

Towards a Unified Framework of IR Tasks and Strategies

Suresh K. Bhavnani

School of Information
University of Michigan
Ann Arbor, MI 48109-1092
Tel: +1-734-615-8281
bhavnani@umich.edu

Karen Drabenstott

School of Information
University of Michigan
Ann Arbor, MI 48109-1092
Tel: +1-734-763-3581
karen.drabenstott@umich.edu

Dragomir Radev

School of Information &
Department of Electrical Engineering and Computer Science
University of Michigan
Ann Arbor, MI 48109-1092
Tel: +1-734-615-5225
radev@umich.edu

1. Abstract

Despite huge advances in making information accessible to vast numbers of users, the effective retrieval of relevant information remains a challenge. User studies of online library catalogues, abstracting and indexing systems, and web search portals, repeatedly show that despite knowledge of basic search techniques, many users do not acquire strategies to find relevant information effectively.

The phenomenon of the ineffective use of retrieval systems despite experience, is not unique; numerous user studies of other complex computer systems such as word processors, spreadsheets, and CAD systems show that despite experience with basic tools, users do not progress to a more effective use of these systems. The effective use of computer systems is essential for improving the overall productivity of information workers and has been in the forefront of our research efforts. This research has led to a framework of effective and general strategies to use complex authoring systems. The framework has been used to develop a new training approach to teach strategic knowledge that has been tested in three universities, and has affected the learning of more than 150 students.

Unfortunately, such a unified framework of effective and general strategies is currently lacking in the field of Information Retrieval (IR). This paper presents our current work and a detailed proposal to build a framework of effective and general strategies based on the following approach: (1) develop a taxonomy of IR tasks based on tasks of real users; (2) develop a taxonomy of strategies based on general and effective IR strategies; (3) develop a descriptive model of expert performance; (4) develop a prescriptive model of effective performance; (5) apply the framework to design a curriculum of effective training of IR strategies. We have begun to perform the above research through a multidisciplinary effort of researchers from three critical domains: Human-Computer Interaction, Library Science and Information Retrieval. The research will consist of the analysis and categorization of reference questions from real users, experiments involving the observation and modeling of experts, and controlled experiments to evaluate the efficacy of teaching effective IR strategies to university students.

The proposed work has the potential of broad impact in three areas: (1) development of an explicit user-based approach to model information retrieval; (2) an approach to train users to be effective retrievers of information; (3) a systematic method to guide designers to identify and implement functionalities that enable users to execute effective IR strategies.

2. Introduction and motivation

Despite huge advances in making information accessible to vast numbers of users, the effective retrieval of relevant information remains a challenge. User studies of online library catalogs (Marchionini, 1989), abstracting and indexing systems (Kirby and Miller, 1986), and web search portals (Jansen et al., 1998), repeatedly show that despite knowledge of basic search techniques, many users do not acquire strategies to find relevant information effectively.

The phenomenon of the ineffective use of retrieval systems despite experience, is not unique; numerous user studies of other complex computer systems such as word processors, spreadsheets, and CAD systems show that despite experience with basic tools, many users do not progress to a more effective use of these systems. These studies include business students using spreadsheet applications (Nilsen et al., 1993), architects and mechanical engineers using CAD systems (Bhavnani and John, 1996; Lang et al., 1991), and secretaries using word processors (Rosson, 1983). Even computer science students with formal training and many years of experience do not use UNIX commands effectively (Doane et al., 1990). As stated by the National Research Council report on Information Technology Literacy: “Many who currently use information technology have only a limited understanding of the tools they use and a (probably correct) belief that they are underutilizing them.” (p.1, Committee on Information Literacy, 1999).

2.1. The ineffective use of authoring systems

To understand why complex computer systems are difficult to use, our prior research has investigated the use of authoring systems such as CAD, word processors, and spreadsheets. We used ethnographic techniques to observe experienced CAD users in their professional environment (Bhavnani et al., 1996). We then used cognitive modeling techniques to analyze our real-world observations, uncovering why the users performed as they did and how they might be helped to use the tools more effectively (Bhavnani and John 1996, 1997, 1998). While searching the human-computer interaction (HCI) literature, we discovered that our observations and analyses were confirmed by empirical studies in many computer domains besides CAD. Our analyses have identified a specific type of knowledge, *strategic knowledge*, which is missing from many real-world users’ repertoires and from the books and curricula that teach complex systems. We have also shown that neither redesign of complex systems, nor experience with the system, can guarantee the acquisition of this important type of knowledge (Bhavnani, 1998, Bhavnani and John, 2000).

We then identified a framework of general and effective strategies to use complex computer applications (Bhavnani, 2000) and used that framework to design a new training approach to impart strategic knowledge to users (Bhavnani et al., in press). This approach has been empirically tested at Carnegie Mellon University (Bhavnani et al., in press), the University of Western Australia (Thomas and Foster, in press), and a third experiment is currently underway at the University of Michigan. Our research has therefore directly affected more than 150 students in three universities.

Based on the above results, we now have a better understanding of how to address the ineffective use of complex authoring tools that is so common in the world. Because the phenomenon of ineffective use by large numbers of users also plagues the critical domain of Information Retrieval (IR), this paper describes how we aim to use our experience in improving the ineffective use of authoring systems, to address the ineffective use of IR systems.

2.2. The ineffective use of IR systems

The explosion of knowledge in the last century, combined with huge advances in making that knowledge available through the Internet, has transformed our meaning of “knowing.” As stated by Herbert Simon, the meaning of “knowing” has shifted from being able to remember and repeat information, to being able to find and use it (Simon, 1996). This makes information retrieval central to our functioning as an informed society, and retrieving information effectively is increasingly becoming a critical skill in doing everyday tasks.

However, despite huge advances in content and access, the effective retrieval of relevant information for most users remains a challenge. User studies of information retrieval systems such as online library catalogs (Marchionini,

1989), abstracting and indexing systems (Kirby and Miller, 1986), and web search portals (Jansen et al., 1998), repeatedly show that despite knowledge of basic search techniques, many users do not acquire strategies to find relevant information effectively.

For example, an extensive study of Excite user queries (Jansen et al., 1998), found that only 5% of the queries involved searches using the feature MORE LIKE THIS. This technique, of starting from a good source and finding similar sources is known as the *Citation Pearl Growing* strategy. This strategy is frequently used by expert librarians to quickly retrieve relevant information (Drabenstott, 2000).

Our own investigation has confirmed the difficulty end users appear to have when retrieving relevant information. The Internet Public Library (IPL), a well-known digital reference service developed by researchers at the University of Michigan, has been collecting data of user activity since 1995 (Carter and Janes, in press). In a random selection of 100 emails from the IPL database, we found 50% of the requests had explicit statements describing a failed attempt at either searching IPL, or the use of general purpose search portals. The following demonstrates a typical example of an email reference question sent to IPL:

Question: My question is about poetry from the sixteenth century. I think that it was a poem by a guy named Bishop, but I'm not sure. I heard this poem or saying at the end of the movie "Forces of Nature," the movie with Sandra Bullock and Ben Affleck. Anyway Ben is the one that recites it at the end and I really liked it. I wanted to find out who wrote it and where could I find it. If you could help in any way I would appreciate it.

Reason: I'm into poetry and this one has managed to stump me!

Sources consulted: I have looked in the poetry sections and I have tried under the sixteenth century!

The above example suggests that despite knowing of the existence of various search sites such as IPL, and functions of how to use them, the user was unable to achieve the goal of finding the poem. Given the fuzzy information the user has, such a search can be problematic. Searching on the very common name "Bishop" leads to a large number of hits in Google, nothing in the IPL poetry section (as experienced by the user). Furthermore, searching for the movie and actors also leads to a huge and irrelevant list of hits. The user is therefore "stumped".

A more effective strategy (as used by the IPL expert who responded to the email question) is searching only within a specialized collection. This requires three steps: (1) Find a collection that specializes in movies through the subject index of any search portal, or by recalling from memory the name of the Internet Movie Database (<http://us.imdb.com/>); (2) perform a targeted search just within the subtopic of movie quotations using the key word "Bishop;" (3) find the movie in the recalled list by scrolling or doing a string search on the movie name. This overall strategy of searching within a specialized collection, instead of in a general catalog, is a strategy well known by reference librarians.

The overall conclusion of user studies on IR systems is that "while users seem to master the command language with no difficulties, they lack the expertise needed for formulating search strategies" (Fidel, 1991). We believe that experts can find information more effectively than most users because they have acquired effective strategies that go beyond knowing how to entering a keyword in a search engine. Strategic knowledge allows experts to organize their search effectively to quickly and accurately retrieve relevant information. The systematic identification and teaching of strategic knowledge therefore becomes central to making users more effective in information retrieval.

3. The need for a unified framework of IR tasks and strategies

In order to identify and teach effective strategies, we need to understand what users are searching for, the kind of strategies experts use, and how the strategies are related to specific tasks. This requires a unified framework of users' tasks, effective strategies, and a descriptive and prescriptive model of effective search performance. However, as discussed below such a unified framework is currently absent in the IR field.

3.1. Taxonomy of IR tasks

Because the central goal of this proposal is to improve how users retrieve information effectively, it is critical to understand the nature of search tasks that users perform. However, the few attempts at defining a taxonomy of IR tasks are either sparse, not empirically based, or not coupled to strategy selection.

Basic texts for educating reference librarians divide tasks into *factual* questions and *source* questions. Factual questions have a definite answer, and, in some cases, a particular source has been published to answer them. For example, most reference librarians consult *Statistical Abstracts* to answer questions such as the average height of US males. In contrast, source questions are broader and require many more pieces of information to address the question. Librarians might consult several sources from which they construct an answer, or from which they form an opinion about a subject. Alternatively, they might suggest these sources to users who can study them in order to arrive at the “answer” on their own. Such source questions include: What are the Republican law maker views on trade with China? What would be the effect of legalizing drugs on the student population? Carter and Janes (in press) analyzed three thousand questions submitted to the Internet Public Library and categorized them as *factual* or *source* questions.

In the course of promoting a seven-step strategy to web searchers, Pffafenberger (1996) divided tasks based on the amount of information needed for a topic into three types of questions: (1) *Finding specific information* (similar to a factual question), (2) *Collecting a few sources* of high quality information, and (3) *Collecting everything on a topic* that is, conducting an exhaustive search to retrieve all available material. Spool (1999) defined four types of questions: (1) *Simple fact questions*, the simplest type of question for which there is only one correct answer (similar to factual questions), (2) *Judgment questions*, where the user must not only locate potential answers from sources, but also analyze them to formulate an opinion based on collected information (similar to source questions), (3) *Comparison of fact questions*, where the user researches two or more questions in order to arrive at an answer, and (4) *Comparison of judgment questions*, situations involving comparisons and judgments.

Furnas (personal communication) conducted interviews with a wide range of professionals in order to analyze information retrieval tasks in the context of larger tasks such as writing a paper, or making a medical diagnosis. He categorized tasks in terms of user traits: how much users care about the outcome, how often the task is repeated, the degree of expertise of the user, and how focused the user’s goal was.

While a review of the published literature has revealed several classifications of tasks, these attempts have several deficiencies. The taxonomies of Pffafenberger and Spool are an important first step but there has not been an empirical confirmation of either their use or their properties. For example fact questions vary along other dimensions such as how much the user knows already about that fact; fuzzy information about a fact question requires a different strategy compared to more accurate information. Furnas’s taxonomy is empirically grounded but is too abstract to allow for strategy selection. For example how does one operationalize the degree of expertise of a user with respect to selecting a strategy? Therefore none of the taxonomies that we have been able to identify is sufficient for our goal of strategy selection.

3.2. Taxonomy of IR strategies

Research in understanding the interactions between librarians and library users can be characterized by at least two levels. At one level, there has been the development of overall models of interactions between librarians and library users such as the seminal work done by Taylor (1968), Kuhlthau (1991), Belkin (1980), and Dervin (1983). At another level, research has focused on a single aspect of the overall search process, specifically the strategies people use to interact with an information retrieval system. The latter research has produced many different categories of strategies.

Markey and Atherton (1978) recommended five search strategies to intermediary searchers, which were published in their Ontap manual for ERIC database searchers: (1) Building Block, (2) Most Specific Facet First, (3) Highest Postings Facet First, (4) Successive Fractions (or Big Bite), and (5) Citation Pearl Growing. They advised intermediary searchers to choose a strategy based on a facet analysis of the concepts that characterized users’ questions. Harter (1986) added Pairwise Facets and Citation Indexing Strategies to the Ontap strategies. To our knowledge, no empirical research has been done to determine whether expert searchers use the Ontap strategies; however, most textbooks cover the Ontap search strategies, and, presumably, intermediary searchers use such strategies in view of the formal training they receive in LIS schools (Meadow and Atherton, 1981; Harter, 1986; Borgman, 1984; Walker and Janes, 1999).

Bates (1979) introduced the notion of search tactics—a move made to further a search. Bates (1978) based search tactics on empirical data and designed them to be of assistance to the searcher while in the process of searching. Fidel (1985) recommended moves to alter sets that were too large, too small, off target, or to increase precision and recall. When libraries made online catalogs available to library users, Bates (1985) introduced the “Berry picking” model of information retrieval which demonstrated (1) that users enlisted many different approaches to satisfying an information need, and (2) that their needs changed and evolved in the course of searching and reading retrieved documents. The Berry picking model included strategies that searchers commonly use such as chasing footnotes, citation searching, and scanning journals’ tables of contents.

When web searching became available in the mid 1990s, search engines were not much different from online catalogs because they too responded to the individual queries that searchers entered. Pfaffenberger (1996) advised web searchers to choose amongst three strategies based on the amount of material they wanted to retrieve: (1) the *locating strategy* advised web searchers to use a search engine to find specific types of information, (2) the *sampling strategy* advised web searchers to choose a web directory or subject tree to find a few sources of high-quality information, and (3) the *deep searching strategy* advised web searchers to use search engines’ Boolean operators, truncation, and proximity operators to perform a highly effective search. Drabenstott (2000), defined seven new strategies for searching the web. Her objective was to make strategy recommendations that would work across all search engines and browsing directories, but would not require end users to perform Boolean searching. This was based on her observations that it was difficult for a layperson to learn how to use Boolean operators and related techniques, e.g., truncation, proximity operators, sets, delimiters.

While the above attempts at identifying search strategies are useful in themselves, no unified framework has emerged. The strategies are typically specific to a particular retrieval system such as web search engines or online catalogs. Although some strategies are new and untested, they remain abstract and difficult to impart to users and even more difficult to remember. The objectives of this proposal are to overcome these limitations by unifying the known strategies into a framework that facilitates the selection of specific search strategies to accomplish specific tasks.

3.3. Use of IR strategies

Several studies have analyzed the details of search behavior of expert intermediaries (such as librarians), and of end users. Early studies focused on the informal observation of expert searchers, and of novices learning to be intermediaries in libraries (Markey and Atherton, 1978, Belkin, 1980). When transaction logs became available, researchers were able to get a more detailed understanding of search behavior. These logs of end users interacting with online library catalogs revealed that users performed many more subject searches compared to title and author searches. However, a large percentage of these subject searches failed because they did not match the catalog’s controlled vocabulary even though the vocabulary was available to the users. Furthermore, users repeatedly did not use advanced features to perform searches effectively (Peters, 1993, Drabenstott, 1991, Kirby and Miller, 1986). The general conclusion is that while users mastered some basic commands to conduct searches, they did not advance to a more effective use of IR systems.

The above conclusion is not unique to users of library catalogs. Several studies of web search have shown that advanced features are typically ignored (Jansen et al., 1998, Abdulla et al., 1998, Spink and Saravecic, 1997). For example despite the availability of the MORE LIKE THIS feature in Excite (which automatically conducts a search using the Citation Pearl Growing strategy), hardly 5% of web searchers used it. In contrast, 11% of the intermediaries used this advanced feature (Spink and Saravecic, 1997).

While the above studies begin to reveal details about user search behavior, they are limited by the nature of the analyzed data. The transaction logs, and queries used in these studies, are snapshots of user behavior and do not capture the richness of interaction over an entire search session. Several studies have attempted to overcome this limitation by analyzing search behavior of users and intermediaries over an entire search session. Fidel (1985) analyzed seven expert online searchers while they performed ninety searches. She found that expert searchers changed their queries in two ways: (1) *Operational moves* reformulated queries without changing the meaning of the sets they retrieved, and (2) *Conceptual moves* reformulated queries by modifying the meaning of the concepts the retrieved sets represented. Experts used these moves when the retrieved sets were too large, too small, or when

the set was off target. Fidel concluded that expert searchers “fine-tune their decisions about how to proceed with a specific search strategy by determining which type of move, operational or conceptual best serves their purposes” (p. 73). In contrast to studying moves within a strategy, Xie (1997) observed that users *shifted* strategies based on several variables ranging from the user’s immediate and long-term goals, to environmental and situational factors.

The above studies have therefore begun to reveal the widespread ineffective use of IR systems, in addition to revealing the complex behaviors involved in search. A framework of strategies must therefore model how experts shift and move within strategies, in a dynamic interaction with the results of a search. However, these descriptive studies of strategy phenomena are only a first step in our larger goal of making such knowledge available to users. To achieve our larger goal, the strategic knowledge known by experts needs to be operationalized before it can be used in a prescriptive way to improve the search behaviors of end users. This, we believe, requires a unified framework of IR tasks and strategies.

3.4. Towards a unified framework of IR tasks and strategies

Although several researchers have proposed categories of IR tasks, strategies, and studied how strategies are used to perform tasks, no unified framework has emerged. Furthermore, as discussed earlier, there has not been an effort to tightly connect an empirically grounded taxonomy of tasks, to a taxonomy of general strategies. This explication, we believe is the basis of the unified framework.

The lack of a unified framework affects several types of professionals who could directly affect end user behavior. These include: (1) researchers who could systematically identify and test new strategies, (2) trainers who could impart a cohesive knowledge structure to users, and (3) designers to make sure their systems support the use of effective strategies. Furthermore, the above strategies have been identified in the absence of a taxonomy of IR tasks which are central to understanding why and how the strategies should be employed.

Based on our previous success at identifying a framework for effective and general strategies for authoring tools, we believe a similar effort is critical to develop a framework for effective and general IR strategies. This framework should have the following characteristics:

1. The development of a taxonomy of IR tasks at the right level of granularity to enable a tight coupling with general strategies.
2. The development of a taxonomy and rationale of effective IR strategies that are general across the three major information retrieval systems: Online library catalogs, abstracting and indexing systems, and web search portals.
3. A descriptive model of expert performance which reveals how experts select, sequence (related to the phenomenon of shifts), and modulate strategies (related to the phenomenon of moves) for real-world tasks described in the taxonomy of IR tasks.
4. A prescriptive model of effective search strategies that operationalizes expert behaviors and makes explicit the utility conditions for each strategy. These utility conditions will describe when each strategy is useful, how they should be sequenced, and how they can be modulated.

The explication of the taxonomy of tasks and strategies, and the utility conditions for each strategy are expected to have direct applications to training and design. Our proposed work will elaborate how we intend to develop the framework and apply it to training and to design.

4. Current and proposed work

Our current and proposed research consists of five parts: (1) The development of a taxonomy of IR tasks based on the IR tasks of real users. (2) The development of a taxonomy of IR strategies that are general across different IR systems. (3) The development of a descriptive model of expert performance. (4) The development of a prescriptive model of effective performance. (5) The design and evaluation of a pedagogical approach to teach the effective use of IR systems.

4.1. The development of a taxonomy of IR tasks

Our approach to develop a taxonomy of IR tasks is grounded in the real-world tasks of users. The goal is to develop a taxonomy of IR tasks that lends itself specifically to the selection of strategies. Towards that goal, we will use the following steps to analyze and categorize emails stored in the IPL database:

1. Develop a coding scheme to categorize IR tasks. We will analyze 100 randomly selected emails from the IPL database and develop a coding scheme to categorize them based on their task characteristics. For example, our preliminary analysis of 100 emails has enabled us to extend the taxonomies developed by Pffafenberger and Spool which were based only on the nature of the question. We believe that a taxonomy of IR tasks must contain at least two dimensions: (1) What the user knows about the information being sought, and (2) what the user requires to do with that information. We have found that what the user knows about the search tasks can vary between fuzzy and accurate information. What the user requires can vary between a pinpointed factual answer, a sampling of documents in the field, or an in-depth or exhaustive collection of information. We will develop this initial understanding into a more coherent coding scheme.

2. Categorize real-world IR tasks based on the coding scheme. Fully realizing that the process of categorization is an iterative process, we will begin to categorize the emails based on our initial coding scheme. Figure 1 shows an initial attempt to categorize the emails to enable the selection of strategies. However, in our past experience, such initial attempts tend to be far too simplistic. For example, we have realized that the information that is known by the user has many more complex characteristics such as whether the accurate information

What the user requires from the search

	Factual	Sample	In-depth or Exhaustive
What the user knows	Fuzzy or incomplete information My question is about poetry from the sixteenth century. I think that it was a poem by a guy named Bishop, but I'm not sure. I heard this poem or saying at the end of the movie "Forces of Nature,"...	I am interested in finding out the history of the town of Pomeroy, Washington. Any newspapers, local records, etc. would be helpful.	I am trying to gather information on the cave dwellings of early american indians in the southwest, i believe the tribe was anasazzi, but not sure.
	Accurate or precise information I am looking for the amount of gold that is in Fort Knox as well as the amount of gold geologist think is still in the earth.	I need a sound file of the lord's prayer spoken in Danish.	Names of tests used in Washington State to place a child in a higher grade, in elementary school, than his age dictates. i.e Kindergarten aged child into Grade 1.

Figure 1. A taxonomy of tasks derived from a preliminary analysis of 100 email requests randomly selected from the IPL database. Each cell of the table contains an example of a categorized search task.

provided by the user is unique. The uniqueness of the terms in the question plays a dominant role in how a query is constructed. To understand these subtle variations within dimensions, we will analyze how the IPL expert solved the search task for the user. For example the following response by the IPL expert reveals that the expert focused on the uniqueness of the term *Anasazzi* (which was misspelled by the user), and that such information is most probably retrieved from an encyclopedia:

Thank you for your question about the Native American cave dwellings in Arizona. I had heard of an ancient American tribe called the Anazassi, but had no luck finding information on the Internet. Playing around with the spelling of the word on the online version of Encyclopedia Britannica at <http://www.britannica.com> turned up two articles on your topic:...

We believe our approach to analyze the rich data recorded in the IPL database will exploit such subtle but important variations typical in real-world tasks.

3. Verify if coding scheme is meaningful to IR experts. Our first attempt at categorizing the emails will be discussed through open-ended interviews with four information retrieval experts to determine if the categorization scheme is meaningful to select a strategy. Random samples from each category will be selected and each of the library experts will be asked to discuss how they would address the task. Particular emphasis will be placed to see if strategies differ within the same category, prompting a rethinking of the categorization scheme. This part of the

study is formative in nature and is deliberately designed to be open-ended to encourage a free flow of ideas. This process will be repeated with each of four experts until the categories begin to predict the kind of strategies that an expert would use.

4. Perform an inter-coder reliability experiment with IR experts. Having categorized each of the emails, we will develop definitions for each of the categories and then conduct an inter-rater reliability experiment with four different information retrieval experts. These experts will be trained in the categorization scheme on a sample set, and then asked to categorize another set of randomly selected email questions. The goal of this experiment is to determine if the experts reliably agree on the categorization scheme.

4.2. The development of a taxonomy of general IR strategies

Our research on the taxonomy of IR strategies will focus on the identification of *general* strategies. By general we mean that the strategies must be effective across the three main IR systems: Online library catalogs, abstracting and indexing systems (consisting of a database such as INSPEC with an interface), and web search portals. The focus on general strategies is critical because of our larger goal to help users develop knowledge that is *flexible* in that it can transfer across systems, and be *relevant* despite rapid changes in interfaces and features offered in systems. The value of identifying and teaching such general strategies has been demonstrated in our previous research (Bhavnani et al., in press). The following steps are designed to achieve that goal:

1. Consolidate known IR strategies. As described in section 2.2, the identification of strategies over the last two decades has been ad hoc and there is a need for a framework that provides a clear description and rationale for the strategies. We will collect all the known IR strategies described in the literature and identify a set that is generally useful across online library catalogs, abstracting and indexing systems, and web search portals.

Figure 2 shows an example of how we have attempted to categorize some general strategies reported in the literature (Bhavnani et al., in review). As shown in the figure, our preliminary analysis of known IR strategies suggests that many can be grouped based on their goal to exploit *regularities* in information spaces. Regularities include the tendency of authors to cite other authors that do related work, and could therefore be of high relevance. Such regularities are *emergent* because they are based on the behavior of information authors. For example, Strategy-1 exploits such emergent regularities by reminding users to search for relevant information by accessing works cited by a relevant author. This strategy can be used by following footnotes in books retrieved through an online library cataloging system, through reading the reference list of papers acquired through full text retrieval in abstracting and indexing systems, or by using the FIND MORE OF THE SAME feature (a rudimentary form of the strategy) in a web search portal. Another form of regularity in an information space is *imposed* to create collections or consistency in the storing of information sources. For example, Strategy-5 exploits the imposed regularities of any indexed collection such as the INSPEC database containing mainly papers in engineering. The strategy calls for the use of simple keywords in successive queries to narrow down a retrieved set, an approach also known as the *Successive Fractions Strategy* (Drabenstott, 2000).

Figure 2 therefore (1) consolidates several known IR strategies, (2) provides the rationale underlying classes of strategies, and (3) demonstrates their general nature. We will extend this initial framework by consolidating other strategies described in the literature, and describe their rationale and general nature.

2. Experiment to identify the strategies experts use while performing real-world IR tasks. While the emerging framework of general IR tasks described in Figure 2 attempts to consolidate strategies, they are based on descriptions of strategies in the literature that are not tied to a systematic analysis of their actual use. Therefore, we need to analyze if and how these strategies are actually used in the context of real tasks. Furthermore, we discovered that most strategies discussed in the literature have focused on the construction of the search query itself. However, there appears to be other strategies that attempt to enhance the process of retrieval. For example, while doing exhaustive searches, an effective strategy is to organize results of a search in named hierarchies to enable easy access at a future time. This strategy is useful regardless of the IR system one uses. Therefore there may be other effective and general strategies that have not been elaborated in the literature.

To determine if the known strategies are actually used by experts, and if there exist other strategies that have not been discussed, we will do a protocol study of 15 IR experts while they perform tasks. We will observe 15 librarians on the University of Michigan campus as they perform search tasks using three types of systems: Web search portals (e.g. Yahoo, Hotbot), an online library catalog (e.g. Mirlyn), and an abstracting and indexing system (e.g. INSPEC). The tasks for the web search portals will be randomly selected from different categories in the taxonomy developed earlier. The tasks for the online library catalog and abstracting and indexing system will be selected from Mirlyn transaction logs. The librarians will be asked to talk aloud as they perform their tasks and their verbal protocols and interactions will be recorded.

General Strategies	Online Library Catalogs	Abstracting & Indexing Systems	Web Search Portals
Emergent regularities			
1. Identify other authors through references	Retrieve paper and identify other authors	Retrieve paper and identify other authors	Use FIND MORE OF THE SAME to identify other sources
2. Identify works that cite current work	Access citation index to identify authors that cite a book	Access citation index to identify authors that cite a journal paper	Use link indexing to identify sites that point to current site
3. Identify other works of same author	Search by author name of known book	Search by author name of known paper	Search by organization name of known site
Imposed regularities			
4. Identify relevant collections	Select CMU library catalog when looking for books on robotics	Select INSPEC when looking for papers in Engineering	Use ditto.com when searching for graphic images
5. Search a collection using indexing scheme	Search within subject headings	Search within index	Search subject trees
6. Search a collection using keywords on successive sets of recalled hits	Search database using single keyword, then search within that result	Search database using single keyword, then search within that result	Search database using single keyword, then use "search within results" to narrow down search
7. Search multiple collections simultaneously	Use online computer library center (OCLC) to locate book in any university	Use DIALOG One Search to search multiple databases	Search using meta search engines
8. Guess name based on fields or naming conventions	Locate book by typing: guessed name "for dummies"	Locate journal by typing guessed field, "Journal"	Locate web site by typing: "www.", guessed name, ".com"

Figure 2. General and effective IR search strategies, and examples of their use across different information retrieval systems.

The verbal protocols of the librarians will be transcribed and analyzed. The analysis will focus on the identification of general strategies in addition to how they interact with domain knowledge that the librarians use to identify and search different sources. We will iterate with the taxonomy of tasks developed in the previous section to determine if taxonomies of tasks and taxonomy of strategies have correspondence.

4.3. The development of a descriptive model of expert performance

While the taxonomy of tasks and strategies provide an empirically-based starting point for our understanding of what users need, and how experts respond to those needs, there needs to be an explication of how experts actually connect the two. Our analysis of strategies has revealed an important difference between IR strategies, and strategies to use authoring tools. Strategies to use authoring tools, such as CAD systems, guarantee that the goal is achieved if the strategy is executed correctly and under normal conditions of application use (e.g. if you set up dependencies correctly, they will be automatically updated). In contrast, strategies of IR are probabilistic in nature and do not guarantee retrieving relevant information. This makes IR strategies heuristic in nature. For example, the strategy of selecting a search engine and restricting the search to sound files does not guarantee that the sound file for the Lord's Prayer is retrieved. For example, a search of Yahoo using this strategy does not retrieve the file, whereas the same search in Altavista does.

Because IR strategies are heuristic in nature, experts select strategies and then abandon them when they do not achieve the goal. Experts therefore appear to know how to *sequence* strategies based on the nature of the retrieved results. Furthermore, even within a strategy, experts *modulate* their strategies by changing their search strings to be more precise or more specific depending on the nature of the results. The sequencing and modulation of

strategies are based on a dynamic understanding of the items being retrieved. To rigorously understand these phenomena, we will model and test interaction as described below.

Analyze interaction through Problem Behavior Graphs. We will model the on-line behavior of the 15 experts described earlier to reveal how they select, sequence, and modulate strategies. This modeling will be based on the pioneering work of Newell and Simon, (1972) in the area of human problem solving through the use of Problem Behavior Graphs (PBG). These graphs helped make salient the search behavior of problem solvers. Figure 3 shows a portion of the problem behavior graph of an expert attempting to find the poem by Bishop on the web. The boxes represent knowledge states, and the arrows represent operators that take the user from one knowledge state to another. The graph is read from left to right and from top to down.

Figure 4 shows the PBG of a novice doing the same task. The vertical lines represent when a search path is abandoned and when either a new strategy is attempted, or when a strategy is modulated to control the precision and recall of the retrieved items. As shown, the PBG can reveal many aspects of a user's search behavior. Long thin PBGs (such as in Figure 3) leading to success expresses expert behavior where selection of the right strategy did not require sequencing or modulation. On the other hand, long thin graphs leading to failure and abandoning the search process, could reveal the lack of sequencing and modulation strategies in novice behavior. Similarly, bushy graphs (such as in Figure 4) may reveal the difficulty a user is having in finding relevant information.

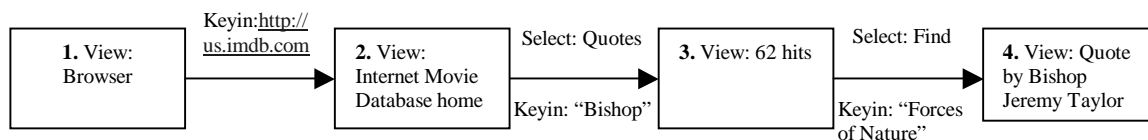


Figure 3. Problem Behavior Graph of an expert finding the poem by Bishop in the movie “Forces of Nature”

The PBGs can also reveal the strategies being used. For example, the PBG of the expert in Figure 3 shows the expert going directly to a collection, an important strategy for tasks where the user has precise information and requires a factual pinpointed answer. On the other hand, the PBG of the novice in Figure 4 who does not know this strategy flounders searching for the same information by directly typing the search string into a web search portal. Such kinds of rigorous modeling have been rare in IR research, but is necessary to do if we are to understand the intricacies of search behavior.

4.4. The development and evaluation of a prescriptive model of effective performance

The observations and modeling of expert behavior will provide us an understanding of the subtleties involved in selecting, sequencing, and modulating IR strategies. We will consolidate these observations into a parsimonious set of general strategies and utility conditions, and evaluate them based on the following steps:

1. Operationalize behaviors observed in expert behavior. Although the behaviors observed in the PBGs provide a rich description of expert behavior, they need to be abstracted and operationalized into a set of utility conditions to use the strategies. Research on the use of strategies emphasize the importance of making explicit the conditions of use before strategies can be understood and retrieved in appropriate contexts (Singley and Anderson, 1989). Towards that goal we will abstract sequences of knowledge states and operators from the PBGs described in the last section into prescriptive heuristics. For example, these heuristics will specify how to select strategies based on characteristics of tasks and intentions, when to switch strategies based on the results of a search, and how to modulate a strategy to increase precision or recall. The heuristic rules will also include aspects that go beyond the

query formulation to ways to enhance the retrieval process such as organizing the hits in hierarchies for future use.

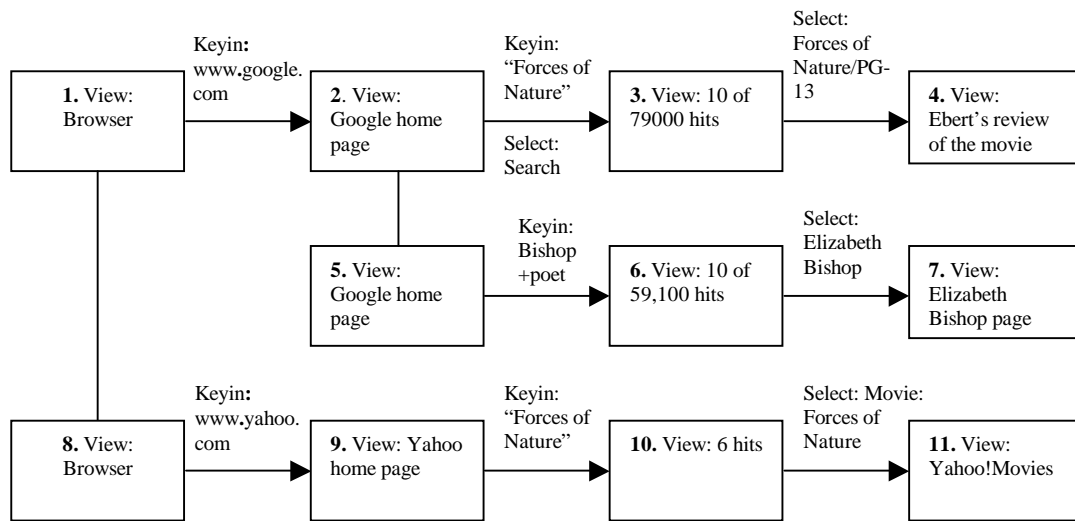


Figure 4. The Problem Behavior Graph of a novice attempting to find the poem by Bishop in the movie “Forces of Nature”

2. Experiment to test the heuristic rules of strategy selection, sequencing, and modulation. Having operationalized the heuristics that connect our taxonomy of tasks to the taxonomy of strategies, (based on how experts use to select, sequence, and modulate strategies) we will then test whether the framework can lead end users to find relevant information. In this experiment we will give users a new set of randomly selected tasks along with strategies generated from the framework to perform those tasks. Fifteen users will be asked to use the strategies in the provided sequence to complete the retrieval task. This experiment will demonstrate: (1) that the rules generalize to tasks other than those from which they were derived, (2) that users can implement these strategies, and (3) missing pieces of knowledge when users fail to complete the tasks despite the use of strategies generated from the framework. This will make explicit aspects of effective behavior that could have been missed in the earlier research steps. Given the heuristic nature of IR strategies, and the dynamic nature of the knowledge states during IR, we are under no illusions that the explication of search expertise is a simple “rule-oriented” task. However, we believe the fine-grained analysis we advocate is necessary to pinpoint exactly where the complexity lies.

4.5. Design and evaluation of an approach to teach effective IR strategies

The above three sections described the development and testing of a unified framework for IR strategies. However our principal goal is to improve how end users search for information using IR systems. Based on our experience in teaching general and effective strategies for using authoring tools (Bhavnani et al., in press), we will use the following steps to teach general IR strategies:

1. Develop a curriculum to teach general IR strategies. Our prior research on the ineffective use of complex computer systems led to the development of a course that teaches general strategic knowledge in combination with specific commands to use applications such as word processors, spreadsheets, and web authoring systems (Bhavnani et al., in press). Our approach has been tested out in three universities, the most recent of which is the *Computing Skills and Concepts* course offered by the School of Information (Course# SI-101) at the University of Michigan. Starting in January 2001, the first offering of this course was limited to students from the Schools of Art and Design, and Music. These schools tend to have low levels of computer skills, and have classes that typically have more women than men. Our current experiment tests if the approach is effective to teach students with low levels of computer skills how to use complex computer systems effectively in a short period of time.

We will add four new lectures to SI-101 that will teach effective strategies for information retrieval. The focus of these new classes will be to teach art students general and effective strategies to retrieve information from library and Internet sources. The goal is to enable art students to create their own web portfolios by using resources on the Internet, and through the extensive library system at the University of Michigan. We will use our explication of the prescriptive model of effective strategies described in the last section to design our course. The design will be based on four types of knowledge that is the core of strategic training:

- (1) Students must learn about the existence of specific strategies to execute frequent tasks. In cognitive terms this is the declarative knowledge related to the existence of effective strategies. We make this knowledge clear by enunciating effective strategies in their general form through a printed handout each student uses throughout the course.
- (2) Students must learn when to use a particular strategy. In cognitive terms this translates into a selection rule that picks an appropriate strategy. We impart this knowledge by presenting students with tasks where there are multiple ways to perform the task. The students will be asked to describe orally what is the most effective method to complete that task and why they picked one way above others. This engagement in problem solving will assist students to understand the appropriate contexts for using strategies.
- (3) Students must know how to execute a strategy. In cognitive terms this means users must learn the procedural knowledge to sequence commands in the strategy, and to sequence operators to execute commands. We will impart this knowledge through demonstration and practice of executing commands and strategies through motivating examples.
- (4) Students must learn to use the strategies across applications. In cognitive terms the knowledge must be learned at a general level to facilitate transfer. We will impart this knowledge by showing explicitly in the handout how the same strategy can be used in different applications.

The above four types of knowledge will be incorporated into the design to teach effective and general strategies for information retrieval. This attempt at using our pedagogical approach originally developed for authoring tools, to the domain of information retrieval will demonstrate the generality of our framework and its usefulness to an underrepresented group.

2. Experiment to evaluate the efficacy of teaching general and effective IR strategies. We will test the efficacy of the new course content in a controlled experiment. Two sections of 22 students each will be taught general strategies and commands to use IR systems by instructors who will be specially trained to teach strategic content. These students will be compared to another group who will be taught only commands to use IR systems by another set of specially trained instructors. Both sets of instructors will follow a script that uses the same commands and examples, but differ in explicit strategic instruction.

The two groups will have an equal number of students and will be balanced by prior experience, major and year in college. The midterm exam will serve as the post-test. The post-test tasks will provide opportunities to use specific strategies, and will test if there is any difference between the two groups in terms of strategy use. The goal of the experiment is to understand which IR strategies require explicit training, and which ones are automatically acquired just by using commands. The experiment will pinpoint those strategies that are difficult to acquire by end users and guide our research towards improving the curriculum design. Through post-hoc evaluation we will analyze if there are any differences in the performance between males and females in the two groups. Based on prior research on gender differences (Margolis and Fisher, 1997), we believe the strategy instruction will have a positive effect on how women acquire knowledge of effective IR strategies.

4.6. Expected impact

Our research should have impact in theory development, education, and the design of IR systems as described below:

1. Development of a unified framework of IR tasks and strategies. The taxonomy of tasks empirically derived from real users' tasks, the taxonomy of general strategies empirically derived from experts who perform the user

tasks and from the existing literature, and the explicit connection between them based on a detailed analysis of problem behavior graphs, will provide a unified framework of IR strategies. This framework of empirically grounded effective and general strategies will be a contribution to the theoretical development of IR research and will be a direct response to Saracevic's call "to concentrate more of research in IR on interactions to more resemble what is actually going on in practice." (Saracevic, 1996).

2. Education of users to be more effective in the use of IR systems. The application of the framework to education will directly benefit real users. This has been a central goal of all of our past research, and will continue to be a focus in the proposed research. Furthermore, our experience in understanding how to teach the Fine Arts student population (which has traditionally had a high percentage of women) the sophisticated use of technology has the potential of widespread application as we have already demonstrated by our past research in authoring applications. The approach of teaching general strategies in combination with specific commands is demonstrably an important method to teach application knowledge that is retained and can be transferred across applications. Our commitment to empirically evaluate each step we take responds to Saracevic's second call "to involve interaction in a major way in test and evaluation of IR systems." (Saracevic, 1996).

3. Consistency in the design of IR systems. We believe that a taxonomy of strategies as described in Figure 2 could help in making IR systems more consistent in the functionality they offer. For example, even our first attempt at identifying general strategies enabled us to systematically analyze whether 6 web search portals (Yahoo, Lycos, Infoseek, HotBot, Excite, and AltaVista) supported the 9 strategies described (Bhavnani et al., in review). We found that although the strategies were well known, they were not consistently supported. For example, neither Yahoo nor Excite provide a way to identify web sites that link to a page of interest (Strategy-2). These general strategies can therefore help designers systematically identify and make salient missing functionality to include on their web sites. This step responds to Saracevic's third call "to apply whatever was found in interactive research to the design and development of improvements in IR interfaces and processes." (Saracevic, 1996).

4.7. Conclusion

This paper has described current and proposed research aimed towards developing a unified framework of task and strategies based on the analysis of real-world tasks, observation and modeling of IR experts, and controlled experiments. We believe this effort will be successful because it involves the interdisciplinary effort of faculty members from the School of Information trained in Human-Computer Interaction (Bhavnani), Library and Information Science (Drabenstott), and Computer Science (Radev). Our team will use prior experience in the development of computational cognitive models and controlled experiments (Bhavnani and John, 1996, 1997, 1998, 2000), the identification of effective strategies for information retrieval (Drabenstott, 2000) and the development and testing of advanced systems of retrieval such as question-answering systems, and document summarization (Radev et al., 2000, Prager et al., 2000). This multidisciplinary research should have direct impact on the design of training focused on making users more effective in use of IR systems, and on the design of IR systems to ensure that they are conducive to the use of effective strategies.

4.8. References Cited

- Abdulla, G., Liu, B., & Fox, E. (1998). Searching the World-Wide Web: implications from studying different user behavior. <<http://video.cs.vt.edu:90/~abdulla/webnet98/query2.html>> (1999 February 6).
- Bates, M. (1978). The testing of information search tactics. *Proceedings of the American Society for Information Science* 15: 25-27.
- Bates, M. (1979). Information search tactics. *Journal of the American Society for Information Science* 30, 4 (July): 205-214.
- Bates, M. (1985). The design of browsing and berrypicking techniques for the online search interface. *Online Review* 13, 5 (October): 407-424.

- Belkin, N. (1980). Anomalous states of knowledge as a basis for information retrieval. *Canadian Journal of Information Science* 5 : 133–143.
- Bhavnani, S. (1998). How Architects Draw with Computers: A Cognitive Analysis of Real-World CAD Interactions. Unpublished Ph.D. dissertation, Carnegie Mellon University, Pittsburgh.
- Bhavnani, S. (2000). Designs Conducive to the Use of Efficient Strategies. *Proceedings of the DIS' 00 Conference*, 338-345. New York: ACM.
- Bhavnani, S., Flemming, U., Forsythe, D., Garrett, J., Shaw, D., & Tsai, A. (1996). CAD Usage in an Architectural Office: From Observations to Active Assistance. *Automation in Construction* 5, 243-255.
- Bhavnani S., & John, B. (1996). Exploring the Unrealized Potential of Computer-Aided Drafting, *Proceedings of the CHI' 96 Conference on Human Factors in Computing Systems*, 332-339. New York: ACM.
- Bhavnani, S., & John, B. (1997). From Sufficient to Efficient Usage: An Analysis of Strategic Knowledge. *Proceedings of the CHI' 97 Conference on Human Factors in Computing Systems*, 91-98. New York: ACM.
- Bhavnani, S., & John, B. (1998). Delegation and Circumvention: Two Faces of Efficiency. *Proceedings of the CHI' 98 Conference on Human Factors in Computing Systems*, 273-280. New York: ACM.
- Bhavnani, S., & John, B. (2000). The Strategic Use of Complex Computer Systems. *Human-Computer Interaction*, 15, 107-137.
- Bhavnani, S., John, B. & Flemming, U. (1999). The Strategic Use of CAD: An Empirically Inspired, Theory-Based Course. *Proceedings of the CHI' 99 Conference on Human Factors in Computing Systems*, 42-49. New York: ACM.
- Bhavnani, S., & John, B. (in press). Beyond Command Knowledge: Identifying and Teaching Strategic Knowledge for Using Complex Computer Applications.. *Proceedings of the CHI' 01 Conference on Human Factors in Computing Systems*. New York: ACM.
- Borgman, C. (1986). Why are online catalogs hard to use? Lessons learned from information retrieval studies. *Journal of the American Society for Information Science* 37, 6: 387–400.
- Card, S., Moran, T., & Newell, A. (1983). *The Psychology of Human-Computer Interaction*. Hillsdale, NJ: Lawrence Erlbaum Associates.
- Carter, D., and Janes, J. (in press). Unobtrusive data analysis of digital reference questions and service at the Internet Public Library: an exploratory study. *Library Trends*.
- Committee on Information Technology Literacy (1999). *Being Fluent with Information Technology*. Washington D.C : National Academy Press.
- Dervin, B. (1983). *An overview of sense-making research: Concepts, methods, and results to date*. Seattle: School of Communicatiins, University of Washington.
- Doane, S., Pellegrino, J., & Klatzky, R. (1990). Expertise in a Computer Operating System: Conceptualization and Performance. *Human-Computer Interaction*, 5, 267-304.
- Drabenstott, Karen M.(1991) Online catalog user needs and behavior. In: *Think tank on the present and future of the online catalog: Proceedings*, edited by Noelle Van Pulis, 59–83. Chicago: American Library Association.

- Drabenstott, K. (2000). Web search strategies. In : *Saving the User's time through subject access innovation; Papers in honor of Pauline Atherton Caochrane*, edited by William J. Wheeler, 114–161. Champaign, Ill.: University of Illinois.
- Fidel, R. (1985). Moves in online searching. *Online Review*, 9, 1, 61-74.
- Fidel, R. (1991) Searchers' Selection of Search keys: 1. The Selection Routine. *Journal of the American Society for Information Science* 42, 7 490-500.
- Harter, S. (1986). *Online information retrieval: Concepts, principles, and techniques*. Orlando, Fla.: Academic Press.
- Jansen, B., et al., (1998). What do they search for on the web and how are they searching: a study of a large sample of Excite searches. *Proceedings of SIGIR '98: 21st Annual International ACM SIGIR Conference on Research and Development in Information Retrieval*, 5–17.
- Kirby, M., & Miller, N. 1986. MEDLINE searching in Colleague: Reasons for failure or success of untrained end users. *Medical Reference Services Quarterly* 5: 17–34.
- Kuhlthau, C. (1991). Inside the search process: information seeking from the user's perspective. *Journal of the American Society for Information Science* 42, 5 (June): 361–371.
- Lang, G., Eberts, R., Gabel, M., & Barash, M. (1991). Extracting and Using Procedural Knowledge in a CAD Task. *IEEE Transactions on Engineering Management*, 38, 257-68.
- Marchionini, G. (1989). Information seeking in electronic encyclopedias, *Machine-Mediated Learning*, 3, 3, 21-26.
- Margolis, J. and Fisher, A. (1997). Geek Mythology and Attracting Undergraduate Women to Computer Science. In *Impacting Change Through Collaboration*, Proceedings of the Joint National Conference of the Women in Engineering Program Advocates network and the National Association of Minority Engineering Program Administrators.
- Markey, K., and Atherton, P. (1978). *ONTAP: Online training and practice for ERIC data base searchers*. Syracuse, NY: ERIC/Information Resources Clearinghouse. 198 p. ED 160109.
- Meadow, C., and Atherton, P. (1981). *Basics of online searching*. New York: Wiley.
- Newell, A., & Simon, H. (1972). *Human Problem Solving*. New Jersey: Prentice Hall.
- Nilsen, E., Jong, H., Olson J., Biolsi, I., & Mutter, S. (1993). The Growth of Software Skill: A Longitudinal Look at Learning and Performance. *Proceedings of INTERCHI'93*, 149-156. New York: ACM.
- Peters, T. (1993). The history and development of transaction log analysis. *Library Hi Tech* 11, 2: 41–66.
- Rosson, M. (1983). Patterns of Experience in Text Editing. *Proceedings of CHI' 93 Conference on Human Factors in Computing Systems*, 171-175, New York: ACM.
- Pfaffenberger, B. (1996). *Web search strategies*. New York: MIS Press.
- Prager, J., Brown E., Coden, A and Radev, D. (2000). Question-answering by predictive annotation. *Proceedings of the 23rd Annual International Conference on Research and Development in Information Retrieval*.

- Radev, D., Prager, J., and Samn, V. (2000). Ranking potential answers to natural language questions, *Proceedings of the 6th Conference on Applied Natural Language Processing*.
- Radev, D., Blair-Goldensohn, S., Zhang, Z., Revathi, S. (in review) Interactive, Domain-Independent Identification and Summarization of Topically Related News Articles.
- Radev, D., Fan, W., and Zhang, Z. (in review). A Personalized Web-Based Multi-Document Summarization and Recommendation System.
- Saracevic, T. (1996). Modeling Interaction in Information Retrieval (IR): A Review and Proposal. *Proceedings of the 59th Annual Meeting of the American Society of Information Science*, 3-9.
- Singley, M., & Anderson, J. (1989). *The Transfer of Cognitive Skill*. Cambridge, MA: Harvard University Press.
- Simon, H.A. 1996 Observations on the Sciences of Science Learning. Paper prepared for the Committee on Developments in the Science of Learning for the Sciences of Science Learning: An Interdisciplinary Discussion. Department of Psychology, Carnegie Mellon University.
- Spink, A., and Saracevic, T. (1997). Interactive information retrieval: Sources and effectiveness of search terms during mediated online searching. *Journal of the American Society for Information Science* 48, 8 741-761.
- Spool, J. et al. (1999) Web site usability: A designer's guide. San Francisco: Morgan Kaufmann Publishers.
- Taylor, R. (1968). Question-negotiation and information seeking in libraries. *College & Research Libraries* 29: 178-194.
- Thomas, R. and Foster, M. (2001) A pilot study of teaching the strategic use of common computer applications. *Australasian User Interface Conference 2001 - Australian Computer Science Communications*, Volume 23, Number 5, IEEE Press, 2001.
- Walker, G., and Janes, J. (1999). *Online retrieval: A dialogue of theory and practice*. 2nd ed. Englewood, Colo.: Libraries Unlimited.
- Xie, H. 1997. Planned and situated aspects in interactive IR: Patterns of use, interactive intentions and information seeking strategies. In: *Proceedings of the 60th ASIS annual meeting*, 101-110. Medford, N. J.: Information Today.