues

One way of collaborating for search is through the usage of contextual cues. The notion of context has been introduced to enhance search tools and refers to a diverse range of ideas from specialty-search engines to personalization. Contextual information can be information related to the user’s task, the problem at hand, what the user knows, his/her domain knowledge, his/her environment, the system capabilities, his/her familiarity with the system, and so forth. There could be several instances of the term ‘context’ outside information retrieval as well. For example, in ubiquitous computing research, *context-aware computing* may be defined as any attempt to use knowledge of a user’s physical, social, informational, and even emotional state as input to adapt the behavior of one or more computational services (Abowd, Dey, Abowd, Orr, & Brotherton, 1997). Ingwersen and Jarvelin (2004) say that the searcher’s need is a complex context consisting of the perceived work task or interest, as well as perceptions and interpretations of knowledge gap and relevance, uncertainty and other emotional states, the potential sources for the solution (if any) of the work task or interest, the intentionality (i.e., goals, purposes, motivation, etc.), information preferences, strategies, pressures (costs, time), self (i.e., own capabilities, health, experiences), systematic and interactive features, and information objects.

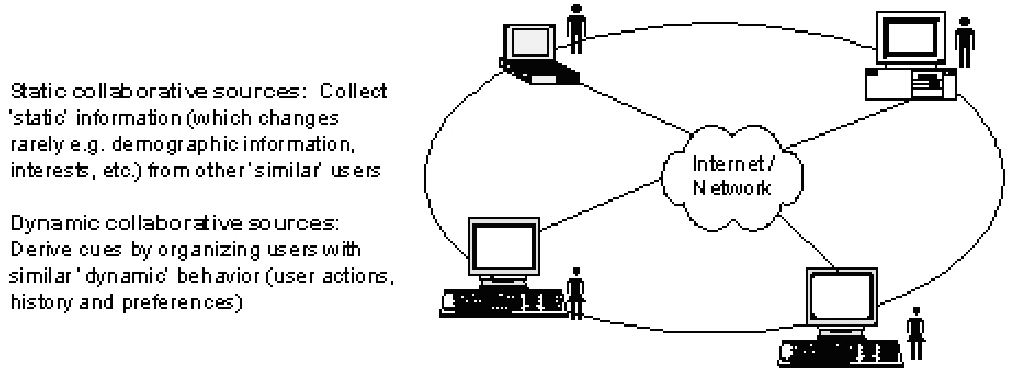
If the search system knows such attributes of the searcher, it can greatly enhance the relevance of search results and lead to a more satisfied searcher. Search would be more effective because the set of relevant results would increase, while the set of non-relevant results would decrease. However, "typically, the cost of acquiring full context is simply too high, compared to the benefits, let alone possible privacy issues" (Hawking, Paris, Wilkinson, & Wu, 2005).

Goh, Poo, and Chang (2004) propose a framework that helps to incorporate contextual cues in information systems. The framework draws on existing studies in user profiling and information filtering to suggest four sources of contextual information. User profiling is the ability to represent and reason about the interests or preferences of a user (Goh et al., 2004). Information filtering refers to tools/techniques to remove irrelevant data and present only the adequate and relevant information to the user that will satisfy his or her information requirements (Belkin & Croft, 1992). The four sources suggested by Goh et al. (2004) are *static content sources, dynamic content sources, static collaborative sources*, and *dynamic collaborative sources.* Static content sources are contextual cues derived from the information that changes rarely such as the demographic information of the user and his/her interests. Dynamic content sources are cues derived from the dynamic changes in the behavior of users, such as the user's actions, history, and preferences.

As the focus of this chapter is collaborative search, we will look at the two collaborative sources of cues—static and dynamic, in some detail.

Automated collaborative filtering systems predict a person's affinity for items or information by connecting that person's recorded interests with the recorded interests of a community of people and sharing ratings between like-minded persons. Unlike a traditional content-based information filtering system, filtering decisions are based on human and not machine analysis of content. Thus, such systems are less error prone. Each user rates items that he or she has experienced, in order to establish a profile of interests. The system then matches the user together with people of similar interests. Then ratings for those similar people are used to generate recommendations for the user. Examples of automated collaborative filtering systems are GroupLens (Konstan et al., 1997; Resnick, Iacovou, Suchak, Bergstrom, & Riedl, 1994), Ringo (Shardanand & Maes, 1995), Video Recommender (Hill et al., 1995), and MovieLens (Dahlen et al., 1998; Herlocker, Konstan, & Riedl,

Figure 4. Static and dynamic collaborative sources



2000). As collaborative filtering relies heavily on user clusters, its effectiveness depends highly on how well the clustering of profiles correlates with those of users (Goh et al., 2004).

Static and dynamic collaborative sources draw on such automated collaborative filtering systems.

Static collaborative sources are contextual cues derived from the information that changes rarely (static content sources) after organizing users with similar profiles into peer groups. These cues may either be obtained automatically (implicitly by the system) or via a user's explicit request. Every time a new user is added to the system, the system collects information about the user and constructs a user's profile that will aid the system to serve the user's needs. Terveen, McMackin, Amento, and Hill (2002) observed that users wanted novel recommendations that closely related to what they were interested in. Thus, gathering of user profile information must be supported by collaborative filtering so that users receive support in finding like-minded users.

In order to cluster static sources, users are grouped according to the static content sources such as the information the user provided during registration. Such clustering can be performed by the system using some supervised machine learning or clustering algorithm. Based on the algorithm, the system will recommend groups that the user may be interested in joining. Cues can be derived from static collaborative sources in two ways—explicit or implicit. Individuals can explicitly provide feedback on items, which can be shared with other users exhibiting similar behavior. The system can also automatically or implicitly adjust the similarity rating of a user with other users based on the matching of certain terms in the profiles of the users.

The notion of static collaborative sources can be expanded to include dynamic sources.

Dynamic collaborative sources are contextual cues derived from organizing users with similar actions and behavior into peer groups, and filtering information pertaining to the group's interest. The technique is similar to that used in static sources, but the difference is that the system performs clustering based on dynamic sources (i.e., via the user's behavior or actions), instead of simply relying on the user's profile. This could also be done in two ways—explicit or implicit. For introducing dynamic collaborative cues explicitly in the system, the system can automatically cluster a user's click stream data, recommend items of interest to the user, and allow him/her to indicate his/her interest. The system could also implicitly introduce dynamic collaborative sources by automatically adjusting the relevance of results presented to the user when the user issues a search query. Here, the relevance scores are derived from the actions and behaviors of other 'similar' users. Terveen et al.'s (2002) observation supports the use of dynamic collaborative cues for personalization in information systems.

Table 1. Strategies to derive contextual cues (Adapted from Goh et al., 2004, p. 480)

	Static Collaborative Sources	Dynamic Collaborative Sources
Explicit	User selects cluster based on work interests.	User selects search query suggested by system based on relevant items of other like-minded users.
Implicit	System clusters users based on work interests.	System clusters users' profiles based on loans and reservations. System assigns higher relevance score to items found in same cluster derived from like-minded individuals' loans and reservations.

Goh et al. (2004) identified the sources of contextual cues in an electronic repository of a library system. Static collaborative sources were identified as clustering *demographic information* of users. Dynamic collaborative sources were identified as clustering *loans and reservations information* of users. They also considered the means to incorporate these sources (see Table 1).

Goh et al. (2004) conducted an experiment with 20 subjects where they had to assess the pages returned by an electronic repository with a search engine with or without incorporating contextual cues. The results showed an improvement for a majority of users in relative search precision (improved percentage of relevant records) and an average reduction of total relevant records, or both, by incorporating cues.

In the subsection "Cues and Specialty Search in the Context of Searcher Modes," we will see how cues can be applied to a searcher with changing needs at different points in time, depending on his/her context of search. For now, let us look at another collaborative mechanism.

Specialty Search

Specialty search is another mechanism that could facilitate collaboration. Also referred to as topical search, "vertical" search, or "vortal" (Sullivan, 2000), specialty search helps provide information specific to an area or domain—for example (add dash instead of comma to break the long sentence), a search engine to be used exclusively by doctors or the medical community, saving them from having to weed out basic health/fitness information meant for the lay man and helping them focus on specific issues like the latest advances in medical science or medical job opportunities.

As highlighted by Lawrence and Giles (2000), the coverage of a general-purpose search engine is limited. Bharat and Broder's (1998) study saw the largest search engines covering just 50% of all Web pages, with a maximum overlap of 30%. Mori and Yamada (2000) also contend that a user cannot search well based on a single general search engine. If the big search engines are unable to deliver comprehensive access to the entire Web, perhaps the time has come for more focused sites to offer near comprehensiveness in their own chosen fields (Kawin, 2003; Khoussainov & Kushmerick, 2003; Battelle, 2004; Sullivan, 2000).

Specialty search can be considered an extension of an important Internet phenomenon—virtual communities, where groups of people communicate, interact, and collaborate with each other, often with a commonality of interest or intent (see discussion on communities of practice in the subsection "Why Collaborative Search?"). It is now much easier to build such virtual communities without much technological know-how, and a lot of these spring forth binding informal groups together. Examples of systems catering to such virtual communities are online groups, discussion forums, and the newly coined Weblogs or blogs. Online groups and discussion forums usually evolve from a need to share knowledge on a common platform. Blogs, on the other hand, usually cater to a group of readers, with the 'bloggers' deciding on the subjects of interest and contributing most of the content. Online groups, discussion forums, and blogs are not information retrieval systems or specialty search engines as such, but they do help bring a diverse group of people together to collaborate in different ways to share information. A lot of the information retrieved is through answers from *human* sources to queries put across on a forum or newsgroup.

Specialty search engines, on the other hand, could be considered as more formal and perhaps better organized than a lot of informal virtual communities. There are a large number of specialty search engines today. Gordon and Pathak (1999) say that of the 1,800 or so search engines estimated in 1997, most of those were specialty search engines that only cover a specific subject like automobiles or sports. Table 2 lists a small number of specialty search engines where different sets of individuals can collaborate socially for

Table 2. Specialty search engines

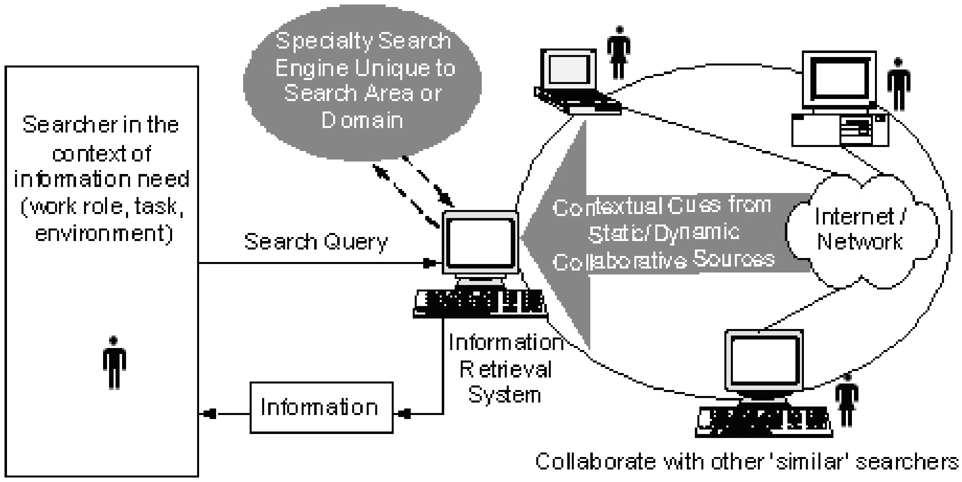
Specialty	Collaborators	Specialty Search Engines
Science	Individuals, students, teachers interested in science	Scirus (*www.scirus.com/srsapp/*); Sciseek (*www.sciseek.com*); Search4science (*www.search4science.com*)
Medical	Doctors, medical students, healthcare workers	HONMedhunt (*www.hon.ch/MedHunt*); MedicineNet (*www.medicinenet.com/script/main/hp.asp*); MedlinePlus (*http://medlineplus.gov/*); WebMD (*www.webmd.com/*)
Biology	Biology students, teachers, professionals	Biocrawler (*www.biocrawler.com*)
Chemistry	Chemistry students, teachers, specialists	Chemie.DE (*www.chemie.de/*)
Mathematics	Mathematicians, students, teachers interested in math	IntegerSequences (*www.research.att.com/~njas/sequences/*)
Civil Engineering	Civil Engineers	iCivilEngineer (*www.icivilengineer.com*)
Law	Lawyers, advocates, those involved in court cases	FindLaw (*http://lawcrawler.findlaw.com*); Law.com (*www.law.com*)
Art	Artists, art lovers, gallery managers, art sellers	Art-Bridge (*www.art-bridge.com/directory/abdir.htm*); Artcyclopedia (*www.artcyclopedia.com*); MuseumStuff [specific to museums] (*www.museumstuff.com*)
Finance	Financial analysts, brokers, businessmen	Business.com (*www.business.com*); Inomics (*www.inomics.com*); DailyStocks (*www.dailystocks.com*); TradingDay (*www.tradingday.com*); EarningsBase (*www.earningsbase.com*); MoneyWeb (*www.moneywebsearch.com*)
Research Papers	Academia, researchers, PhD candidates	GoogleScholar (*http://scholar.google.com*); CiteSeer (*http://citeseer.ist.psu.edu*)
Journalism	Journalists, reporters	Journalist's Toolbox (*www.americanpressinstitute.org/pages/toolbox/*)
Maps/Atlas	Geography teachers, students, individuals looking for maps	MapsArea (*www.mapsarea.com*)
Books	Students, professors, researchers	The Online Books Page (*http://digital.library.upenn.edu/books/*); AddALL (*www.addall.com*)
Jobs/Employment	Job Seekers, Employers	Monster (*www.monster.com*); JobWeb (*www.jobweb.com*); CareerBuilder (*www.careerbuilder.com*); JobsDB (*www.jobsdb.com*) (Singapore); BioView [specific to Biotechnology/Life Sciences] (*www.bioview.com/bv/servlet/BVHome*)
Origin	Genealogists	TheOriginsNetwork (*www.originsnetwork.com*)
Alumni	Alumni of a school, university, or institution; former school friends	Classmates (*www.classmates.com*); FriendsReunited [specific to UK] (*www.friendsreunited.co.uk*)

continued on following page

Table 2. continued

Specialty	Collaborators	Specialty Search Engines
Research Papers	Academia, researchers, PhD candidates	GoogleScholar (*http://scholar.google.com*); CiteSeer (*http://citeseer.ist.psu.edu*)
Journalism	Journalists, reporters	Journalist's Toolbox (*www.americanpressinstitute.org/pages/toolbox/*)
Maps/Atlas	Geography teachers, students, individuals looking for maps	MapsArea (*www.mapsarea.com*)
Books	Students, professors, researchers	The Online Books Page (*http://digital.library.upenn.edu/books/*); AddALL (*www.addall.com*)
Jobs/Employment	Job Seekers, Employers	Monster (*www.monster.com*); JobWeb (*www.jobweb.com*); CareerBuilder (*www.careerbuilder.com*); JobsDB (*www.jobsdb.com*) (Singapore); BioView [specific to Biotechnology/Life Sciences] (*www.bioview.com/bv/servlet/BVHome*)
Origin	Genealogists	TheOriginsNetwork (*www.originsnetwork.com*)
Alumni	Alumni of a school, university, or institution; former school friends	Classmates (*www.classmates.com*); FriendsReunited [specific to UK] (*www.friendsreunited.co.uk*)
Weather	Travelers, weather forecasters, individuals	WeatherBug (*www.weatherbug.com*)
Mobile	Mobile phone users	SomewhereNear [specific to UK] (*http://somewherenear.com*); Waply (*www.waply.com*)
Travel	Travelers, hoteliers, travel agents, airline companies	Kayak (*www.kayak.com*); Mobissimo (*www.mobissimo.com*); SideStep (*www.sidestep.com*); Orbitz (*www.orbitz.com*); Expedia (*www.expedia.com*); Travelocity (*www.travelocity.com*); DoHop (*www.dohop.com*); IGoUGo (*http://igougo.com*); Travelazer (*www.travelazer.com*)
Dogs	Dog lovers/sellers	Doginfo (*www.doginfo.com*)
Shopping	Shoppers, sellers	BizRate (*www.bizrate.com*); Kelkoo (*www.kelkoo.co.uk*); NexTag (*www.nextag.com*); PriceGrabber (*www.pricegrabber.com*); PriceSCAN (*www.pricescan.com*)

Figure 5. Cues and specialty search



information retrieval in their specialties of interest (Sullivan, 2002; Meyer, 2006; VanFossen, 2005; ISEdb, 2005; Hofstede, 2000).

Figure 5 shows the concepts of cues and specialty search incorporated to Figure 3 discussed earlier. While contextual cues are derived from various collaborative sources *similar* to the user and his/her needs in order to make search results more relevant, specialty search engines serve as repositories/search interfaces specific to a particular domain which a searcher can utilize when searching for something in that particular area or domain.

So far, we have discussed aspects related to the *system*, that is, existing social information retrieval mechanisms, including two specific types of collaborative search mechanisms—cues and specialty search. Let us now look at search from the point of view of the *searcher* or the *user.* An aspect basic to any searcher of information is his or her changing needs.

Changing Needs of the Searcher: Four Searcher Modes

It is extremely important for a search service provider to understand the reasons and circumstances surrounding a search in order to truly satisfy the user. Not only are the goals behind a user's search query (requirement of specific information) important, the user's prior domain knowledge in the area of search must also be taken into account to carry out an effective search and output of results.

Agarwal and Poo (2006) refer to Papadopoullos (2004), who states that search and classification results must satisfy four basic categories of users. We term these four categories of users (searcher modes) as: (a) novice, (b) data gatherer, (c) known-item searcher, and (d) focused searcher. Depending on the context of data one is searching for and the domain knowledge the person has in the field of search, the same person may be in one of the four modes (see Table 3). The surfer is not looking for anything in particular and is just entertaining himself, so we do not count him in our typology of searcher modes. We are interested in addressing the needs of users performing goal-oriented search. For example, a student would be in novice mode when searching for course-related information, a medical doctor searching for latest advances in medicine would be in the data-gathering mode, and a researcher locating a research paper based on the author's name, publication,

Table 3. Four modes/activities of information searchers

No.	Searcher mode during a particular search	Searcher need during a particular search	Prior domain knowledge	Requirement of specific information
1	Novice	Needs information about a topic he is not familiar with in preparation for starting a new project	0 (no)	0 (not yet)
2	Data Gatherer	Needs information about a topic he *is* knowledgeable about and is therefore in data-gathering mode	1 (yes)	1 (yes)
3	Known-Item Searcher	Has a good idea what he is looking for, knows that a given document or piece of data exists, and simply needs to locate it	X (do not care)	1 (yes)
4	Focused Searcher	Needs a very specific answer to a specific question	X (do not care)	1 (yes)
-	Surfer	Does not need anything in particular; searching purely for entertainment	X (do not care)	0 (no)

Table 4. Searcher modes based on domain knowledge and requirement of specific information

Prior domain knowledge	Requirement of specific information	Searcher mode
0 (no)	0 (no)	Novice or Surfer
0 (no)	1 (yes)	Known-Item Searcher or Focused Searcher
1 (yes)	0 (no)	Surfer
1 (yes)	1 (yes)	Data Gatherer, Known-Item Searcher, or Focused Searcher

Table 5. Searcher modes based on five factors

Goal-oriented search	Prior-domain knowledge	Requirement of specific information	Focused search	Known-item search	Searcher mode
0	X	0	X	X	Surfer
1	0	0	X	X	Novice
1	X	1	1	0	Focused searcher
1	X	1	0	1	Known-Item searcher
1	1	1	0	0	Data Gatherer

and year would be in the known-item searching mode. A focused search would require a specific answer to a question, for example, "What are the differences between qualitative and quantitative data in information systems research?" A bored teenager searching for celebrity gossip would be in the surfer mode.

An understanding of the different searcher modes is extremely important to understand the differing needs of searchers at different points in time. Such an understanding will facilitate users in different searcher modes to collaborate effectively for search.

Table 4 below shows the searcher modes based on domain knowledge and requirement of specific information. Again (also in Table 5), the surfer is included for completeness but is outside the scope of this study.

From Table 4, we notice that there is ambiguity between novice or surfer modes (case 00); between known-item searcher and focused searcher modes (case 01); and between data gatherer, known-item searcher, and focused searcher modes (case 11). To resolve this, we add three more factors:

- Looking for anything in particular? *(goal-oriented search)*
- Looking for something you have seen before and know that it exists? *(known-item search)*
- Need a specific answer to a specific question? *(focused search)*

From Table 5, we can see that a surfer mode implies casual search, which is not goal oriented. Within goal-oriented searches, absence of (or insignificant in the view of the searcher) prior domain knowledge and non-requirement of specific information determines the novice mode. A focused search must be goal oriented, requires specific information, and has a specific question needing a specific answer. A known-item search is goal oriented, requires specific information, and is a case where the item has been encountered before and simply needs locating. The data-gathering mode is goal-oriented where the searcher has prior domain knowledge and requires specific information.

Cues and Specialty Search in the Context of Searcher Modes

As mentioned, when we talk about collaborative mechanisms such as specialty search engines and contextual cues, these are mostly provided by the information retrieval system and are from a system perspective. There is a need to understand these mechanisms from a searcher perspective, based on his or her differing needs at different points in time.

Incorporating contextual cues from static/dynamic content or collaborative sources should benefit all the four searcher modes by increasing the set of relevant results and decreasing the set of non-relevant results. Usefulness to the searcher in data-gathering mode may range from medium to high depending on the cues obtained from dynamic collaborative sources by matching the actions of the data gatherer with those of others with similar domain knowledge.

Specialty search would be extremely useful to the searcher in data-gathering mode, as he would be able to access the search engine directly relevant to his prior domain knowledge and the domain of search. It would greatly benefit the focused searcher too, perhaps by pulling out answers from the FAQ (Frequently Asked Questions) section of a specialized portal. It might have medium to high utility for the novice depending on whether the specialty search engine provides background or basic information that could be understood by a novice. Specialization would have low utility for the known-item searcher, unless the known-item he is seeking resides within the specialty search engine.

We can also try to map the other collaborative mechanisms discussed with the four searcher modes. Collaborative tagging and folksonomies should be highly useful to the novice who can

Table 6. Usefulness of collaborative mechanisms to the different searcher modes

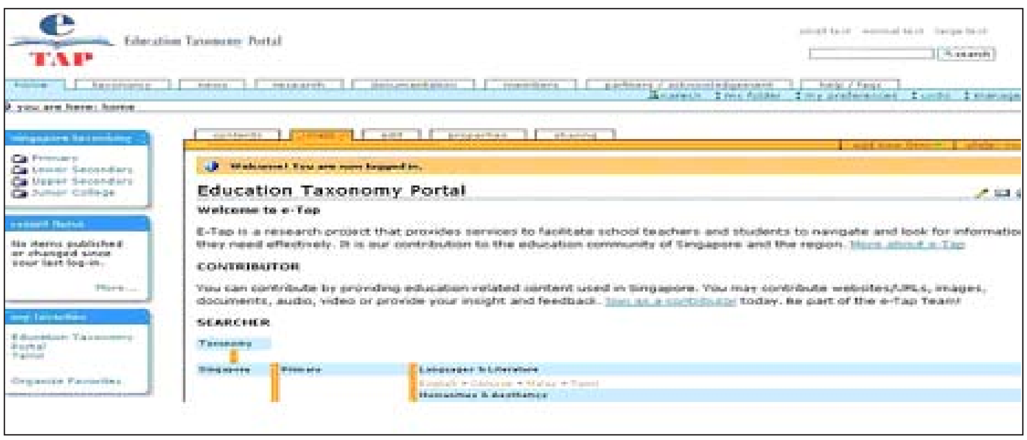
Usefulness: ↓ Low ~ Medium ↑ High	Searcher Modes			
Collaborative Search Mechanisms	Novice	Data Gatherer	Known-Item Searcher	Focused Searcher
Contextual Cues/Recommendation Systems	↑	~ ↑	↑	↑
Specialty Search	~ ↑	↑	↓	↑
Other Collaborative Mechanisms				
Collaborative Tagging/Folksonomies	↑	↑	↓	↓
Social Bookmarking	↑	↑	~ ↑	↓
Social Networking	↑	↑	↓	↓

search based on tags put by other expert users. A focused searcher can help to tag content but may find limited use of folksonomies, unless there is content that has been tagged by other focused searchers and can help answer the focused searcher's question. It should also be useful to a data gatherer who can gather data based on tags put forth by other users. It should have limited applicability for the known-item searcher.

Social bookmarking should greatly benefit the novice as well as the data gatherer, who can access relevant links based on bookmarks by other expert searchers. The known-item searcher should find medium to high usefulness for social bookmarking in tracking down content that he/she has encountered before. It could be less useful for the focused searcher who needs a specific answer to his/her question, which may not be provided by social bookmarks.

Lastly, social networking should greatly interest the novice, who can get the profiles of other searchers with similar needs or who are experts in the area of search. It should help a data gatherer by linking him/her with someone else within his/her domain of search. It should have low utility for the known-item searcher, unless the specific item he/she is looking for is part of the social network. It may not be very useful for the focused searcher as well, unless he is linked to a person who can answer his/her question.

Figure 6. Snapshot of ETaP: Education Taxonomy Portal (http://etap.comp.nus.edu.sg)



Let us now look at a case study to illustrate the collaborative mechanisms we have seen, as well as the four searcher modes, in the context of education.

COLLABORATING FOR EDUCATION-RELATED SEARCH: A CASE STUDY

The Education Taxonomy Portal (ETaP) is an online digital repository being developed for the Singapore Education community; it is still in its infancy. Accessible from http://etap.comp.nus.edu.sg (see Figure 6), ETaP provides services to facilitate schoolteachers and students to collaborate in contributing, searching, navigating, and retrieving education-related content effectively. Information retrieved is specific to users' local needs while enabling them to contribute and share their contents. Apart from search, a taxonomy based on the prescribed education curriculum helps provide browsing facilities.

Singapore teachers looking on the Internet for teaching materials and information relevant to their courses are almost always presented with a huge amount of data. Gathering required information is a time-consuming process, which may take hours. Students who want to search for information for project work or to supplement their course materials are similarly presented with a huge array of non-relevant data.

There are many education-related professionals, teachers, and schools that, in the past couple of years, have compiled their own frequently used education material as well as useful links gathered while browsing. Different organizations/individuals have their own small repositories. The project aims to provide a countrywide repository for gathering such material (Web sites, images, audio, video, journals, etc.) and classifying it into different categories for effective search.

ETaP is targeted specifically for Singapore. The scope will subsequently be expanded to include other countries in the ASEAN region. The portal aims to help teachers, students, parents, and all associated with the education community in Singapore to collaborate and perform quality search to be better satisfied with their search results. ETaP is available free for everyone's use.

The four searcher modes will be built into ETaP. A searcher will be able to specify whether he or she is a novice, data gatherer, known-item searcher, or focused searcher, depending on the context of search.

We intend to apply the different types of *contextual cues* in ETaP. Static content sources can be added by utilizing a database from participating schools containing users' information (name/majors). Dynamic content sources can be captured using a system that logs the users' actions. Users of the system can create a record of users whom they know so as to utilize the contextual cues that can be obtained from static collaborative sources. With such information, dynamic collaborative sources can also be obtained by matching the actions of the users with those of users with similar interests.

ETaP could implement collaborative filtering and serve as a recommender system. Educational resources found useful and recommended by students of a particular batch are likely to be useful to the next batch of students the following year. Similarity between users can be based on the grade of the student when the recommendation was made. For example, a Secondary 3 student recommends a Web site as useful for the Physics exam paper of Secondary 3. The following year, ETaP can recommend the Web site to another Secondary 3 student based on its usefulness, and having Secondary 3 as the similarity measure. Recommendation could also be based on expertise. A teacher recommending a study material can be viewed as useful by a student looking for material on a certain area.

ETaP provides *specialization* by focusing on the education domain. All the information is specific to the needs of students, teachers, as well as other stakeholders such as parents and tutors, owners of tutoring agencies, and so forth. Anybody who has an interest in an educational aspect can come to ETaP and get specific results, instead of searching a general-purpose search engine. The portal will eventually be expanded to gather relevant education-related material from major search engines and combine the result set based on user needs (specialty search).

Other collaborative mechanisms could also be implemented in ETaP.

With *collaborative tagging* and *folksonomies* implemented in ETaP, students could tag various resources, Web sites, and educational materials based on their specific needs. For example, a specific Algebra tutorial could be variously tagged as 'Algebra', 'Good tutorial on Algebra', 'Important for exams', and so forth. Teachers could tag the resources depending on how relevant they are to the subjects they teach. Teachers could also learn about the way students perceive educational resources from the way they tag them. This could help them in their own teaching. A study of the folksonomy of educational content tagged by students and teachers might provide useful insights into education from a student perspective, and help in evolving educational curriculum, method, and techniques. Implementation of collaborative tagging in ETaP will also require it to be able to group synonymous tags together and to moderate for spam.

The storing and categorizing of Web page links (pointing to educational resources), along with other content such as documents, images, tutorials, and so forth, can be likened to *social bookmarking*. Tools can be developed for the browser that allows a student to automatically bookmark a relevant educational site in ETaP and suggest categories under which the bookmark can be stored.

Social networking would also find useful application in ETaP. A student's profile could contain links to the profiles of other students in the school, as well as to teachers he/she has been taught by. The system could allow students to participate in communities of specific subjects and projects. By looking at a student profile, a new visitor will be able to know the subjects the student has taken in the past. Those willing (and not having privacy concerns) may also share their grades on the subjects to reflect level of expertise (or perhaps only those with high grades may choose to share their grades, as they are less likely to have inhibitions in sharing their grades). This will help connect new students not just to content and resources, but also to those students who have taken certain subjects in the past, and have received good grades. This will help enhance the support network among students. Peers could collaborate to form study groups. There could be sections and communities on previous-year exam papers related to specific subjects that students can try to solve and which can be overseen by teachers. Schools utilizing such social networks for education can allocate a small percentage of marks towards the level of participation and collaboration displayed by students in the educational network. This will help foster a sense of sharing among students.

With its focus on collaborative searching and retrieval, ETaP aims to bring together a diverse range of people (teachers, students, parents, etc.) to collaborate effectively for knowledge sharing.

While a novice would provide more questions than answers, he or she could contribute by bringing in new insights and different ways of looking at a topic or problem. A data gatherer in any educational topic (e.g., a student embarking on a new school project) would benefit from the repository of past experience in doing such projects, and from teachers or seniors more experienced in the area. Focused searchers and known-item searchers are also likely to be experienced in the different areas of the educational domain, and thus are useful to the novice searcher in providing answers and ways of approaching a problem at hand.

ETaP, we hope, will serve as an example to show how collaborative searching is enhanced when the needs of the searcher in different modes are matched with the right collaborative mechanism that allows collaborating with like-minded searchers.

FUTURE TRENDS

The world is at a phase where countries are coming together and collaborating, drawn by the power of economics and common goals, rather than traditional political and military agendas. ASEAN, European Union, and Free Trade Agreements among countries are a case in point. An unprecedented binding factor to bringing diverse thoughts and ideas together has been the connectivity provided by the Internet. Social approaches to information searching seeks to harness the most important phenomenon arising out of the growth of the Internet—bringing people of diverse nationalities, temperaments, personalities, and needs together in one common network. However, the Internet has also brought in ‘info-glut’, where too much information puts a huge cost on time and money, leaving individuals to sieve the important from the unimportant, the wanted from the unwanted, and the relevant from the non-relevant.

By harnessing the power of social networks, collaborative search mechanisms will make information comprehension easier and help reduce associated costs. The future will only see more and more collaborative mechanisms built into information retrieval systems.

The next generation of search engines will not just provide personalized searches, which will take into account the user’s prior domain knowledge, experience in the area of search, experience with the search technology or search engine, and the searcher’s task at hand; the user’s interests and social affiliations will play a major role too. Virtual experts will be at hand to solve the problems faced by the majority. Already there are initiatives such as the Knowledge-Community (K-Comm) project (*http://kcomm.redirectme.net*) at the National University of Singapore and About-Experts (*http://experts.about.com*). K-Comm is an initiative by the authors of this chapter to harness the tacit knowledge residing in different individuals. By recognizing that every individual is good at and has experience in some area or the other, K-Comm seeks to harness the latent expertise hidden in every individual and brings out a feeling of self-worth in everyone. This feeling is enhanced as users share more and more of their knowledge with others and collaborate to seek, as well as contribute to knowledge.

The experiments and approaches so far hold a lot of promise for collaborative search. The tremendous success of communities of practice shows how collaboration comes across naturally in the virtual world and can easily be extended to search. Active research and implementation will see the benefits reach all seekers of information.

As highlighted in this chapter, future search engines and information retrieval systems must also take into account the varying needs of the searcher at different points in time, and build collaborative mechanisms to serve that need.

From the perspective of this chapter, more research is needed into static and dynamic sources of collaborative cues, as well as the phenomenon of specialty search, to align them with different searcher modes and to best search the varying needs of the searcher. This alignment with searcher modes could also be explored further for the other collaborative mechanisms discussed.

CONCLUSION

In this chapter, we have seen how the search engines and search mechanisms of today are good but not the ideal. Information overload and difficulties in query formulation remain a major problem, and an average search still takes about

11 minutes. We posit that collaborative approaches to searching will provide an important way to help a user connect the dots and make sense of information.

For collaboration to be successful, the collaborator or collaborative mechanisms must be able to help the searcher either through expertise or experience, or similarity of needs. While providing collaborative mechanisms, an information provider must also take into account the different modes a searcher is in at different points in time (based on the characteristics of the task at hand or the qualities/expertise of the searcher), and provide technology features that match the task and searcher characteristics reflected from the user mode.

In this chapter, we reviewed concepts related to collaboration, as well as existing collaborative mechanisms that are finding a high level of success and are being widely adopted. We also discussed two collaborative mechanisms—cues and specialty search. These can be built into a system and will allow a searcher to retrieve information collaboratively with other like-minded searchers. However, simply building collaborative mechanisms is not enough. These mechanisms must also be viewed in the context of the different modes a searcher is in at different points in time.

An illustration was provided using a case study of an educational taxonomy portal.

It is our sincere hope that the world will see collaboration in more and more spheres, including the common, but ubiquitous activity of looking for information.

REFERENCES

Abowd, G.D., Dey, A.K., Abowd, G., Orr, R., & Brotherton, J. (1997). *Context-awareness in wearable and ubiquitous computing.* GVU Technical Report GIT-GVU-97-11. Retrieved September 2, 2006, from http://www-static.cc.gatech.edu/fce/pubs/iswc97/wear.html

Agarwal, N.K., & Poo, D.C.C. (2006). Meeting knowledge management challenges through effective search. *International Journal of Business Information Systems, 1*(3), 292-309.

Alavi, M., & Leidner, D.E. (2001). Review: Knowledge management and knowledge management systems: Conceptual foundations and research issues. *MIS Quarterly, 25*(1), 107-136.

Asmus, J., Bonner, C., Esterhay, D., Lechner, A., & Rentfrow, C. (2005, December 6). *Instructional design technology trend analysis.* Denver: University of Colorado. Retrieved September 1, 2006, from http://elgg.net/collinb/files/1136/2967/TrendAnalysisWeb.pdf

Battelle, J. (2004). GlobalSpec: Domain specific search and the semantic Web. *John Battelle's Searchblog.* Retrieved May 24, 2006, from http://battellemedia.com/archives/000519.php

Belkin, N.J. (1984). Cognitive models and information transfer. *Social Science Information Studies, 4*(2-3), 111-129.

Belkin, N.J. (2000). Helping people find what they don't know. *Communications of the ACM, 43*(8), 58-61.

Belkin, N.J., & Croft, B.W. (1992). Information filtering and information retrieval: Two sides of the same coin? *Communications of the ACM, 35*(12), 29-36.

Bharat, K., & Broder, A.Z. (1998). A technique for measuring the relative size and overlap of Web search engines. *Proceedings of the 7th International World Wide Web Conference (WWW7)* (pp. 379-388).

Billsus, D., & Pazzani, M. (1998). Learning collaborative information filters. *Proceedings of the 15th International Conference on Machine Learning* (pp. 46-54). San Francisco: Morgan Kaufmann. Retrieved May 28, 2006, from http://www.ics.uci.edu/~pazzani/Publications/MLC98.pdf

Blake, C., & Pratt, W. (2002, November 18-21). Collaborative information synthesis. *Proceedings of ASIST 2002* (pp. 44-56), Philadelphia.

Case, D.O. (2002). *Looking for information: A survey of research on information seeking, needs, and behavior.* San Diego: Academic Press, Elsevier.

Chi, E.H., & Pirolli, P. (2006, February). Social information foraging and collaborative search. *Proceedings of the HCIC Workshop,* Fraser, CO.

Cox, A. (2005). What are communities of practice? A comparative review of four seminal works. *Journal of Information Science, 31*(6), 527-540.

Cross, R., Borgatti, S.P., & Parker, A. (2002). Making invisible work visible: Using social network analysis to support strategic collaboration. *California Management Review, 44*(2), 25-46.

Dahlen, B.J., Konstan, J.A., Herlocker, J.L., Good, N., Borchers, A., & Riedl, J. (1998). *Jump-starting movielens: User benefits of starting a collaborative filtering system with 'dead data'.* University of Minnesota TR 98-017.

Deerwester, S., Dumais, S.T., Furnas, G.W., Landauer, T.K., & Harshman, R. (1990). Indexing by latent semantic analysis. *Journal of the Society for Information Science, 41*(6), 391-407.

Dretske, F. (1981). *Knowledge and the flow of information*. Cambridge, MA: MIT Press

Ellis, D., Allen, D., & Wilson, T.D. (1999). Information science and information systems: Conjunct subjects, disjunct disciplines. *Journal of the American Society for Information Science, 50*(12), 1095-1107.

Fang, X., Chen, P., & Chen, B. (2005). User search strategies and search engine interface design. In R.W. Proctor & K.-P.L. Vu (Eds.), *Handbook of human factors in Web design* (pp. 193-210). Mahwah, NJ: Lawrence Erlbaum.

Fayyad, U.M., Piatetsky-Shapiro, G., Smyth, P., & Uthurusamy, R. (Eds.). (1996). *Advances in knowledge discovery and data mining.* Menlo Park, CA: AAAI Press.

Fowell, S., & Levy, P. (1995). Developing a new professional practice: A model for networked learner support in higher education. *Journal of Documentation, 51*(3), 271-280.

Freyne, J., Smyth, B., Coyle, M., Balfe, E., & Briggs, P. (2004). Further experiments on collaborative ranking in community-based Web search. *Artificial Intelligence Review, 21*(3-4), 229-252.

Fu, L., Goh, D.H.-L., Foo, S.S.-B., & Supangat, Y. (2004, September 12-17). Collaborative querying for enhanced information retrieval. *Proceedings of ECDL 2004* (pp; 378-388), Bath, UK.

Goh, J.M., Poo, D.C.C., & Chang, K.T.T (2004, July 8-11). Incorporating contextual cues into electronic repositories. *Proceedings of the 8th Pacific-Asia Conference on Information Systems* (pp. 472-484), Shanghai, China.

Golder, S., & Huberman, B.A. (2005). *The structure of collaborative tagging systems.* Retrieved May 28, 2006, from http://www.hpl.hp.com/research/idl/papers/tags/tags.pdf

Google Technology. (2004). Our search: Google technology. *About Google*. Retrieved May 27, 2006, from *http://www.google.com/technology/*

Gordon, M., & Pathak, P. (1999). Finding information on the World Wide Web: The retrieval effectiveness of search engines. *Information Processing and Management, 35,* 141-180.

Green, A. (1990). What do we mean by user needs? *British Journal of Academic Librarianship, 5,* 65-78.

Hammond, T., Hanny, T., Lund, B., & Scott, J. (2005). Social bookmarking tools (I). *D-Lib Magazine, 11*(4). Retrieved May 28, 2006, from http://www.dlib.org/dlib/april05/hammond/04hammond.html

Hawking, D., Paris, C., Wilkinson, R., & Wu, M. (2005, August 9). Context in enterprise search and delivery. *Proceedings of the ACM SIGIR 2005 Workshop on Information Retrieval in Context* (IRiX) (pp. 14-16), Salvador, Brazil.

Herlocker, J.L., Konstan, J.A., & Riedl, J. (2000, December 2-6). Explaining collaborative filtering recommendations. *Proceedings of the 2000 ACM Conference on Computer Supported Cooperative Work (CSCW'00)* (pp. 241-250). Philadelphia.

Herlocker, J.L., Konstan, J.A., Terveen, L.G., & Riedl, J.T. (2004). Evaluating collaborative filtering recommender systems. *ACM Transactions on Information Systems, 22*(1), 5-53.

Hill, W., Stead, L., Rosenstein, M., & Furnas, G. (1995a). Recommending and evaluating choices in a virtual community of use. *Proceedings of the Human Factors in Computing Systems (CHI '95) Conference* (vol. 1, pp.194-201). New York: ACM.

Hill, W., Stead, L., Rosenstein, M., & Furnas, G.W. (1995b). Recommending and evaluating choices in a virtual community of use. *Proceedings of the Conference on Human Factors in Computing Systems* (CHI'95) (pp. 194-201). Denver, CO.

Hofstede, M. (2000). A collection of special search engines. *Universiteitsbibliotheek.* Retrieved May 28, 2006, from http://www.leidenuniv.nl/ub/biv/specials.htm

Ingwersen, P., & Jarvelin, K. (2004, July 29). Information retrieval in contexts. *Proceedings of the ACM SIGIR 2004 Workshop on Information Retrieval in Context* (pp. 6-9), Sheffield, UK.

ISEdb (Internet Search Engine Database). (2005). *Specialty search engines.* Retrieved May 28, 2006, from http://www.isedb.com/html/Internet_Search_Engines/Specialty_Search_Engines/

Karamuftuoglu, M. (1998). Collaborative information retrieval: Toward a social informatics view of IR interaction. *Journal of the American Society for Information Science, 49*(12), 1070-1080.

Kawin Interactive. (2003, January 1). *Kawin Vortalbuilding.com: Your gateway to the universe.* Retrieved May 24, 2006, from http://elibrary.line56.com/detail/RES/1098686731_907.html

Khoussainov, R., & Kushmerick, N. (2003). Learning to compete in heterogeneous Web search environments. *Proceedings of the 18th International Joint Conference on Artificial Intelligence (IJCAI-03).*

Komarjaya, J., Poo, D.C.C., & Kan, M.-Y. (2004, September 12-17). Corpus-based query expansion in online public access catalogs. *Proceedings of ECDL 2004* (pp. 221-231). Bath, UK.

Konstan, J. (2004). Introduction to recommender systems: Algorithms and evaluation. *ACM Transactions on Information Systems, 22*(1), 1-4.

Konstan, J.A., Miller, B.N., Maltz, D., Herlocker, J.L., Gordon, L.R., & Riedl, J. (1997). GroupLens: Applying collaborative filtering to Usenet news. *Communications of the ACM, 40*(3), 77-87.

Lancaster, F.W. (1968). *Information retrieval systems: Characteristics, testing and evaluation.* New York: John Wiley & Sons.

Lawrence, S., & Giles, C.L. (2000). Accessibility of information on the Web. *Intelligence, 11*(1), 32-39.

Lederman, P. (2005). Implementing a taxonomy solution. *AIIM E-Doc, 19*(2), 25-26.

Levy, D.M., & Marshall, C.C. (1994). What color was George Washington's white horse? A look at assumptions underlying digital libraries. *Proceedings of Digital Libraries '94* (pp. 163-169). College Station, TX.

Machlup, F. (1980). *Knowledge: It's creation, distribution, and economic significance* (vol. 1). Princeton, NJ: Princeton University Press.

Marchionini, G. (1997). *Information seeking in electronic environments*. Cambridge, UK: Cambridge University Press (Cambridge Series on Human-Computer Interaction).

Mathes, A. (2004, December). Folksonomies: Cooperative classification and communication through shared metadata. *Proceedings of the Computer Mediated Communication Doctoral Seminar* (LIS590CMC), Urbana-Champaign, IL. Retrieved September 2, 2006, from http://www.adammathes.com/academic/computer-mediated-communication/folksonomies.html

Meyer, H.A. (2006). *Topic search engines*. Retrieved May 28, 2006, from http://www.all-searchengines.com/

Mika, P. (2005, November 6-10) Ontologies are us: A unified model of social networks and semantics. *Proceedings of the 4th International Semantic Web Conference* (ISWC 2005), Galway, Ireland. Retrieved September 2, 2006, from http://www.cs.vu.nl/~pmika/research/papers/ISWC-folksonomy.pdf

Mitchell, J.C. (1969). The concept and use of social networks. In J.C. Mitchell (Ed.), *Social networks in urban situations* (pp. 1-50). Manchester, UK: University of Manchester Press.

Morahan-Martin, J., & Anderson, C.D. (2000). Information and misinformation online: Recommendations for facilitating accurate mental health information retrieval and evaluation. *CyberPsychology & Behavior, 3*(5), 731-746.

Mori, M., & Yamada, S. (2000, October 30-November 4). Adjusting to specialties of search engines using Meta Weaver. *Proceedings of the World Conference on the WWW and Internet (WebNet 2000)* (pp. 408-412). San Antonio, TX.

Morville, P. (2002, February 21). *Social network analysis*. Retrieved August 30, 2006, from http://semanticstudios.com/publications/semantics/000006.php

Narayanan, S., Koppaka, L., Edala, N., Loritz, D., & Daley, R. (2004). Adaptive interface for personalizing information seeking. *CyberPsychology & Behavior, 7*(6), 683-688.

Nichols, D.M., Pemberton, D., Dalhoumi, S., Larouk, O., Belisle, C., & Twidale, M.B. (2000, September 18-20). DEBORA: Developing an interface to support collaboration in a digital library. *Proceedings of ECDL 2000* (pp. 239-248). Lisbon, Portugal.

North, E.J., North, J., & Benade, S. (2004). Information management and enterprise architecture planning—a juxtaposition, *Problems and Perspectives in Management, 4,* 166-179.

Papadopoullos, A. (2004). Answering the right questions about search. *EContent leadership series: Strategies for…search, taxonomy & classification* (supplement to July/August 2004 *EContent and Information Today*, pp. S6-S7). Retrieved May 29, 2006, from http://www.procom-strasser.com/docs/Convera_Right_Questions.pdf

Procter, R., Goldenberg, A., Davenport, E., & McKinlay, A. (1998). Genres in support of collaborative information retrieval in the virtual library. *Interacting with Computers, 10*(2), 157-175.

Pujol, J.M., Sanguesa, R., & Bermudez, J. (2003, May 20-24). Porqpine: A distributive and collaborative search engine. *Proceedings of the 12th International World Wide Web Conference* (WWW2003, p. S25), Budapest, Hungary. Retrieved May 27, 2006, from http://www2003.*org/cdrom/papers/poster/p341/p341-pujol.html.html*

Resnick, P., & Varian, H.R. (1997). Recommender systems. *Communications of the ACM, 40*(3), 56-58.

Resnick, P., Iacovou, N., Suchak, M., Bergstrom, P., & Riedl, J. (1994). GroupLens: An open architecture for collaborative filtering of Netnews. *Proceedings of the 1994 Conference on Computer*

Supported Collaborative Work (CSCW'94) (pp. 175-186).

Robertson, S.E. (1981). The methodology of information retrieval experiment. In K. SparckJones (Ed.), *Information retrieval experiment* (pp. 9-31). Butterworths.

Scott, J. (2000). *Social network analysis: A handbook* (2nd ed.). London/Thousand Oaks, CA: Sage.

Shardanand, U., & Maes, P. (1995). Social information filtering: Algorithms for automating 'word of mouth'. *Proceedings of the Conference on Human Factors in Computing Systems (ACM CHI'95)* (pp. 210-217). Denver, CO.

Sherman, C. (2006, January 19). *Survey: Google, Yahoo still favorites in North America*. Retrieved May 29, 2006, from http://searchenginewatch.com/searchday/article.php/3578491

Smyth, B., Balfe, E., Boydell, O., Bradley, K., Briggs, P., Coyle, M., & Freyne, J. (2005, July 31-August 5). A live-user evaluation of collaborative Web search. *Proceedings of the 19th International Joint Conference on Artificial Intelligence (IJCAI-05)* (pp. 1419-1424). Edinburgh, Scotland.

Smyth, B., Balfe, E., Briggs, P., Coyle, M., & Freyne, J. (2003). Collaborative Web search. *Proceedings of the 18th International Joint Conference on Artificial Intelligence* (IJCAI-03) (pp. 1417-1419). San Francisco: Morgan Kaufmann.

Starr, B., Ackerman, M.S., & Pazzani, M. (1996). Do-I-Care: A collaborative Web agent. *Proceedings of the Human Factors in Computing Systems Conference (CHI '96)* (pp. 273-274). New York: ACM.

Sullivan, D. (2000, April 4). The vortals are coming! The vortals are coming! Retrieved May 24, 2006, from http://searchenginewatch.com/sereport/article.php/2162541

Sullivan, D. (2002, February 20). *Specialty search engines*. Retrieved March 23 and May 28, 2006, from http://searchenginewatch.com/links/article.php/2156351

Swearingen, K., Lyman, P., Varian, H.R., Charles, P., Good, N., Jordan, L.L., & Pal, J. (2003). *How much information?* Berkeley: University of California School of Information Management and Systems. Retrieved May 30, 2006, from http://www.sims.berkeley.edu/research/projects/how-much-info-2003/

Terveen, L., McMackin, J., Amento, B., & Hill, W. (2002). Specifying preferences based on user history. *Proceedings of the Conference on Human Factors in Computing Systems* (CHI'02). Minneapolis, MN.

Tkatchenko, M. (2005). *Combining reputation and collaborative filtering systems.* Retrieved September 1, 2006, from http://www.cs.ubc.ca/~kevinlb/teaching/cs532a%20-%202005/Class%20projects/Maria.pdf

Twidale, M., & Nichols, D. (1998). Designing interfaces to support collaboration in information retrieval. *Interacting with Computers, 10*(2), 177-193.

van de Wijngaert, L. (1999, August 13-15). A policy capturing study of media choice: The effect information [sic] of needs and user characteristics on media choice. In T.D. Wilson & D.K. Allen (Eds.), *Information Behavior: Proceedings of the 2nd International Conference on Research in Information Needs, Seeking and Use in Different Contexts* (pp. 463-478). Sheffield, UK: Taylor Graham.

Vance, D.M. (1997, August). Information, knowledge and wisdom: The epistemic hierarchy and computer-based information system. In B. Perkins & I. Vessey (Eds.), *Proceedings of the 3rd Americas Conference on Information Systems*, Indianapolis, IN.

VanFossen, L. (2005). *Top specialty search engines list.* Retrieved May 28, 2006, from http://www.cameraontheroad.com/index.php?p=196

Wang, J., Pouwelse, J., Lagendijk, R.L., & Reinders, M.J.T. (2006, April 23-27). Distributed collaborative filtering for peer-to-peer file sharing systems. *Proceedings of the 21st Annual ACM Symposium on Applied Computing (SAC'06),* Dijon, France.

Wasserman, S., & Faust, K. (1994). *Social network analysis: Methods and applications.* New York: Cambridge University Press.

Wenger, E. (1998). *Communities of practice: Learning, meaning and identity.* Cambridge: Cambridge University Press.

Wenger, E., McDermott, R.A., & Snyder, W. (2002). *Cultivating communities of practice: A guide to managing knowledge.* Boston: Harvard Business School Press.

Wu, X., Zhang, L., & Yu, Y. (2006, May 23-26). Exploring social annotations for the semantic Web. *Proceedings of the 15th International World Wide Web Conference* (WWW 2006), Edinburgh, Scotland.