Interactive Information Retrieval

Ian Ruthven
University of Strathclyde

Introduction

Information retrieval is a fundamental component of human information behavior. The ability to extract useful information from large electronic resources not only is one of the main activities of individuals online but is an essential skill for most professional groups and a means of achieving competitive advantage.

Our electronic information world is becoming increasingly complex with more sources of information, types of information, and ways to access information than ever before. Anyone who searches for information is required to make more decisions about searching and expected to engage with an increased number and variety of search systems. Even a standard desktop personal computer comes equipped with numerous search tools (desktop search, e-mail search, browsers to help search the Internet, embedded search tools for specific file formats such as PDF [portable document format] or Word, and specific document types such as help manuals). A standard day, if one is electronically enabled, may involve many searches across different search systems accessing different electronic resources for different purposes. The Internet, in particular, has revolutionized the ability to search, especially in the commercial arena where we have the choice of using different search systems to search essentially the same electronic resources but with different interactive functionalities. The search decisions a human is required to make before encountering any information involve not only how to search *this* resource using *this* system but also how to choose a system or resource to search in the first place. These decisions are complicated because skills learned using one type of system do not always transfer simply to searching a different type of system (Cool, Park, Belkin, Koenemann, & Ng, 1996). Neither does information literacy in one domain of expertise necessarily help when searching on unfamiliar topics.

The variability of data available, and the explicit or implicit structures of the data, also place a burden on both the searchers and system designers. How does searching within a Weblog, for example, differ from searching within technical manuals; or does all searching involve the same activities and require the same user support? As research shows

(e.g., Florance & Marchionini, 1995; Ford & Ford, 1993; Kim & Allan, 2002) people often come to information retrieval (IR) systems with existing approaches to information seeking and processing and develop strategies for using specific systems. Neither search success nor a searcher's satisfaction with a system necessarily depends solely on what interactive features a system offers or on how it encourages searchers to employ these features; success and satisfaction instead depend on how well the system supports the searcher's personal strategies and how well it leads the searcher to understand how the system operates (Cool et al., 1996). Many authors have pointed out that individual differences affect interaction with information and information systems (e.g., Chen, Czerwinski, & Macredie, 2000; Ford, Miller, & Moss, 2005; Slone, 2002), that different stages of the search process require different kinds of assistance (Belkin, Cool, Stein, & Thiel, 1995; Kuhlthau, 1991), and that differences in the search context affect the interactive support required—for example searching in secondary languages requires more support in the process of document assessment and querying (Hansen & Karlgren, 2005; López-Ostenero, Gonzalo, & Verdejo, 2005).

The area of interactive information retrieval (IIR) covers research related to studying and assisting these diverse end users of information access and retrieval systems. IIR itself is shaped by (a) research on information seeking and search behavior and (b) research on the development of new methods of interacting with electronic resources. Both approaches are important; information seeking research provides the big picture on the decisions involved in finding information that contextualizes much of the work in IIR; research on methods of interacting with search systems promotes new understandings of appropriate methods to facilitate information access. The latter aspect of IIR is the main area covered in this chapter, the aim of which is to study recent and emerging trends in IIR interfaces and interactive systems.

Scope

People can find or become aware of useful information in many ways. We can receive recommendations through information filtering (Robertson & Callan, 2005) or collaborative filtering (Konstan, 2004), both of which push information toward us based on some model of our information preferences. We can follow information paths by traversing a series of items that have been manually or automatically linked to provide a narrative (Shipman, Furuta, Brenner, Chung, & Hsieh, 2000) or by creating our own information paths through browsing (Cove & Walsh, 1988). We can, of course, also find information by chance: looking for one piece of information and uncovering an unexpected piece of useful information. As Foster and Ford (2003, p. 337) note: "Perceptions [of the study participants] of the extent to which serendipity could be induced were mixed. While it was felt that some element of control could be exercised to attract 'chance encounters,' there was a perception that such

encounters may really be manifestations of the hidden, but logical, influences of information gatekeepers—inherent in, for example, library classification schemes." This suggests that IIR systems could be designed to help find useful information by chance by reusing existing techniques for purposely finding information.

More commonly, electronic support for information seeking and retrieval consists of two types of systems: query-based and browse-based. Query-based systems differ from filtering systems as they force searchers to pull information out of the stored resource by expressing a request. Browsing systems—systems that are designed to support as opposed to simply permit browsing—help searchers understand and navigate an information space. In this chapter I deal with both types of systems, concentrating more on querying systems.

Dealing with the interactive issues involved in all types of information access is too wide an area to cover in one chapter. I focus specifically on the idea of a person interacting with a dedicated search system and the interaction engendered and supported by the system and interface design rather than discussing general search behavior, although, as will be seen, these two areas are linked. So, although this chapter discusses issues such as assessment of relevance and information behavior where appropriate, it does not discuss, in depth, issues such as work tasks or general information seeking behavior. The aim is to produce a chapter that is complementary to those by Vakkari (2002) on task-based information searching and Case (2006) on information seeking.

Scoping research on IIR is also problematic because research on developing better interactive systems often has an impact not only at the interface or dialogue level but also on the design of the whole system. Similarly, many articles discuss systems that have a novel interface, from which we can learn something about interaction, but the main aim of the research is neither the interface nor interaction. Finally, one could argue that almost all IR is interactive; most IR systems have some kind of interface and searchers are required to engage in some form of interaction. In deciding what to cover, I have tried to concentrate on systems where the novel features are interface- or interaction-related or where there is a human-centered evaluation to assess the interactive quality of the system; that is, where the intention behind the research is to investigate new forms of interaction, evaluate existing forms, or exploit user interaction for improved search effectiveness. Much of the research reviewed in this chapter is evaluated by experiments or studies with human participants. The variability of the experimental details and participants involved in these studies makes it difficult to compare directly, at a quantitative level, the results obtained. Therefore, although I discuss the relative success or failure of various approaches, I mostly compare the studies at a qualitative level.

Because IR is not an isolated field, another scoping issue arises as developments outside IR naturally have an impact on solutions to interactive IR questions. The rise of ontologies, for example, as part of the

Semantic Web initiative in artificial intelligence has provided new impetus to the area of knowledge-based query expansion (e.g., Legg, 2007; Navigli & Velardi, 2003). Similarly, technological advances in mobile computing have stimulated research in the area of contextual information retrieval, where context includes location, user tasks, and personal preferences. I do not touch on the technical sides of these developments but consider, where appropriate, the interactive issues raised.

The chapter concentrates on research published since 2000, mentioning early influences on current research where relevant.

Sources of Information

Interactive information seeking and retrieval is of interest to many communities and, as a result, work in this area is diffused across academic and practitioner fora. The main IR journals such as *Journal of the American Society for Information Science and Technology*, *Information Processing & Management*, *Journal of Documentation*, *Journal of Information Retrieval*, and *ACM Transactions on Information Systems* all regularly present high-quality research articles on IIR as do the leading journals in human–computer interaction (HCI), including *ACM Transactions on Computer–Human Interaction*, *Interacting with Computers*, and, to a lesser extent, *Human–Computer Interaction*.

Conferences are also a good source of material. The main IR conferences, the Association for Computing Machinery (ACM) Special Interest Group on Information Retrieval (SIGIR), the European Conference on Information Retrieval (ECIR), the International Conference on Information and Knowledge Management (CIKM) contain work on IIR, although the emphasis of late has been less on interfaces and more on system components such as relevance feedback, personalization, and techniques that could form part of an interface (e.g., summarization or clustering). The ACM Special Interest Group on Computer–Human Interaction (CHI), the Annual Meeting of the American Society for Information Science and Technology (ASIST), the World Wide Web (WWW) and Digital Library conferences, notably the Joint Conference on Digital Libraries (JCDL), and the European Conference on Digital Libraries (ECDL) also contain work on interactive information retrieval.

TREC (Text REtrieval Conference, trec.nist.gov) has dedicated efforts on interactive searching, notably the Interactive Track (1995–2003) and the HARD track (2003–2005), although tracks such as the video TREC-VID (from 2001) have also influenced interactive work in TREC. All TREC proceedings are available from the TREC Web site, and Dumais and Belkin (2005) provide a useful history of the TREC approach to interaction updating the previous history by Over (2001). Other initiatives such as CLEF (Cross-Language Evaluation Forum, www.clef-campaign.org) and INEX (INitiative for the Evaluation of XML [Extensible Markup Language] Retrieval, inex.is.informatik.uni-duisburg.de) also contain regular interactive tracks.

These are the main sources of materials on IIR, the ones I have used primarily for this chapter, but most conferences in the wide areas of IR, information science, librarianship, HCI, and the Web, as well as other less obvious places, such as conferences on social computing, will include occasional papers reflecting the pervasive nature of information access.

There is no single monograph dealing solely with IIR although there are a number of dedicated monographs or collections of edited works addressing related areas. Numerous "how-to" books on optimizing end-user searching strategies and awareness (e.g., Hill, 2004) indicate the need for user support in searching. Hearst's (2000) chapter in *Modern Information Retrieval* is still worth reading. *The Turn*, by Ingwersen and Järvelin (2005), serves as a companion to Ingwersen's (1992) earlier work, which set out to provide a cognitive account of interactive information seeking. Other contributions teach us about information seeking and behavior, which, in turn, help specify the role of IIR and define the broader context in which these systems are used. Examples include the two recent collections edited by Spink and Cole on human information behavior (Spink & Cole, 2005b) and cognitive information retrieval (Spink & Cole, 2005a). Cognitive information retrieval, in this context, is focused on the human's role in information retrieval.

The question does arise of whether IIR is a distinct research area or simply a subfield of HCI (Beaulieu, 2000). Obviously one's own position does lend a particular view; but it is clear that interactive IR is more than simply developing interfaces for searching (Shneiderman, Byrd, & Croft, 1998) and that the strength of good research in IIR comes not only from a technical knowledge of interactive systems development but also from a knowledge of people's search behavior and search context, including the environmental factors that influence behavior (Fidel & Pejtersen, 2004). A particular strength of information seeking and retrieval as a hybrid domain is the awareness of the importance of the information objects themselves; not simply the media type being searched but also the generation, use, and storage of these objects (Blair, 2002). The notion of an information collection as more than simply a searchable grouping of objects is a powerful concept often under-utilized in IIR systems.

HCI and IIR come from different traditions; HCI, for example, places more emphasis on the published literature on usability whereas IR emphasizes effectiveness. Both, of course, are important, as a system with low usability will typically have low effectiveness and we probably care little about the usability of a poor system. Interactive IR does not stop at the interface and, as Bates (2002) and others point out, IIR system design is a coherent whole rather than a set of units. However, the two fields can learn from each other and the best research in IIR often reflects best practice in HCI as well as IR.

Themes

All research fields have stereotypes, idealized views of the aims and role of the activities within the field that are used to focus its intellectual debates and research agendas. Interactive information retrieval is no exception. The idealized IIR session is conceptually and practically simple: An end user creates some form of information request; this request is put to an electronic search system that retrieves a number of information objects, or references to these objects; the end user then assesses the set of retrieved results and extracts the relevant objects or information. For many searches, especially for straightforward types of Web search, this idealized view suffices and the interactive process is simple for the searcher. However, for the system designer, even this most simple view of searching raises interactive design issues—how does the system facilitate good queries or make it easier for the searcher to assess the retrieved material, for example?

A simplistic account of the interaction involved in searching eliminates many of the aspects that make interactive searching difficult for both searchers and designers of search systems. It also ignores the fact that information seeking and retrieval are usually only part of some larger activity and not ends in themselves. This larger activity, variously termed the task, domain task, or work task (Ingwersen, 1992, p. 131), influences our interaction with a system and our expectations of the interaction.

Although searches are commonly viewed and described at the session level—a series of interactive steps within a fixed time frame or terminated by a specific action such as the searcher leaving the system—we often repeat searches at different intervals (Spink, 1996; Vakkari, 2001). This can be to re-find the same information (Dumais, Cutrell, Cadiz, Jancke, Sarin, & Robbins, 2004), to update ourselves on new information provided within a dynamic information resource (Ellis, 1989; Ellis, Cox, & Hall, 1993), or because we are engaged in an ongoing activity and require additional information on the same topic (Vakkari, 2001; Vakkari & Hakala, 2000; Vakkari, Pennanen, & Serola, 2003). We may also be forced to repeat searches across different systems because no single source can completely satisfy an information need (Bhavnani, 2005). A repeated search can, therefore, be a request for the same information, for new information, or for different information—even though the search requests may appear very similar. Lin and Belkin (Lin, 2005; Lin & Belkin, 2005) demonstrate elegantly how complex is the nature of successive searching compared to the idealized one-iteration model.

Even within a single search session, the individual steps involved in completing a search may be interactively simple but not cognitively simple. We do not, for example, always know what information we require in advance of creating an information request or we may find it difficult to express our need for information as a searchable request (Belkin, 1980). The material retrieved may be too large to analyze easily and may

require refinement resulting in a need for multiple query iterations. These refinements may be difficult to create and, even if the retrieval system offers the capability, it may be difficult to recognize good refinements (Ruthven, 2003). Assessing the retrieval results to select relevant material may be simple if we can easily recognize the relevant, or correct, information. On the other hand, it may be much more difficult if we have less certainty regarding the quality or accuracy of the information returned; and here the tasks that initiated the search in the first place may affect which criteria we use to assess the retrieval material (Barry & Schamber, 1998).

Searching involves a series of decisions; each decision may be influenced by the task, the collection, and factors relating to the person engaged in the search. Consequently, designing interactive systems that support how people search and, more importantly, how they want to search raises many intellectual challenges. Historically, there have been two dominant lines of research on helping people search for information: a major research thrust on automating or semi-automating tasks that humans may find difficult to perform on their own, and an equally important line of research on providing enhanced interactive functionality to allow humans more control over and input into their own search process. Both of these fields are still very much evident in recent research in IIR and the discussion presented here focuses on the research in both areas.

Improving Interaction

In the first line of research—improving interactive support for searchers—we see both novel interfaces and novel interactive functionality that help users organize information, structure their investigation of an information resource, or make interactive decisions.

The rise of the World Wide Web and the availability of Internet search engines such as MSN Search, Google, and AltaVista have radically changed perceptions of searching. Web search engines have changed the search landscape by making the ability to search more widely available than before. The effects of this availability have raised new challenges not least because the users of these systems are extremely diverse. The popularity and availability of Web search engines are particularly important in creating users' models of how search engines operate and users' expectations of search engines in general (Muramatsu & Pratt, 2001).

Web search engines, although freely available, are driven by commercial interests. This means that user interfaces developed for this type of searching may well have different aims than more traditional interfaces, but the search engine providers have an unparalleled opportunity to test new interactive techniques on a very large sample of end users. The dominance of particular behaviors in Web searching, for example, short queries, few page accesses, and little use of advanced

search features, translate into new interactive challenges (Jansen & Spink, 2006). If, for example, most people use very short queries, how do we gain better representations of their information needs? Techniques such as query intention analysis have been suggested for this purpose (Kang & Kim, 2003). Similarly, if people look at only a very few results on a search page, what techniques will help users optimize the information they obtain? Here, techniques such as clustering and novel surrogates have attracted great attention.

Designing interfaces that support more difficult interactive decisions, such as selecting good query refinements, is also challenging for searchers who have learned to expect easy answers via the Web. The rise of the Web itself has had a huge impact on the development of new interactive retrieval systems and interfaces, with much of the recent work on general search interfaces using the Web as a source collection.

The prevailing model has been the query-driven approach in which a human enters a query and retrieves a list of references to information objects; this is the model favored by most search systems. Consequently, this section starts with a discussion of query formulation/reformulation and also of surrogates. These two areas represent the inputs and outputs of the querying approach: how to obtain queries and how to present the results of the retrieval. I then discuss the major alternatives to query models such as clustering, categorization, and visualization approaches incorporating some notion of information organization at the interface level. Finally, I discuss some newer trends in the literature, specifically work on personal information management, subobject retrieval, and systems for specialized retrieval tasks.

Query Formulation

Searchers often begin with a query and query-driven search interfaces rely on searchers being able to form an information request in a manner understandable by the underlying search engine. The typical querying interface accepts as a query a natural language keyword-based statement, one without operators such as Boolean connectives. Creating a good initial query is regarded as important for many reasons; it can increase search effectiveness and searcher satisfaction (Belkin, Kelly, Kim, Kim, Lee, Muresan, et al., 2003) and it facilitates automatic techniques for improving query quality such as pseudo-relevance feedback (Lynam, Buckley, Clark, & Cormack, 2004; Ruthven & Lalmas, 2003).

A system definition of a good query is one that helps discriminate objects that the searcher will judge relevant from those that the searcher will judge non-relevant, and can prioritize retrieval of the relevant material. A more human interpretation of a good query is one that returns appropriate or expected results. Depending on the user's stage of the search process, the notion of appropriate search results may be very different. Individuals carrying out an initial search, or with little knowledge of the topic being searched, may be satisfied with results that

inform them about the information space being searched (Kuhlthau, 1991). Alternatively, a searcher who has good topical knowledge, good knowledge of the information problem being tackled, and a clear view of what information is required may have very specific criteria in mind for the end result (Kuhlthau, 1991). A good retrieval result, therefore, is related to the searcher's expectations of the search.

Interactive systems can help searchers construct good queries in various ways. Options include automatically modifying the searcher's query by relevance or pseudo-relevance feedback or, more radically, by replacing the searcher's query through query substitution (Jones, Rey, Madani, & Greiner, 2006; Kraft, Chang, Maghoul, & Kumar, 2006). Other options allow queries to develop through interaction, as in faceted browsing interfaces. Yet others are more interactive—either offering query suggestions to searchers or allowing searchers to be more precise in how they construct queries by developing complex querying languages or using advanced search features.

Complex Query Languages

Complex, or structured, query languages can facilitate more precise access to complex objects. Complex languages can be useful where the searcher wants to be very precise through the use of a detailed query (e.g., Pirkola, Puolamäki, & Järvelin, 2003) or where the data themselves are complex, for example, music data, which comprise different attributes such as timbre, tone, and pitch (Downie, 2004), each of which might be expressed as individual query components. Niemi, Junkkari, Järvelin, and Viita (2004) provide an example of the latter approach.

Structured query languages have attracted attention through the increased use of XML as a general description language for Web information (Chinenyanga & Kushmerick, 2001; Führ & Großjohann, 2001). Evidence for the success of complex querying languages is mixed. Järvelin and Kekäläinen (2000) have shown that structuring the content of the query is generally beneficial and that good structures can facilitate additional techniques such as query expansion. However, query languages that allow for mixing content and structural information about the document are often not easy for searchers to create and, as explained by Kamps, Marx, de Rijke, and Sigurbjörnsson (2005) and O'Keefe and Trotman (2004), query languages that are difficult to conceptualize can lead to more semantic mistakes within the query, especially if the searcher is not aware of the document structures being searched.

"Advanced" Search

An alternative to complex query languages is to offer form-based support in which searchers are asked questions about the material they wish to retrieve. Answering these questions, always assuming the searchers can answer them, produces a more sophisticated and precise query than a simple keyword request. The most common instantiation of

interactive support is the advanced search features of search engines, which allow for the inclusion of metadata reflecting non-content aspects of the objects being searched. Google and AltaVista, for example, offer date range, file type, and domain restrictions among their advanced search features. Typically these restrict the objects returned in some way, cutting down the number of results rather than prompting the searcher with new ideas for queries or the content of queries. As such, these search facilities may not seem very advanced but I retain this term as the one most commonly advertised by search engine interfaces. Typically, search engines will also offer query operators such as phrase matching and Boolean-like operators ("all of these words," "none of these words," etc.).

Interestingly, the intended effect of these operators does not always match their actual effect on retrieval (Eastman & Jansen, 2003), meaning the searcher may have problems using these operators effectively. The effect of any particular operator is dependent on the implementation of the individual system, which can vary. Such lack of consistency between operators reflects earlier concerns about the usability of Boolean IR systems (Borgman, 1996). Topi and Lucas (2005a, 2005b) suggest that interfaces that support a greater number of query operators and independent training on using such operators *generally* help in improving query quality; however, this improvement is not consistent and it can be difficult to predict what training will help. Of course, most Web searchers do not have any training in online searching.

It is commonly reported that searchers often do not use advanced search features, such as Boolean operators or relevance feedback, to any great extent (e.g., Spink & Jansen, 2004). This could be because they do not understand how to use them, are not aware that they are available, or because the actual support is not viewed as being useful. However, even if the utilization is low, the fact that people try these features suggests that users often want *something* to support query formulation. Whatever the reason for the low use of advanced search features, we have to consider different styles of interactive query support or automating this support to be in the area of query intention analysis.

Asking for More Information

A common finding in Web studies is that users enter short queries, perhaps one to three query terms, which may require immediate reformulation. Belkin et al. (2003) consider the degree to which this might be a problem and how to persuade users to enter longer and more effective queries. Simply asking users for more information on their tasks helps them enter longer queries and results in shorter searches, although with equal effectiveness. Kelly et al. in a robust follow-up ask searchers for more information regarding their prior knowledge of the topic, the intended purpose of the information being sought, and additional search terms (Kelly, Deepak, & Fu, 2005). The results show such an approach—

simply asking for more information—to be very successful, outperforming pseudo-feedback approaches.

The key here is that searchers often know more about what information they want than they provide in a query. By asking good questions, interfaces that prompt searchers to enter more information can improve retrieval effectiveness. Determining which questions are good, however, can vary according to the task being performed or the domain being searched (Toms, 2002); and specialized interfaces for searching within individual domains may be more appropriate than generic one-size-fits-all interfaces.

Offering Suggestions

The system itself can offer suggestions for query creation. Google's Suggest feature proposes queries using a form of predictive text entry. As the searcher types a query, the system tries to match it to previously submitted queries. White and Marchionini (2007) attempted to replicate and evaluate Google's Suggest facility. Their comparison was similar to Koenemann and Belkin's (1996) investigation of the effects of offering query suggestions either before or after running a query. As with the Koenemann and Belkin study, offering query suggestions before running a query improves results, but not all suggestions prove to be good suggestions. Nevertheless, such a mechanism would seem to be a useful step in supporting query creation.

Query Reformulation

Queries are often reformulated by searchers after an initial, perhaps tentative, query has been run. Interactive reformulation—where the searcher controls how the query is reformulated—is a core area for IIR and a stream of research has investigated how to select good reformulation suggestions. Recently, the trend has been toward more complex refinement suggestions instead of single query terms. Kruschwitz and Al-Bakour (2005) automatically extract concepts, essentially phrases, to provide a domain model of a corpus. These concepts are mixed with terms from top-ranked documents (the terms providing new information not present in the domain model) for presentation to the searcher. The results are mixed but clearly show that the participants are willing at least to experiment with the novel interface. However, searcher attitude is important. Some participants appreciate the attempt to support query reformulation but others appear to have a low tolerance of inappropriate reformulation suggestions.

Query reformulation is supported on some Web search engines but less than the advanced search features mentioned earlier. Although Web search engines are very influential and contribute heavily to developing people's experiences of searching, the actual mechanisms are often not described or evaluated in the public literature and we must infer their design principles. There are some exceptions; for example Anick (2003)

examines terminological support—suggesting key phrases derived from a pseudo-relevance feedback approach—on AltaVista. This implementation is similar in spirit to Koenemann and Belkin's (1996) notion of transparent interactive query expansion—showing expansion units for user selection after a query has been run. In this case the phrases are based on linguistic analysis, not derived from user-identified relevant documents. A major finding is that people use this terminological support, continue to use it in later searches, and generally use it successfully. However, most query reformulations continue to be manual.

As with many large-scale analyses, these findings are based on log analysis, using cookies to track individual users; deeper analysis shows that people can become confused about the nature and role of the phrases offered. This confusion can be resolved in part by providing more information about the reformulations (Ruthven, 2002) but a limiting factor in any interface is screen space. In Web or commercial interfaces, extra screen space may come at the expense of advertisements (hence revenue), requiring developers to be even more imaginative in deciding how to support searchers.

Bruza, McArthur, and Dennis (2000) compare interfaces that offer linguistically well-formed phrases for refinement with more traditional search interfaces and make two points generalizable to any method of suggesting refinements. First, although refinement suggestions can make searchers aware of useful search concepts, searchers also must realize the benefits of such refinement. That is, searchers will need to understand why refining a query could be a useful undertaking (Dennis, McArthur, & Bruza, 1998). Support for assessing the effect of any particular refinement on a search would also be useful (Ruthven, 2003). Second, a given individual interactive technique may lead to more effective searching; it also needs to be attractive to searchers—a technique that entails more work may not be used unless the benefits are very clear (Bruza et al., 2000).

Phrases are not the only unit that can be offered to searchers, although they are easy for searchers to interpret. Historically single terms have been the most studied means for interactive query expansion. Efthimiadis (2000) demonstrates the general effectiveness of interactive query suggestion and expansion and also its power in stimulating new ideas for search requests. D. He and Demner-Fushman (2004) also indicate that interactive query refinement is useful in cases where few relevant documents are available. This strength of interactive query refinement—its ability to support difficult search tasks—is endorsed by Fowkes and Beaulieu (2000) who show interactive refinement to be more effective and appropriate for complex tasks. Sihvonen and Vakkari (2004), investigating suggestions from a thesaurus rather than relevance information, find an increased use of terms for difficult tasks. This study indicates strongly that topical knowledge improves thesaurus use in that searchers are better at selecting which aspects of a topic are

important and are more informed about which terms are likely to be appropriate.

How to offer reformulation suggestions has been less investigated, with most interface approaches simply presenting lists of suggestions. These may be structured in some way, for example using facets (Hearst, 2006b), but generally do not support much decision making on the quality or appropriateness of the suggestions. Rieh and Xie (2006), while examining query reformulation, note that most interaction takes place at the surface level; that is dealing with queries and results rather than deeper cognitive aspects such as searcher intent or attitude. However, the intent behind formulation is obviously important and needs support. Using a log analysis and original coding scheme, Rieh and Xie attempt to categorize query reformulation behavior and find, perhaps not surprisingly, that content reformation is most common and strict synonym replacement is rare. Specialization of queries and parallel movements, tackling different aspects of a query, are much more common—the latter being essentially multitasking (see Spink et al., 2006); this reinforces the call for more interactive support for this type of searching. If we can reliably recognize different types of query reformulation behavior, then it would be useful to see if we can predict query reformulations that support this behavior and make these suggestions clearer at the interface level.

An alternative to interactive query reformulation is, of course, to provide automatic support to refine queries, using either some form of knowledge-based query reformulation (Liu, Lieberman, & Selker, 2002; Pu, Chuang, & Yang, 2002) or techniques such as pseudo-relevance feedback (Crouch, Crouch, Chen, & Holtz, 2002). True user relevance feedback—searchers giving explicit feedback on the relevance of retrieved items—has remained popular, especially for non-textual objects such as images where the required objects are easy to recognize although perhaps harder to describe. There has been less work recently on the usability of relevance feedback, as opposed to the underlying algorithms, but also much more work on implicit feedback, to be discussed later.

Surrogates

After submitting a search request one is usually presented with a set of results; an important aspect of searching is assessing these results. For some objects, such as images, the complete objects themselves are displayed and assessed. More often surrogates are employed—keyframes for video; sentences, titles, or abstracts for documents; thumbnails for Web pages; and so on—and these can be created manually (document titles) or automatically (such as summaries).

The design and role of these surrogates within searching is of interest to IIR, especially to facilitate quick review of retrieval results and access to useful information. Novel elicitation methods such as eyetracking allow us to learn more about how people use such surrogates in searching. Lorigo et al. (2006) indicate that, in more than half of the Web

searches they investigated, users reformulate queries based on scanning the surrogates without examining any pages and that navigational (Web site-finding) queries are often answered by surrogates alone. Surrogates can be useful in these cases if searchers are willing to accept occasional false hits. False hits in this case are pages appearing to be relevant because the surrogate misrepresents a page's content; this can arise from surrogates being created from cached pages (automatically generated surrogates) or deliberate misrepresentation of the page's content (Lynch, 2001). False hits can occur with most types of surrogates and also in non-Web environments; for example, Ruthven, Tombros, and Jose (2001) report a similar finding with query-biased summaries. The quick response speeds of most search engines may mean that such mistakes are not important because people can recover from them with little cost.

Summarization approaches are particularly popular for creating surrogates (e.g., Chuang & Yang 2000; Tombros & Sanderson, 1998; White, Jose, & Ruthven, 2003). Most summaries are text-based, although summaries are also possible for non-textual and mixed media, such as music videos (Xu, Shao, Maddage, & Kankanhalli, 2005). In Web research several studies have compared the relative effectiveness of the standard text-based summaries and summaries that incorporate more visual aspects commonly found in Web pages. Woodruff, Faulring, Rosenholtz, Morrison, and Pirolli (2001) report that different types of surrogates work well for different types of search task. However, some form of aggregate surrogate, incorporating Web page thumbnails and text, is best for most tasks and appears to be a safe default. Dziadosz and Chandrasekar (2002) also find that certain types of surrogates work better for certain tasks (e.g., thumbnails can be less effective than textual summaries when searching for unknown items) and report that the presence of both thumbnails and text lead to more predictions of relevant material but also more incorrect predictions of relevance.

This issue of prediction is important; a good surrogate should allow searchers to make informed decisions about the content of the object being represented. Vechtomova and Karamuftuoglu (2006) examine the quality of query-biased sentences as surrogates: assessors being asked to predict the decisions they would make on the relevance of documents based on sentence surrogates. Assessors are generally fairly good at predicting relevance, although the actual results may vary depending on the quality of the sentence selection mechanism. In a separate study Ruthven, Baillie, and Elsweiler (in press) show that sentences, in this case leading sentences from newspapers, can result in good prediction of relevance but this depends on the personal characteristics of the individual making the assessment. When given the choice, some assessors would rather not make a prediction and variation in characteristics such as the assessor's knowledge level can lead to very poor predictions. However, Bell and Ruthven (2004) show that sentence surrogates are useful in reducing the complexity of searches by allowing searchers to

see an overview of the retrieved material without having to access individual documents serially.

Where summaries are particularly useful is in the area of mobile IR, or more precisely IR performed on hand-held devices with small screens and requiring different methods of information presentation. Thumbnails can be used here as well as typical text summaries, which may (Buyukkokten, Kalijuvee, Garcia-Molina, Paepcke, & Winograd, 2002) or may not use information on the structure of the document being summarized (Sweeney & Crestani, 2006).

Sweeney and Crestani (2006) point to an interesting distinction between effectiveness and preference. In their study of optimal summary length for handheld device presentation they find that people prefer longer summaries on larger devices but this does not make them more accurate at using summaries to predict relevance. Radev, Jing, Styś, and Tam (2004), also looking at summary length, find that assessors can be in agreement on the most important sentences in a summary, but longer summaries can, in certain cases, reduce agreement. As summaries become longer less important sentences are included. Document retrieval provides a similar analogy. Typical IR systems will first retrieve documents upon whose relevance most people would agree followed by more marginally relevant documents (unanimity is lacking) (Voorhees, 2000). What makes a good surrogate, then, depends on the searcher's context.

Evaluation of summarization systems follows two approaches: so-called intrinsic evaluations measure the quality of summaries directly (e.g., by comparison to a manually created *ideal* summary) and extrinsic methods evaluate how well the summaries support a person or system in some predefined task such as searching (Mani, 2001). The latter is more of interest to current work within IIR although intrinsic tasks are used in the annual Document Understanding Workshops (duc.nist.gov). Evaluating what makes an effective summary or surrogate for searching is not trivial; the idea of a good summary depends very much on the role the summary is intended to play within the search process. The authors of manually created summaries or abstracts perhaps had a similar aim—to construct one good, objective summary that would represent the content of an object for all potential (and unknown) readers and with unknown search tasks and personal differences. Such a summary may be sub-optimal for all readers but would be usable by them all.

We could argue that what searchers really want is a surrogate that will help them make decisions about the represented material: Should I read this document? Can I safely ignore it? Is it different from these other retrieved documents? These decisions may be very subjective and personal. As we can automatically create many different representations of the same object, we can potentially create different surrogates at different points in a search and for different purposes. White and Ruthven's (2006) interface for Web searching presents a layered approach to using multiple surrogates. In their interface, shallow sentence surrogates are

used to give an overview of the retrieved set of pages and more detailed surrogates are employed to drill down into searcher-selected parts of the page content. Such an approach is useful in increasing the effectiveness of a search and, as a side effect, in characterizing stages within a search based on the level of use of different surrogates (White, Ruthven, & Jose, 2005).

Surrogates need not always be representations of a single object but can be representations of multiple objects. Multi-document summarization (e.g., Harabagiu & Lacatusu, 2005; McKeown, Passonneau, Elson, Nenkova, & Hirschberg, 2005; Radev et al., 2004) is a popular technique although the evaluation is decidedly non-trivial. McKeown et al. (2005) compared a number of methods for creating summaries; the results indicate the positive effect of high-quality summaries in a task-based evaluation.

Maña-López, DeBuenga, and Gómez-Hidalgo (2004) use a mixture of techniques centered on instance recall tasks: finding as many aspects of a topic as possible. Their approach also combines summaries of multiple document sets and the results show some behavioral improvements—more useful interaction may have taken place because of the combined summarization and clustering. More importantly, they indicate that such a supportive interface, one that structures access to information, helps searchers who have little familiarity with the search topic.

Surrogates for specific collections can be very inventive. For selected books, Amazon offers an intriguing range of surrogates (all of which can be used as search keys) such as a concordance of the 100 most frequently used words in the book, a list of statistically improbable phrases (phrases common to an individual book but uncommon in a collection of books) and capitalized phrases (phrases consisting of capitalized words occurring often within the text). These kinds of surrogates may, or may not, be as useful for searching and assessment as traditional surrogates. However, their presence does add to the sense of fun and engagement with the material being assessed. Norman (2004) points to such emotional appeal as a core factor in the success of search engines such as Google.

Clustering, Categorization, and Browsing

Surrogates not only are useful to help searchers assess individual objects but also to understand relationships among a collection of items or to structure their investigation of an information space. Such approaches typically assist a searcher either by presenting information before searching to aid request creation or after searching to aid interpretation of the results or provide suggestions for search refinement. For both tasks, clustering and categorization are popular approaches. Although the terminology is not always used consistently, categorization typically refers to the manual or automatic allocation of objects to

predefined labels, whereas clustering generally refers to automatic groupings of objects through inter-object similarity.

One of the advantages of offering searchers information on the collection being searched *before* any other interaction takes place is that searchers may find browsing easier than producing search terms (Borgman, Hirsch, Walter, & Gallagher, 1995). Browsing is usually initiated by some information display; services such as Yahoo! (dir.yahoo.com) and Wikipedia (en.wikipedia.org) offer browsable categories as well as freetext searching to inform searchers of the location of additional information. A particularly useful form of categorization is faceted search (Yee, Swearingen, Li, & Hearst, 2003) in which metadata is organized into categories to allow searchers to explore objects interactively and drill down to an area, or set of objects, of interest (Hearst, 2006b). Such faceted approaches also facilitate the creation of complex queries through natural interaction and exploration (Hearst, Elliot, English, Sinha, Swearingen, & Yee, 2002). In situations where the searcher is less certain of the information required or is less informed about the information space, such as the area of exploratory search (White, Kules, Drucker, & Schraefel, 2006), categorization and browsing could be particularly useful to help the searcher structure the investigation.

As summarized by Hearst (2006a), clustering and categorization have advantages and disadvantages. Clustering requires no manual input, it can adapt to any genre of data, and there is a range of good, well-understood algorithms for implementation. Unfortunately, the labels assigned to clusters may not be semantically intuitive and, depending on the algorithm used, the clusters may be badly fragmented. Depending on the nature of the clustering—what objects are being clustered and how many are being clustered—clustering algorithms can also result in interaction delays. Categorization, on the other hand, generally results in superior quality groupings of objects, has clearer and better motivated semantic relationships, and is popular with end users (Hearst, 2006a). A drawback to categorization approaches is the need for an external categorization scheme, although some work has examined automatically creating hierarchies of concepts (Joho, Sanderson, & Beaulieu, 2004).

Clustering approaches can be used to select better sets of objects for presentation to the searcher; work in this area has shown the effectiveness of clustering approaches that use the query as an additional input (e.g., Iwayama, 2000; Tombros, Villa, & van Rijsbergen, 2002) rather than clustering independently of the query. Tombros et al. also report that query-biased clustering approaches can improve retrieval effectiveness over standard inverted file searches (Tombros et al., 2002). However, the interface, in particular the intuitiveness of the information display, is important in maximizing these benefits; a searcher needs to be able to understand the relationships being presented by the clustering to avoid losing the potential benefits of the clustered organization (Wu, Fuller, & Wilkinson, 2001b).

This use of clustering and categorization for displaying search results is also helpful. In particular, clustering for visualization—automatically detecting similarities between objects for graphical representations—is popular. Where the objects being clustered are easy to assess for relevance, such as images or video key-frames, the objects themselves are usually displayed (e.g., Heesch & Rüger, 2004). Where the objects are more complex, clusters will typically have some form of surrogate to label the grouping and aid the searcher's understanding of the grouping (Roussinov & Chen, 2001). Visualizations can be useful but can also create usability problems if there is insufficient support for navigation and searchers' decisions about their own search process (Wiesman, van den Herik, & Hasman, 2004). However, solid work has been done to investigate usability issues in category systems based on a searcher's criteria for using categories (e.g., Hearst, 2006b). Systems such as Kartoo (www.kartoo.com) that integrate multiple interactive features—clustering and visualization, summarization, and query refinement suggestions—may be more robust in supporting the searcher's decision making. Toms (2002) evaluated a novel interface to Google that combined Google's query interface with its directory categories. Their subjects preferred different interaction models for different types of searches (in this case, searches in different domains), with travel or shopping searches favoring the category-based approach and research-style searches favoring querying.

Although clustering can be used for visualization, clusters are more commonly used to facilitate other types of interactive support. For example, both the Wisenut (www.wisenut.com) and Vivisimo (vivisimo.com) Web search engines use clustering approaches to extract and display query refinements. Käki's (2005) Findex system offers categories with which to filter search results; a longitudinal study indicates that searchers will use categories, although not as a default mechanism. However, even if the categories were used only in a minority of searches, the categories could help in more difficult searches—searches where the queries are poor. In Käki's interface, categories are displayed alongside search results; in an earlier study Chen and Dumais (2000) use categories to structure the display of search results. Their study also hints at the utility of categories for more difficult searches; this is more readily apparent in a later study (Dumais, Cutrell, & Chen, 2001).

Visualization

As noted, visualization of information can help the searcher understand relationships among objects or sets of objects. Visualizations can be useful at many different levels. Visual representations of documents, for example, can aid the searcher by graphically representing some information about the content of the document. These representations can be representations of the document such as Hearst's (1995) TileBars, which represent the shift in a document's topical structure; representations

that are relative to the query, such as Reid and Dunlop's (2003) RelevanceLinkBar; or within document representations, that is, visual representations to aid searchers as they read a document (Harper, Koychev, Sun, & Pirie, 2004). Although the mechanisms behind these visual representations are different, they share a common aim of helping searchers find which documents are most likely to be useful, identifying where in a document relevant information may be located, and giving the searcher short cuts to accessing the most useful parts of a document. With highly structured documents, especially those that have an explicitly pre-defined structure, such as plays, even more complex visual representations are possible. Crestani, Vegas, and de la Fuente (2002) present a layered ball metaphor by exploiting a rich collection-specific structure.

Visualization of multiple objects can be useful in representing relationships among objects, grouping together images with similar colors, for example, or linking the content of multiple Web sites through shared concepts as in the case of Kartoo. One particular use of visualization is to help understand similarities and differences among complex objects. Liu, Zhao, and Yi (2002) examined visualization approaches for comparing Web sites. Web sites are complex objects consisting of multiple pages and are difficult to compare using query-driven approaches. Visualizations can help by presenting overviews of Web site content, showing where information is located in a Web site, and presenting comparative views, showing which Web site contains more information on a topic or covers more areas of interest.

How information is visualized is usually decided by the system designer. However, approaches that allow searchers to manipulate and organize information while they are searching can also be useful. Interfaces such as the pile metaphor suggested by Harper and Kelly (2006)—searchers develop piles of documents as they search—help searchers by visualizing what aspects of a search they have covered and how much material they have collected. Buchanan, Blandford, Thimbleby, and Jones (2004) also find spatial displays and metaphors useful when searchers can organize their own search activities and outputs. Such visualizations are not restricted to being passive displays of information, as these interactions with the visualizations can be used to mine useful information about the searcher's interests. Heesch and Rüger (2004) use a searcher's interactions with image visualizations, for example, as a way of neatly gaining information for relevance feedback.

Visualizations may also help searchers remember search interactions. In complex searches involving a series of interactions it may be difficult for searchers to remember what objects they have already seen or where they saw a particular piece of information. Harper and Kelly's pile metaphor allows searchers to organize relevant material as it is encountered. Often, however, a previously viewed document is realized to be useful and the searcher must backtrack to re-find it. Milic-Frayling, Jones, Rodden, Smyth, Blackwell, and Sommerer

(2004) consider interface support for such backtracking on Web search engines. Their system demonstrates how even such an apparently simple concept as going back to a previous page is far from trivial and can benefit from good interface design. Campbell's (1999) path-based browser also supports session-based searching through the use of retrieval paths, which visualize the order in which objects were selected and viewed. This interface was particularly successful in allowing for multiple paths, each representing a different thread in the retrieval session.

As noted, visualizations can be applied to many stages of the interactive retrieval process. As with browsing approaches, the strength of visualization is allowing people to identify useful relationships or relevant information rather than having to recall pertinent keywords as in the querying approaches.

These techniques can be applied to most data sets and used for most retrieval tasks. Recently, new research directions have opened up to deal with novel retrieval tasks and methods of retrieval. In the remainder of this section, I discuss three in detail: support for re-finding information, part-of-object retrieval, and task specific support.

Re-Finding Information

The ability to store so much information electronically on personal computers means that we have to manage the information in such a way as to be able to re-find it later. Our ability to manage our information space constructively and our willingness to devote time to creating useful structures such as folder hierarchies are doubtful (e.g., Whittaker & Sidner, 1996). Hence the attention on how personal re-finding should be supported.

Re-finding personal information is part of personal information management (PIM) and covers the retrieval of information previously stored or accessed by an individual searcher. PIM was covered in detail in an *ARIST* chapter by Jones (2007), so here I summarize some of the features of PIM as they relate to IIR.

Re-retrieval is different from most retrieval scenarios in that what is being retrieved is not new information but information objects one has previously encountered and therefore can be partially recalled. Hence, even though the queries put to PIM systems may appear similar to those put to standard search engines, they describe what one remembers about an information object and are not a description of the information one requires. Features that can be remembered and used for searching may not be features typically supported by standard interfaces, for example, temporal information or information on previous use. Rodden and Wood (2003), in a study on personal image management, show that people can remember many different features of images (context, color, objects, semantic associations), which can be used as query components. Gonçalves and Jorge (2004), in a study based on textual documents rather than images, also indicate the range of features that people can

remember including time, task, and related documents. In their QUILL system, Gonçalves and Jorge (2006) acknowledge such contextual clues by allowing searchers to tell stories (e.g., narratives describing the documents they would like retrieved). Contextual elements in the narratives, such as time or authorship information, were used to trigger retrievals. Apart from being easy to use, the interface increased the length of queries (the narratives) submitted to the system.

Personal information objects are more heavily influenced by their surrounding context than are non-personal objects. The surrounding context can contain elements from the context of the information object's creation (e.g., personal documents), their access (e.g., Web bookmarks), or their use and reuse. Such context can be used to aid retrieval as in the Haystack system (Adar, Kargar, & Stein, 1999). Successfully re-finding an object often depends on a searcher being able to step out of his current task context and remember previous contexts in which he stored or used an object—"What was I doing, or thinking, when I stored that e-mail?" How people think about their personal objects can affect the kind of retrieval support that might be effective.

Boardman and Sasse (2004) note that different personal media have different characteristics that affect how people store and retrieve objects. Bookmarks, for example, are often seen as being less personal than e-mail and as being pointers to information rather than containers of information. E-mails and files and folders, on the other hand, are often seen as more personal. Boardman and Sasse describe how people use different strategies for managing different media types. This raises the question of whether we want different tools for different media or unified interface support for all personal information objects, regardless of media type. Historically, the preference may have been for the former—media-dependent systems—but more recently the trend is toward comprehensive systems that work across all genres of personal information. There is a range of desktop search systems such as Google's (desktop. google.com) or MSN's Desktop Search (toolbar.msn.com). A common and popular theme in these systems is to relieve the searcher of having to remember *where* an item may be stored (Cutrell, Robbins, Dumais, & Sarin, 2006).

Elsweiler, Ruthven, and Jones (2005) describe an interface for retrieval of personal images, which they claim exploits features of human memory to aid the re-retrieval of personal information. Their interface displays *clues* on context, stored as object annotations, to help people remember additional features of the images they want to retrieve and also create better queries. As the searcher interacts with the information objects, the interface prompts the user with clues on previous contexts. The authors claim that the interface can help create so-called "retrieval journeys," one piece of information aiding in the recall of other useful contextually related information.

Dumais et al. (2004) employ the unified systems approach in their system, Stuff I've Seen, which presents a unified index over all information

objects stored on a desktop machine. At the interface, Stuff I've Seen shows contextual information such as file type, date of access, and author. A large-scale, longitudinal evaluation of Stuff I've Seen showed positive results especially for hard-to-find items and vague recollections.

Cutrell et al. (2006) expand the Stuff I've Seen unified data approach but consider more of the user interface issues involved in facilitating access to personal information archives. As with Elsweiler et al., their system, Phlat, exploits the idea that people may remember various attributes of objects such as date, file type, or author information. Supporting searching by these attributes helps make the process more flexible. Phlat also allows for user-defined tags representing searchable concepts, which can be used to filter objects in searching, although at present it allows tagging only of objects already retrieved. Phlat's filtering and tagging allow searchers to create complex queries very simply.

Part-of-Retrieval

Information objects can be complex. Entities such as Web pages may be constructed from more than one component (images, tables, text) whereas objects such as documents, video, or speech frequently have some internal structure: video samples can be deconstructed into component scenes, documents into sections, speech into speakers. The complexity of these objects raises challenges for searchers; documents may be long and contain multiple topics, meaning that the searcher may have to perform extra work to find relevant material contained within them. Surrogates aid searchers in making initial assessments—I may want to investigate an object in more detail and summarization techniques, in particular, permit a quick overview of content but, as documents become longer or more complex, summaries may be less useful.

However, explicit structures within retrievable objects can be utilized to facilitate quick access to complex objects and an alternative to whole object retrieval is to allow the retrieval system to deal with sub-object units, returning parts of objects instead of complete ones. The two most common media types for sub-object retrieval are video and text. Smeaton (2004) has elegantly summarized the retrieval of video components including interactive video browsing and retrieval, so in this section I concentrate on retrieval of document components.

Previously, passage retrieval—retrieving the best window of text (Callan, 1994)—was the most common technique for retrieving parts of documents. Currently, thanks to the INEX initiative, structured document retrieval is receiving more attention (inex.is.informatik.uni-duisburg.de). Structured document retrieval, unlike simple passage retrieval, acknowledges the author-defined structure of the document (sections, subsections, title, etc.) to select and display the best component of the document to the searcher; the best component in INEX is the one that covers as many aspects of the query as possible with the minimum of non-relevant information.

Structured document retrieval raises a number of interesting retrieval and interaction questions. When searching, does the notion of the best component change if the search situation changes? How should an interface relate different components from the same document in an intuitive way for the searcher? Recent research in interactive retrieval from complex objects such as structured documents has followed three approaches: first, visualization approaches such as that of Harper, Muresan, Liu, Koychev, Wettschereck, and Wiratunga (2004), which aim to help searchers assess complex objects, in particular navigating to the most relevant parts of objects; second, employing complex querying languages, which, as noted before, can help searchers specify more precisely how the content and structure of an object should be used to retrieve objects (e.g., Kamps et al., 2005); and third, good interface development can help people interact with complex objects.

Research in this area is heavily influenced by behavioral studies of how people interact with structured documents. Reid, Lalmas, Finesilver, and Hertzum (2006a, 2006b) make a useful distinction between the concepts of most relevant component and the best entry points (BEPs) for accessing documents. Whereas the most relevant component may be the part of the document that contains the most relevant information, BEPs are the best place for a searcher to start investigating an individual document. A searcher may, for example, obtain an answer from a short, relevant section but prefer to be shown the containing section (BEP) to contextualize the information given in the relevant section. Reid et al. (2006b) propose different types of BEPs. A container BEP, for example, is a component that contains at least one relevant object, whereas a context BEP is a component that, although not containing any relevant information itself, provides contextual information for a subsequent relevant object.

Reid et al. (2006a, 2006b) empirically investigate these BEPs and general information search behavior in a number of small studies. Although the BEP types are shown to be not useful, it is clear that searchers themselves distinguish, conceptually and behaviorally, the notion of relevance and BEP. Searchers grasped the difference between relevant objects and the interactive steps necessary to access and recognize these objects. Furthermore, which BEP is seen as being useful depended on the search task, and to an extent, the nature of the data being searched.

This behavioral study of interaction in structured document retrieval continues in the Interactive Track of INEX (Tombros, Malik, & Larsen, 2005), still concentrating more on behavior than interface support. Investigating search behavior provides insights into possibly useful interface designs. For example, once users have investigated one component, there is a tendency to examine components of similar granularity (e.g., section followed by section). Whether this is because of preferred size of component or because some components are more useful at different points in the search is not clear. Knowledge of where components are

located within documents is usefully presented at the interface level, and information on overlapping components from the same document may reduce redundancy in searching. However, as Betsi, Lalmas, and Tombros (2006) acknowledge, searchers often want different forms of interactive support for different reasons—they want relevant sections from different documents, especially if one document cannot completely satisfy the information need, but also want information on how components are linked within an individual document.

Research on novel interfaces within this area have typically exploited the structure of documents either by displaying their tables of contents (Malik, Klas, Führ, Larsen, & Tombros, 2006; Szlavik, Tombros, & Lalmas, 2006) or presenting location information as part of the surrogates in the results lists (Gövert, Führ, Abolhassani, & Großjohann, 2003).

Task-Specific Support

As retrieval tasks become more specialized and better defined, so too do the systems to support these tasks. IR now provides solutions for a range of retrieval problems, not just reference retrieval. Dedicated initiatives such as TREC have enabled the development of specialized retrieval systems and also facilitated work on interfaces for specialist retrieval tasks. As a result, there are systems for question answering, topic detection, topic distillation (selecting a good set of home pages for a given topic), large scale retrieval, and cross-language retrieval.

Specialized retrieval systems fall into two broad groups. First, we have systems that perform a specialized retrieval task. The underlying system is designed to handle particular data (e.g., genomic retrieval or people-finding systems) or the task itself is specialized and involves more than simply retrieval (e.g., topic detection, novelty detection, task distillation). Second, there are systems where the interaction is specialized in some way (e.g., structured document retrieval). Naturally this is a rough categorization, as specialized tasks often require specialized interaction and specialized interaction often results from a nonstandard retrieval model.

Each of these retrieval tasks necessitates a different type of retrieval system and also influences the type of searcher interaction that is appropriate or necessary. Cross-language retrieval interfaces often require support for user querying and document analysis (Hansen & Karlgren, 2005), whereas question answering will require support for contextualizing the typically short answers given by such a system (Lin, Quan, Sinha, Bakshi, Huynh, Katz, et al., 2003). These specialized systems contribute to both research themes—automating search processes by developing systems to help searchers perform specific tasks and improving interaction by developing novel interfaces for such tasks.

Wu, Muresan, McLean, Tang, Wilkinson, Li, et al. (2004) consider topic distillation, that is, sourcing and creating a list of resources that can be assembled into a single page to act as an entry point to a topic. In such a task, Web site home pages will be preferred to low level pages, and relevant information may be split across resources rather than contained within one site (Bhavnani, 2005). Wu et al. tackled the general question of whether a dedicated interface to a topic distillation task would perform better than traditional ranked-list interfaces. The results are inconclusive. Although participants tend to prefer a dedicated search interface, employing a specialized search engine improves search results. In a separate study on question answering, however, Wu, Fuller, and Wilkinson (2001a) show that a specialized interface can increase both searcher performance and preference.

Specialized systems increasingly cater to both information organization and retrieval tasks. Swan and Allan (2000) present an interface for browsing events in news articles. Their timeline-based interface supports discovery activities (what events are covered in the corpus, how long these events have been discussed in the corpus, which are surprising events) and organization (which terms are associated with each event, which events are most important) that may be important to people browsing news events. Smith (2002) tackles retrieval of historical data using maps to help visualize important information and recent work on geographical information retrieval has relied heavily on visualization techniques for searching (Jones & Purves, 2005).

Unfortunately, novel interfaces for specialized tasks often lack a corresponding user evaluation, concentrating only on algorithmic measures of effectiveness. A recurring question remains: What types of novel evaluation metrics are appropriate for new retrieval solutions? Wu et al. exemplify attempts to determine whether interfaces that are designed to fit a specialist task can outperform standard retrieval interfaces. The question is not as trivial as it might seem, as people do learn strategies in order to use familiar interfaces for new tasks. As Turpin and Hersh (2001) point out, people can compensate for sub-optimal retrieval systems by adopting different interaction styles. Similarly, as Muramatsu and Pratt (2001) claim, humans are very adaptable and can operate successfully even with a certain lack of understanding of how retrieval systems actually operate.

Automating Search Processes

The second theme of this chapter, automating search processes, deals with research that attempts to provide technical support for search activities that searchers find either difficult or time-consuming. This involves a wide range of solutions, from traditional approaches such as automatic query reformulation and relevance feedback to newer techniques such as collaborative filtering. Conventional techniques like relevance feedback have been treated in the research literature for decades

and basic questions surrounding their use (how to use relevance information to modify a query, how to encourage searchers to engage in feedback, what relevance evidence is useful, and so on) are still being actively investigated, with the questions changing both subtly and radically as the environment in which their use changes.

Investigating how to deploy relevance feedback successfully in Web search environments has opened up new lines of research on the use of implicit search modeling. The large scale use of Web search engines (and related usage data) has also facilitated work on searcher classification, query intention analysis, and prediction of relevant items. Older research questions are still relevant to new IR environments because, even if contexts change, the problems faced by searchers often do not. Creating queries can be difficult whether we are searching bibliographic databases or the Web. Web search engine query constructors can exhibit the same usability problems as Boolean operators, and assessing relevance is affected by the reasons for the search, irrespective of the medium being searched.

In this section I select four main areas for discussion reflecting the increased attention given to them in the recent literature: implicit feedback, query intention analysis, personalization, and automated assistance.

Implicit Feedback

Explicit approaches to relevance feedback require a searcher to make explicit an assessment of (non-)relevance on retrieved information objects and also to request the system to use such assessments. These approaches rely on obtaining sufficient assessments to generalize a good model of the searcher's underlying information need. The notion of sufficient information, the amount of information required to produce a good generalized model, will generally be topic- and collection-dependent.

Early experiments indicated that standard methods of relevance feedback could perform reasonably well with small numbers of feedback documents; at least, small amounts of relevant information were better than no relevance information (Spärck Jones, 1979). However, the amount of evidence supplied by searchers is still sparse and feedback approaches that have access to multiple examples of relevant information tend to outperform those optimizing very little information (Smucker & Allan, 2006). Without sufficient relevance evidence, the system may make weak query modification decisions resulting in poor effectiveness and—potentially worse from an interactive perspective—low levels of confidence in relevance feedback as a useful technique (Beaulieu & Jones, 1998; Ruthven, 2002).

One method of overcoming the lack of evidence available from traditional explicit approaches is to exploit *implicit* evidence as a substitute for explicit feedback. The use of implicit evidence is not new in itself; however, the approach has gathered momentum due partly to the ease

with which Web browsers, in particular, can be adapted or extended to capture such evidence. Recent research in the use of implicit evidence has centered around three questions: What evidence about a searcher's interaction is useful to know, how reliable is implicit evidence, and what can we do with such evidence to provide better retrieval or better interaction?

Implicit Evidence and Indicators of Interest

Evidence for implicit feedback could potentially be any evidence gained from the system–searcher interaction, including physiological evidence such as heart rates or brain wave tracing. Evidence is commonly restricted to that available from the human–system interaction. Types of evidence and proposed categorizations are summarized by Claypool, Le, Waseda, and Brown (2001); Oard and Kim (2001); and Kelly and Teevan (2003). A common distinction is between direct and indirect evidence.

Direct evidence—such as bookmarking, printing, or saving a document—represents distinct, and usually discrete, actions performed on an object by a searcher, which could be taken as an indication that the object is of interest to the searcher. Click-through behavior could also be considered as direct evidence, depending on whether the system uses the link clicked or the object clicked to as an indicator of interest. Such direct evidence is usually treated as evidence of the searcher's interest in some information object or, at least, evidence that the object is significant in some way to the search or searcher. Most researchers do not go beyond this and equate interest with relevance; the position is usually that implicit evidence tells us something about the potential significance of an object, not necessarily about its relevance to a search. However, whether we can treat implicit interest as synonymous with relevance information (implicit evidence as a substitute for explicit relevance decisions) is an important question and one that has been investigated in a number of ways.

Direct evidence is usually less abundant than indirect evidence but represents deliberate, objectively observed actions performed by a searcher on a system or object. Indirect evidence, on the other hand—such as scrolling behavior, reading time, or repeat visits to a page—is typically continuous evidence that could be interpreted as evidence of interest if it differs from some form of average behavior. That is, the system could make an inference of searcher interest based on differences in the searcher's behavior from normal behavior. For example, long reading time relative to the length of a document, scrolling the length of a document, or repeated visits to the same Web page might imply that the searcher is interested in the content of the page. Equally, these actions could say nothing about searcher interest. Long reading time might imply that the reader is unfamiliar with the content of the document being read, repeated visits might mean that the page is dynamically

updated, and scrolling might mean the searcher is failing to find any useful information within the document.

The available research on good, implicit indicators suggests that direct evidence is usually more reliable; the more plentiful indirect evidence is more nebulous and requires more careful handling, in two ways. Firstly, we need to perform more validation experiments on our hypotheses regarding useful indicators of interest and, secondly, we need to construct additional components in our systems to handle the reasoning about this evidence. However, the sheer quantity of implicit information, especially indirect evidence, is one of the reasons for the attractiveness of implicit feedback.

Implicit evidence can also supply useful information about general searching behavior for investigative purposes. Lorigo, Pan, Hembrooke, Joachims, Granka, and Gay (2006) demonstrate that new methods of collecting information about the search process can uncover important aspects of how people search. Methods such as eye-tracking can uncover patterns of information assessment previously possible only through log analysis or verbal reporting. There are interesting indications that men and women have different assessment strategies, for example. Lorigo et al.'s study, examining how searchers view surrogates when searching on Google, reveals how little information searchers actually view before reformulating a query. Their findings are important in assessing what type of information searchers actually want presented on results pages (based on what they look at) and also on refining our assumptions underpinning traditional algorithms. We should not, for example, assume that pages that have not been visited by searchers are not relevant; the searcher may not even have considered a page or its surrogates.

Joachims, Granka, Pan, Hembrooke, and Gay (2005) use eye-tracking to show that people look at the top-ranked results far more than any other position, prefer visible ("above-the-fold") results and make decisions very quickly on relatively little information. However, the decisions as to which documents to click depend on the relative quality of the overall results. It is suggested, in line with earlier findings (e.g., Florance & Marchionini, 1995) that implicit judgments of relevance are best viewed as relative rather than absolute.

Reliability of Implicit Information

Several studies have probed which indicators of implicit interest are reliable and can be used for predictive algorithms such as relevance feedback or information filtering. Claypool et al. (2001), in an early study on Web browsing, found strong correlations between time taken to read Web pages, scrolling behavior, and explicit interest in a Web page. White, Ruthven, and Jose (2002) report a relationship between reading time and relevance, but Kelly and Belkin (2004) find no correlation.

Kelly and Belkin's longitudinal study was less artificial than the empirical work of Claypool et al. and White et al. and investigated the

general search behavior of a small number of participants. Kelly and Belkin note the importance of the task relative to search behavior: The tasks an individual searcher performs affect his behavior and consequently the interpretations that should be made of the behavior. This raises the general question of whether implicit evidence should be treated as personal—*this* searcher typically behaves in *this* way in the presence of interesting information—or general evidence—*most* people behave *this* way.

Claypool et al. note the importance of task in implicit feedback and also the context in which the evidence appears. Reading time might be more reliable, for example, when differentiating between documents of a similar type or documents from a similar domain but less reliable for documents such as Web pages, which may vary greatly in style and readability. Similarly, relatively coarse evidence such as repeated visits to a resource might be seen as a more reliable indicator of interest or trust in the resource if the searcher has a choice as to which resource to visit.

Whether quantity of evidence can substitute for quality of evidence is not clear, although quality of evidence can perhaps be established through careful analysis of a sufficiently large data set. As has been noted, Joachims et al. (2005), in a direct comparison of explicit and implicit evidence, produced results indicating that implicit evidence can indeed substitute for explicit evidence, if handled as *relative* evidence rather than absolute evidence.

So, although implicit evidence has the potential to be effective, it often does not display the potential benefits, because it either is poor or needs to be contextualized with other information. In particular, implicit evidence can have a signal-to-noise ratio, with little useful information being presented (Shen, Tan, & Zhai, 2005). However, additive context—context from multiple sources added together—appears to be beneficial; Shen et al. (2005) use query histories and click-through data. In a relatively large sample investigation, Fox, Karnawat, Mydland, Dumais, and White (2005) also investigate interaction; they consider interaction only with the results of a search engine rather than whole searching behavior to compare explicit judgments of satisfaction with implicit feedback indicators. They also show that, overall, combinations of implicit evidence perform better than single pieces of evidence. The impression from most work in this area is that simple measures of implicit evidence may not suffice and some combination of evidence will be necessary to make robust predictions of interest (Fox et al., 2005; Shen et al., 2005; Teevan, Dumais, & Horvitz, 2005).

The area is progressing quickly and there is a move to look at the bigger picture of searching by considering more than isolated units of behavior. Fox et al., for example, analyze temporal actions—sequences of actions modeled by state transitions. Such a holistic approach to modeling implicit behavior could provide more useful clues to searcher satisfaction than simply modeling individual pieces of information.

Use of Implicit Feedback

Once we have evidence, either indirect or direct, we can use it to improve either retrieval performance or user interaction. Often this takes the form of query support, using evidence of interest to suggest new query formulations (White & Ruthven, 2006) or re-ranking of search results for presentation of retrieved material to searchers (Teevan et al., 2005). Teevan et al.'s personalized ranking study demonstrates that such approaches can be successful but not for all queries. Agichtein, Brill, and Dumais (2006) find that, for a large set of Web queries, incorporating implicit factors into the original ranking process is more effective than re-ranking the results. Agichtein et al. observe that some queries benefited from implicit feedback whereas others, particularly navigational queries, did not.

White et al. (2002) also employ implicit feedback, based on time-to-read information, to re-rank sentence surrogates. The use of implicit feedback largely failed in this experiment due to usability issues rather than the effectiveness of the feedback. Teevan et al. (2005) make a similar point—the effects of personalization should be interpretable by the searcher and not work against current user strategies.

Implicit evidence can also be used to direct the system response—what is an appropriate system reaction for this searcher at this point in a search (Ruthven et al., 2003)? White and Ruthven (2006), using a newer interface, employ implicit feedback to determine the level of system response based on a system model of the change in the user's information need as reflected in the interaction. Small estimated changes in an information need would result in modest changes, such as re-ranking of search results; large changes in the perceived information need would result in a more radical response such as running a new search.

As part of the study they investigate factors that can affect the utility of implicit versus explicit feedback using a novel Web search interface, specifically the complexity of a search task, the experience of the searcher, and search stage. Implicit feedback performs well when searchers have difficulty deciding on the relevance of individual items but explicit feedback is preferred in simpler search situations (where it is easy to decide on relevance and easy to choose new query terms). More importantly for implicit modeling, the investigation shows different search behaviors with tasks of varying complexity; searchers spend longer on initial browsing of search results before focusing their search. However, the question still remains of how reliably we can move implicit feedback from a descriptive to a predictive tool—at what level can implicit relevance feedback be consistently useful (White et al., 2005)?

Query Intention Analysis

Jansen and Spink (2006) present findings suggesting that low use of advanced search features is part of a long-term trend in Web searching. If this is the case, then a solution to improving retrieval performance

might be to provide some form of support to improve searcher queries automatically.

An approach that is gaining in popularity, especially in Web research, is query intention analysis: ascertaining the searcher's goal behind the query and generating an appropriate response. For example, Broder (2002) proposes three types of Web search: informational (which corresponds to normal ad hoc search and which was the most common in Broder's study), navigational (home page finding to locate a particular site), and transactional service finding (finding a site where the user can accomplish a particular task such as shopping). Depending on how one gathers and analyzes the data, the proportion of searches within each group can vary; Broder's study on query logs estimates approximately 20 percent of searches as navigational, 30 percent transactional, and almost 50 percent as informational. This classification itself is not static—Rose and Levinson (2004) extended and revised Broder's original classification to produce one comprising twelve search goals. If a system could work out which type of search a searcher was engaged in, then it could optimize retrieval for that kind of search.

In a sense, query intention analysis, or identification, is not fundamentally new to IR. Earlier work in the area of user modeling (e.g., Ingwersen, 1992, chapter 7) tended to stress some notion of user need analysis and it has always been known that searchers carry out different types of search and want different types of responses. What is new is the work toward automatic identification of these goals.

Kang and Kim (2003, 2004) concentrate on informational and navigational searches and show that using various types of scoring techniques gives different results for different types of searches. They propose a method, based on a mixture of techniques, for classifying a query as either navigational or informational based on properties of the query terms used and how these terms have been used in Web pages. Lee, Liu, and Cho (2005) investigate navigational and informational searches but concentrate on link information and searcher clicking behavior. Both studies show reasonable success for some types of queries, especially when combinations of evidence are used, but the more difficult informational queries require more effort to detect. The difficult nature of informational queries is also noted in the TREC Web track (Craswell & Hawking, 2005).

As Azzopardi and de Rijke (2006) note, the query is not the only thing we can try to infer; other possible attributes of the search include the expertise of searcher, unit of document desired, or document type. Liu, Yu, and Meng (2006) try to infer which category of information is most appropriate for a searcher's query; Azzopardi and de Rijke try to infer query structure (or fields) with reasonable success although they note that query ambiguity causes significant problems for retrieval performance.

Personalization

Most retrieval systems assume nothing about the people who use the system. That is, each search iteration is analyzed purely on the session-based interaction with no input about the searcher, his history of interaction, or preferences. IIR systems could help improve retrieval effectiveness automatically by personalizing retrieval for individual searchers or tasks.

An awareness of the wider search context is important because it is not only the interface that affects one's ability to conduct a successful search. Attitude to searching, for example, affects interaction and not all people will react to the same tasks in the same way (Ford et al., 2005; Heinström, 2005). However, Savolainen and Kari (2006) note that, although we have many studies on various factors that might affect searching behavior (age, gender, experience, domain expertise), we have fewer tools (methodological or practical) for carrying out these analyses.

In the HARD track of TREC (Allan, 2005) and the follow-on ciqa (complex interactive question answering) track (Kelly & Lin, 2007), a central interest was how personal information about a searcher could help personalize retrieval for that individual. That is, rather than assuming that there was one average result list that would be good for all searchers, HARD and ciqa investigate whether employing information from individual searchers (in this case TREC assessors) could be used to personalize and improve retrieval performance. In different years, the tracks operated with different information that could be used to personalize retrieval, for example, metadata reflecting personal preferences toward types of article or personal information such as level of topic familiarity.

Both HARD and ciqa are unusual in that they allow limited interaction with the TREC assessors through the use of clarification (HARD) or interaction (ciqa) forms to ask for information from the assessors or to ask the assessor to judge information. Various groups (e.g., Belkin, Cole, Gwizdka, Li, Liu, Muresan, et al. [2005]; Tan, Velivelli, Fan, & Zhai, [2005]; Vechtomova & Karamuftuoglu [2005]) use the forms to investigate interactive query expansion approaches. Others tried novel interfaces; for example, Evans, Bennett, Montgomery, Sheftel, Hull, and Shanahan (2004) investigate a clustering approach and Kelly, Dollu, and Fu (2004) simply ask the assessors more about their information needs.

Topic familiarity was an area that sparked interest in many of the participating groups: How would the results of retrieval differ if the searchers, in this case the TREC assessors, had a high or low level of topical knowledge? Here groups propose and evaluate different hypotheses centered on issues such as readability of documents or the degree to which documents contain specialized vocabularies.

The Rutgers group, for example, show some benefit in presenting highly readable documents, as measured by the Flesch Reading Ease score, for assessors with low topical knowledge (Belkin, Chaleva, Cole, Li, Liu, Liu, et al., 2004) whereas the Robert Gordon group find some

benefit in presenting more specific documents to assessors with high topic familiarity (Harper, Muresan, et al., 2004). Researchers at the University of Strathclyde investigate assessor confidence and interest in the topic being searched, as well as topic familiarity, trying different retrieval algorithms for assessors with different characteristics. Few of these personalized techniques work well, although there are indications that some are more effective for individual assessors though not for individual topics (Baillie, Elsweiler, Nicol, Ruthven, Sweeney, Yakici, et al., 2006).

Input to personalized retrieval systems can come from outside the searcher's own interaction. The interactive issues associated with newer systems are not as well defined as in the more traditional information access models, nor have they received sufficient research attention as yet, although the evaluation issues and profiling issues are interesting (see, for example, the special issues edited by Konstan, 2004, and Riedl & Dourish, 2005). However, collaborative or social filtering systems can be effective in mapping a searcher's interaction to that of other searchers to suggest new material. Systems such as Amazon, which recommends new items to customers based on their previous purchases, are the most visible examples; other researchers, for example Boydell and Smyth (2006), have shown collaborative techniques to be effective in increasing retrieval effectiveness by group-based filtering of useful material. Further, collaborative approaches can help identify related communities so that searchers can obtain recommendations from outside their normal group (Freyne & Smyth, 2006).

Automated Assistance

As has been noted, searchers can and do adapt to the search tools provided. However, such adaptation is not guaranteed. Searchers may simply give up using a search tool, and a lack of understanding and support can lead to poor search strategies (Muramatsu & Pratt, 2001). Savolainen and Kari's (2006) study of Web searching indicates that people face more problems in searching than simple use of search engines. We need to examine search behavior in relation to search interfaces. The area of automated assistance—offering search help—is popular in part to compensate for most searchers' lack of training.

What support should be provided in searching and what form this support should take is not yet clear. Jansen (2005) investigates the role of automated assistance in interactive searching, specifically a system that offers assistance in different aspects of search process (e.g., when issuing a query, the system offers thesaural refinements; when bookmarking or printing a document, the system implements relevance feedback based on the bookmarked or printed document and suggests terms). Jansen reports that the presentation of assistance is important; searchers will take automatic assistance and it is better to provide assistance than have it requested. Ruthven (2002) also demonstrates a

preference for assistance that is offered as default rather than requested. There is, of course, a balance to be struck between increased support and cognitive load. The more complex the interaction becomes, the less useful the support. Not all interactive support is equally useful to all searchers, and for some it could become a distraction rather than a help.

Brajnik, Mizzaro, Tasso, and Venuti (2002) employ a rule-based expert system for offering suggestions to searchers on query reformulation and search tactics such as searching by author. The results are similar to Jansen (2005) and Ruthven (2002), in that automated assistance can be popular and effective but also needs to avoid being too generic. Searchers in all studies request very personal and situation-specific assistance rather than just general search advice.

The notion of situation-specificity is not only important to the area of automated assistance—mobile information seeking, for example, depends very much on a good model of the local context in which searches are being made—but automated assistance also needs to be precise enough to be of use. Studies like those of Jansen and Brajnik help benchmark the quality of other solutions.

Discussion

I started this chapter by contrasting two approaches to supporting end-user searching: automating difficult, interactive tasks and improving interactive functionality. In a sense this is a not a clear-cut distinction but two ends of a spectrum. At one end the IIR system assumes little, if any, knowledge of the person on the other side of the interface and the research objective has been to develop interactive functionality to allow the searcher to make better decisions about searching. At the other end there are approaches (such as query intention analysis) where the searcher sees no difference in the interaction and the research effort is focused on the retrieval machinery.

One way to characterize the difference between these two poles is by how extensively they use contextual information. In situations where the system has access to a variety of contextual information (the searcher's previous interactions, preferences, knowledge, etc.), it is best to develop systems that exploit this context to adapt the system's interaction or retrieval results to the needs of individual searchers. The examples presented in this chapter—implicit feedback, automated assistance, personalization—depend, in some sense, on knowledge about the person searching. Query intention analysis may not reveal what a searcher intends but tries to make an informed guess about what the searcher wants.

At the other end of the spectrum we have situations with very little contextual information; here the IIR system tries to augment the searcher's abilities to access information and there is increased interaction. Techniques such as the development of appropriate surrogates or

specially designed interfaces are intended to make the most of the interaction.

Of course, novel approaches can contribute to both areas. Part-of-object retrieval research, for example, may try to automate the task of finding the most useful section of an object but the interfaces that present these sections allow searchers to interact in new and useful ways. It can also work the other way: Speech-driven interfaces avoid the need for typing and result in more convenient querying (Gilbert & Zhong, 2001), but can also make it easier to think about querying, which results in better initial queries (Du & Crestani, 2004).

Integrated solutions—solutions that combine multiple techniques within a single interface—are becoming more prevalent. Maña-López et al. (2004) use text segmentation, clustering, and summarization in their interface. Similarly, there are trends to combine information organization and retrieval techniques. These have the potential to be useful because, as Xie (2002) notes, people often engage in multiple information-seeking strategies. That is, people adopt varied methods of seeking information or interacting with an interface. Instead, they develop strategies to achieve specific goals and base these strategies on the (often low level) interactive functionality of the systems used. Interfaces that offer more flexible methods to create such strategies could provide more room for individual approaches to retrieval and allow the inclusion of personal information seeking strategies, which might otherwise be hampered by rigid interactions.

Integrating multiple searches within a session, or multitasking (Ozmutlu, Ozmutlu, & Spink, 2003; Spink 2004; Spink, Park, Jansen, & Pederson, 2006), is not well supported by the design of search engine interfaces (although Campbell's [1999] Ostensive Browser allowed multiple search threads within a single search session but not distinct subsearches). As Spärck Jones (2005) reminds us, searching is not a discrete interactive activity and we naturally integrate other activities such as spell-checking within a search. Providing more functionality to support decision making and information management activities within searching raises the overall utility of search systems. This functionality itself does not have to be very sophisticated to be useful—the spelling variation feature in Google is a simple, intuitive, and useful feature.

The move from small studies of isolated interactive features to systems that take a more realistic view of how people search is beneficial. A particular theme that has been gaining popularity, and one that has been central to the information seeking literature for some time, is that of task. As elucidated in an earlier *ARIST* chapter (Vakkari, 2002), task is a concept that has many definitions and uses within the information search and retrieval literature. However, we can point to two general aspects of task that are important to IIR: the work task (or the background activity that initiates the need to search) and the task to be fulfilled by the search itself (to answer a question, to gain as much information as possible, to obtain a useful resource, etc.).

Query intention analysis is a particular attempt to understand what the searcher means by a query—what type of response might be most appropriate—but systems that support user tasks suggest systems with a wider consideration of how people search for information. Many authors point to the importance of individual types of interactive support, either for different tasks, different stages within a search, or different search activities (e.g., Allan, Leuski, Swan, & Byrd, 2001; Kim & Allan, 2002; McDonald, Lai, & Tait, 2001; McDonald & Tait, 2003; White & Ruthven, 2006). Systems that offer the wrong support can work against the decisions a searcher must make. Ranked-list approaches to results presentation, for example, do not offer support for searchers trying to understand the structure of an information space, something that is better handled with visualizations (van der Eijk, van Mulligen, Kors, Mons, & van den Berg, 2004).

Classification and understanding of search tasks can help clarify which research directions might bear most fruit and which interactive functionality might best support these tasks. The more general concept of search task complexity is also interesting as such work provides clues about why retrieval systems are not used (Byström & Järvelin, 1995), why they might appear to be less successful for some tasks than others (Bell & Ruthven, 2004), and why certain interactive features might be preferred to others (Fowkes & Beaulieu, 2000). Examining specific types of search and the support required for successful searching appears to be a useful step forward in IIR design.

Allan, Carterette, and Lewis (2005) suggest that difficult search tasks, or at least difficult search topics, are where the most gains could be expected in IIR system performance; the argument being that today's IR systems perform well with simple tasks and we should look at ways of supporting more difficult tasks. This was the basis of the TREC HARD track (Allan, 2005), which investigated search topics where current IR systems performed poorly and where increased interaction might be the only way of improving the search performance. Search tasks that are difficult for IR systems lead to increased interaction. Research such as that of Kim (2006) show increased interaction, especially increased query reformulation, for difficult tasks. However, we could simplify the interaction by developing systems that respond better. This might mean better presentation of information, as in the case of specialist retrieval systems, or incorporation of more personal (searcher) information into the retrieval process.

In this chapter I have tried to represent the areas of IIR activity with the most recent impetus, the balance of discussion being decided by the amount of published activity within the area. The solutions proposed in searching range from complex sets of components to simpler changes in an interface or algorithm; simple changes can make a big difference. Even persuading searchers to examine more search results can increase retrieval effectiveness (White et al., 2003). What resources searchers are searching is also important and the lessons we learn from how people

search can be used to determine the information architecture of these resources (Rosenfeld & Morville, 2002; Toms, 2002). This ability of human searchers to inform and surprise us is one reason to continue studying IIR as a dedicated field.

References

Adar, E., Kargar, D., & Stein, L. A. (1999). Haystack: Per-user information environments. *Proceedings of the 8th International Conference on Information and Knowledge Management*, 413–422.

Agichtein, E., Brill, E., & Dumais, S. (2006). Improving Web search ranking by incorporating user behavior information. *Proceedings of the 29th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval*, 19–26.

Allan, J. (2005). HARD track overview in TREC 2004: High accuracy retrieval from documents. *Proceedings of 13th Text REtrieval Conference*, 24–37.

Allan, J., Carterette, B., & Lewis, J. (2005). When will information retrieval be "good enough"? *Proceedings of the 28th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval*, 433–440.

Allan, J., Leuski, A., Swan, R., & Byrd, D. (2001). Evaluating combinations of ranked lists and visualizations of inter-document similarity. *Information Processing & Management*, *37*, 435–458.

Anick, P. (2003). Using terminological feedback for Web search refinement: A log-based study. *Proceedings of the 26th Annual International ACM Conference on Research and Development in Information Retrieval*, 88–95.

Azzopardi, L., & de Rijke, M. (2006). Query intention acquisition: A case study on automatically inferring structured queries. *Proceedings of the 6th Dutch-Belgian Information Retrieval Workshop*, 3–10.

Baillie, M., Elsweiler, D., Nicol, E., Ruthven, I., Sweeney, S., Yakici, M., et al. (2006). University of Strathclyde at TREC HARD. *Proceedings of the 13th Text REtrieval Conference*. Retrieved January 4, 2007, from trec.nist.gov/pubs/trec14/papers/ustrathclyde.hard.pdf

Barry, C. L., & Schamber, L. (1998). Users' criteria for relevance evaluation: A cross-situational comparison. *Information Processing & Management*, *34*, 219–236.

Bates, M. J. (2002). The cascade of interactions in the digital library interface. *Information Processing & Management*, *38*, 381–400.

Beaulieu, M. (2000). Interaction in information searching and retrieval. *Journal of Documentation*, *56*, 431–439.

Beaulieu, M., & Jones, S. (1998). Interactive searching and interface issues in the Okapi best match probabilistic retrieval system. *Interacting with Computers*, *10*, 237–248.

Belkin, N. J. (1980). Anomalous states of knowledge as a basis for information retrieval. *Canadian Journal of Information Science*, *5*, 133–143.

Belkin, N. J., Chaleva, I., Cole, M., Li, Y.-L., Liu, L., Liu, Y.-H., et al. (2004). Rutgers' HARD Track Experiences at TREC 2004. *Proceedings of the 13th Text REtrieval Conference*. Retrieved January 4, 2007, from trec.nist.gov/pubs/trec13/papers/rutgers-belkin.hard.pdf

Belkin, N. J., Cole, M., Gwizdka, J., Li, Y.-L., Liu, J.-J., Muresan, G., et al. (2005). Rutgers Information Interaction Lab at TREC 2005: Trying HARD.

Proceedings of the 14th Text REtrieval Conference. Retrieved January 4, 2007, from /trec.nist.gov/pubs/trec14/papers/rutgersu.hard.murensan.pdf

Belkin, N. J., Cool, C., Stein, A., & Thiel, U. (1995). Cases, scripts and information-seeking strategies: On the design of interactive information retrieval systems. *Expert Systems with Applications*, *9*, 379–395.

Belkin, N. J., Kelly, D., Kim, G., Kim, J.-Y., Lee, H.-J., Muresan, G., et al. (2003). Query length in interactive information retrieval. *Proceedings of the 26th Annual International ACM Conference on Research and Development in Information Retrieval*, 205–212.

Bell, D. J., & Ruthven, I. (2004). Searchers' assessments of task complexity for Web searching. *Proceedings of the 26th European Conference in Information Retrieval*, 57–71.

Betsi, S., Lalmas, M., & Tombros, A. (2006). XML retrieval: User expectations. *Proceedings of the 29th International ACM SIGIR Conference on Research and Development in Information Retrieval*, 611–612.

Bhavnani, S. K. (2005). Why is it difficult to find comprehensive information? Implications of information scatter for search and design. *Journal of the American Society for Information Science and Technology*, *56*, 989–1003.

Blair, D. C. (2002). The challenge of commercial document retrieval, part 1: Major issues and a framework based on search exhaustivity, determinacy of representation and document collection size. *Information Processing & Management*, *38*, 273–291.

Boardman, R., & Sasse, M. A. (2004). "Stuff goes into the computer and doesn't come out": A cross-tool study of personal information management. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 583–590.

Borgman, C. L. (1996). Why are online catalogs still hard to use? *Journal of the American Society for Information Science*, *47*, 493–503.

Borgman, C. L., Hirsh, S. G., Walter, V. A., & Gallagher, A. L. (1995). Children's searching behavior on browsing and keyword online catalogs: The Science Library Catalog Project. *Journal of the American Society for Information Science*, *46*, 663–684.

Boydell, O., & Smyth, B. (2006). Capturing community search expertise for personalized Web search using snippet-indexes. *Proceedings of the 2006 ACM CIKM International Conference on Information and Knowledge Management*, 277–286.

Brajnik, G., Mizzaro, S., Tasso, C., & Venuti, F. (2002). Strategic help in user interfaces for information retrieval. *Journal of the American Society for Information Science and Technology*, *53*, 343–358.

Broder, A. (2002). A taxonomy of Web search. *SIGIR Forum*, *36*(2), 3–10.

Bruza, P., McArthur, R., & Dennis, S. (2000). Interactive Internet search: Keyword, directory and query reformulation mechanisms compared. *Proceedings of the 23rd Annual International ACM SIGIR Conference on Research and Development in Information Retrieval*, 280–287.

Buchanan, G., Blandford, A., Thimbleby, H., & Jones, M. (2004). Integrating information seeking and structuring: Exploring the role of spatial hypertext in a digital library. *Proceedings of Hypertext 2004, 15th Annual Conference on Hypertext and Hypermedia*, 225–234.

Buyukkokten, O., Kaljuvee, O., Garcia-Molina, H., Paepcke, A., & Winograd, T. (2002). Efficient Web browsing on handheld devices using page and form summarization. *ACM Transactions on Information Systems*, *20*, 82–115.

Byström, K., & Järvelin, K. (1995). Task complexity affects information seeking and use. *Information Processing & Management*, *31*, 191–213.

Callan, J. P. (1994). Passage level evidence in document retrieval. *Proceedings of the 17th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval*, 302–310.

Campbell, I. (1999). Interactive evaluation of the Ostensive Model, using a new test-collection of images with multiple relevance assessments. *Journal of Information Retrieval*, *2*, 89–114.

Case, D. (2006). Information seeking. *Annual Review of Information Science and Technology*, *40*, 293–327.

Chen, C., Czerwinski, M., & Macredie, R. D. (2000). Individual differences in virtual environments: Introduction and overview. *Journal of the American Society for Information Science*, *51*, 499–507.

Chen, H., & Dumais, S. (2000). Bringing order to the Web: Automatically categorizing search results. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 145–152.

Chinenyanga, T. T., & Kushmerick, N. (2001). Expressive retrieval from XML documents. *Proceedings of the 24th Annual international ACM SIGIR Conference on Research and Development in Information Retrieval*, 163–171.

Chuang, W. T., & Yang, J. (2000). Extracting sentence segments for text summarization: A machine learning approach. *Proceedings of the 23rd Annual International ACM SIGIR Conference on Research and Development in Information Retrieval*, 152–159.

Claypool, M., Le, P., Waseda, M., & Brown, D. (2001). Implicit interest indicators. *Proceedings of the 6th International Conference on Intelligent User Interfaces*, 33–40.

Cool, C., Park, S., Belkin, N. J., Koenemann, J., & Ng, K. B. (1996). Information seeking behavior in new searching environment. *Proceedings of the 2nd International Conference on Conceptions of Library and Information Science*, 403–416.

Cove, J. F., & Walsh, B. C. (1988). Online text retrieval via browsing. *Information Processing & Management*, *24*, 31–37.

Craswell, N., & Hawking, D. (2005). Overview of the TREC 2004 Web Track. *Proceedings of the Thirteenth Text REtrieval Conference*. Retrieved January 4, 2007, from trec.nist.gov/pubs/trec13/papers/WEB.OVERVIEW.pdf

Crestani, F., Vegas, J., & de la Fuente, P. (2002). A graphical user interface for the retrieval of hierarchically structured documents. *Information Processing & Management*, *40*, 269–289.

Crouch, C. J., Crouch, D. B., Chen, Q., & Holtz, S. J. (2002). Improving the retrieval effectiveness of very short queries. *Information Processing & Management*, *38*, 1–36.

Cutrell, E., Robbins, D. C., Dumais, S. T., & Sarin, R. (2006). Fast, flexible filtering with Phlat: Personal search and organization made easy. *Proceedings of the ACM SIGCHI Conference on Human Factors in Computing Systems*, 261–270.

Dennis, S., McArthur, R., & Bruza, P. D. (1998). Searching the World Wide Web made easy? The cognitive load imposed by query refinement mechanisms. *Proceedings of the 3rd Australian Document Computing Symposium*, 65–71.

Downie, J. S. (2004). A sample of music information retrieval approaches. *Journal of the American Society for Information Science and Technology*, *55*, 1033–1116.

Du, H., & Crestani, F. (2004). Retrieval effectiveness of written and spoken queries: An experimental evaluation. *Proceedings of 6th International Conference on Flexible Query Answering Systems*, 376–389.

Dumais, S. T., & Belkin, N. J. (2005). The Interactive TREC Track: Putting the user into search. In E. Voorhees & D. Harman (Eds.), *TREC: Experiment and evaluation in information retrieval* (pp. 123–152). Boston: MIT Press.

Dumais, S., Cutrell, E., Cadiz, J. J., Jancke, G., Sarin, R., & Robbins, D. C. (2004). Stuff I've Seen: A system for personal information retrieval and re-use. *Proceedings of the 27th Annual International ACM Conference on Research and Development in Information Retrieval*, 72–79.

Dumais, S., Cutrell, E., & Chen, H. (2001). Optimizing search by showing results in context. *Proceedings of the ACM SIGCHI Conference on Human Factors in Computing Systems*, 277–283.

Dziadosz, S., & Chandrasekar, R. (2002). Do thumbnail previews help users make better relevance decisions about Web search results? *Proceedings of the 25th Annual International ACM Conference on Research and Development in Information Retrieval*, 365–366.

Eastman, C. M., & Jansen, B. J. (2003). Coverage, relevance and ranking: The impact of query operators on Web search engine results. *ACM Transactions on Information Systems*, *21*, 383–411.

Efthimiadis, E. N. (2000). Interactive query expansion: A user-based evaluation in a relevance feedback environment. *Journal of the American Society for Information Science and Technology*, *51*, 989–1003.

Ellis, D. (1989). A behavioural approach to information retrieval system design. *Journal of Documentation*, *45*, 171–212.

Ellis, D., Cox, D., & Hall, K. (1993). A comparison of the information seeking patterns of researchers in the physical and social sciences. *Journal of Documentation*, *49*, 356–359.

Elsweiler, D., Ruthven, I., & Jones, C. (2005). Dealing with fragmented recollection of context in information management. *Context-Based Information Retrieval: Workshop in Fifth International and Interdisciplinary Conference on Modeling and Using Context*. Retrieved January 5, 2007, from ftp.informatik.rwth-aachen.de/Publications/CEUR-WS/Vol-151/CIR-05_4.pdf

Evans, D. A., Bennett, J., Montgomery, J., Sheftel, V., Hull, D. A., & Shanahan, J. G., (2004). TREC 2004 HARD Track experiments in clustering. *Proceedings of the 13th Text REtrieval Conference*. Retrieved January 4, 2007, from trec.nist.gov/pubs/trec13/papers/clairvoyance.hard.pdf

Fidel, R., & Pejtersen, A. M. (2004). From information behaviour research to the design of information systems: The Cognitive Work Analysis framework. *Information Research*, *10*. Retrieved January 5, 2007, from InformationR.net/ir/10-1/paper210.html

Florance, V., & Marchionini, G. (1995). Information processing in the context of medical care. *Proceedings of the 18th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval*, 158–163.

Ford, N., & Ford, R. (1993). Towards a cognitive theory of information accessing: An empirical study. *Information Processing & Management*, *29*, 569–585.

Ford, N., Miller, D., & Moss, N. (2005). Web search strategies and human individual differences: A combined analysis. *Journal of the American Society for Information Science and Technology*, *56*, 757–764.

Foster, A., & Ford, N. (2003). Serendipity and information seeking: An empirical study. *Journal of Documentation*, *59*, 321–340.

Fowkes, H., & Beaulieu, M. (2000). Interactive searching behaviour: Okapi experiment for TREC8. *Proceedings of the IRSG 2000 Colloquium on Information Retrieval Research*, 47–56.

Fox, S., Karnawat, K., Mydland, M., Dumais, S., & White, T. (2005). Evaluating implicit measures to improve Web search. *ACM Transactions on Information Systems*, *23*, 147–168.

Freyne, J., & Smyth, B. (2006). Further experiments in case-based collaborative Web search. *Proceedings of the 8th European Conference on Case-Based Reasoning*, 256–270.

Führ, N., & Großjohann, K. (2001). XIRQL: A query language for information retrieval in XML documents. *Proceedings of the 24th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval*, 172–180.

Gilbert, J. E., & Zhong, Y. (2001). Speech user interfaces for information retrieval. *Proceedings of the 10th International Conference on Information and Knowledge Management*, 77–82.

Gonçalves, D., & Jorge, J. A. (2004). "Tell me a story": Issues on the design of document retrieval systems. *Proceedings of Engineering Human Computer Interaction and Interactive Systems: Joint Working Conferences* (Lecture Notes in Computer Science), 129–145.

Gonçalves, D., & Jorge, J. A. (2006). Evaluating stories in narrative-based interfaces. *Proceedings of the 11th International Conference on Intelligent User Interfaces*, 273–275.

Gövert, N., Führ, N., Abolhassani, M., & Großjohann, K. (2003). Content-oriented retrieval with HyREX. *Proceedings of the 1st Initiative for the Evaluation of XML Retrieval Workshop*, 26–32.

Hansen, P., & Karlgren, J. (2005). Effects of foreign language and task scenario on relevance assessment. *Journal of Documentation*, *61*, 623–639.

Harabagiu, S., & Lacatusu, F. (2005). Topic themes for multi-document summarization. *Proceedings of the 28th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval*, 202–209.

Harper, D. J., & Kelly, D. (2006). Contextual relevance feedback. *Proceedings of the 1st International Conference on Information Interaction in Context*, 129–137.

Harper, D. J., Koychev, I., Sun, Y., & Pirie, I. (2004). Within document retrieval: A user-centred evaluation of relevance profiling. *Information Retrieval*, 7, 265–290.

Harper, D. J., Muresan, G., Liu, B., Koychev, I., Wettschereck, D., & Wiratunga, N. (2004). The Robert Gordon University's HARD Track experiments at TREC 2004. *Proceedings of the 13th Text REtrieval Conference*. Retrieved January 4, 2007, from trec.nist.gov/pubs/trec13/papers/rutgers-belkin.hard.pdf

He, D., & Demner-Fushman, D. (2004). HARD experiment at Maryland: From need negotiation to automated HARD process. *Proceedings of the 12th Text REtrieval Conference*, 707–714.

Hearst, M. (1995). TileBars: Visualization of term distribution information in full text information access. *Proceedings of the ACM SIGCHI Conference on Human Factors in Computing Systems*, 59–66.

Hearst, M. (2000). User interfaces and visualization. In R. Baeza-Yates & B. Ribeiro-Neto (Eds.), *Modern information retrieval* (pp. 257–324). New York: Addison-Wesley Longman.

Hearst, M. (2006a). Clustering versus faceted categories for information exploration. *Communications of the ACM*, *49*(4), 59–61.

Hearst, M. (2006b). Design recommendations for hierarchical faceted search interfaces. *ACM SIGIR Workshop on Faceted Search*. Retrieved January 5, 2007, from flamenco.berkeley.edu/papers/faceted-workshop06.pdf

Hearst, M., Elliot, A., English, J., Sinha, R., Swearingen, K., & Yee, K.-P. (2002). Finding the flow in Web site search. *Communications of the ACM*, *45*(9), 42–49.

Heesch, D., & Rüger, S. (2004). Three interfaces for content-based access to image collections. *Proceedings of International Conference on Image and Video Retrieval* (Lecture Notes in Computer Science), 491–499.

Heinström, J. (2005). Fast surfing, broad scanning and deep diving: The influence of personality and study approach on students' information-seeking behaviour. *Journal of Documentation*, *61*, 228–247.

Hill, B. (2004). *Google for dummies*. Indianapolis, IN: Hungry Minds.

Ingwersen, P. (1992). *Information retrieval interaction*. London: Taylor Graham.

Ingwersen, P., & Järvelin, K. (2005). *The turn: Integration of information seeking and retrieval in context*. Dordrecht, The Netherlands: Springer.

Iwayama, M. (2000). Relevance feedback with a small number of relevance judgements: Incremental relevance feedback vs. document clustering. *Proceedings of the 23rd Annual International ACM SIGIR Conference on Research and Development in Information Retrieval*, 10–16.

Järvelin, K., & Kekäläinen, J. (2000). IR evaluation methods for retrieving highly relevant documents. *Proceedings of the 23rd Annual International ACM SIGIR Conference on Research and Development in Information Retrieval*, 41–48.

Jansen, B. J. (2005). Seeking and implementing automated assistance during the search process. *Information Processing & Management*, *41*, 909–928.

Jansen, B. J., & Spink, A. (2006). How are we searching the World Wide Web? A comparison of nine search engine transaction logs. *Information Processing & Management*, *42*, 248–263.

Joachims, T., Granka, L., Pan, B., Hembrooke, H., & Gay, G. (2005). Accurately interpreting clickthrough data as implicit feedback. *Proceedings of the 28th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval*, 154–161.

Joho, H., Sanderson, M., & Beaulieu, M. (2004). A study of user interaction with a concept-based interactive query expansion support tool. *Proceedings of the 26th European Conference in Information Retrieval*, 42–56.

Jones, C. B., & Purves, R. (2005). GIR '05 ACM workshop on geographical information retrieval. *SIGIR Forum*, *40*(1), 34–37.

Jones, R., Rey, B., Madani, O., & Greiner, W. (2006). Generating query substitutions. *Proceedings of the 14th World Wide Web Conference*, 387–396.

Jones, W. (2007). Personal information management. *Annual Review of Information Science and Technology*, *41*, 453–504.

Käki, M. (2005). Findex: Search result categories help users when document ranking fails. *Proceedings of the Conference on Human Factors in Computing Systems*, 131–140.

Kamps, J., Marx, M., de Rijke, M., & Sigurbjörnsson, B. (2005). Structured queries in XML retrieval. *Proceedings of the 14th ACM International Conference on Information and Knowledge Management*, 4–11.

Kang, I.-H., & Kim, G. (2003). Query type classification for Web document retrieval. *Proceedings of the 26th Annual International ACM Conference on Research and Development in Information Retrieval*, 64–71.

Kang, I.-H., & Kim, G. (2004). Integration of multiple evidences based on a query type for Web search. *Information Processing & Management*, *40*, 459–478.

Kelly, D., & Belkin, N. J. (2004). Display time as implicit feedback: Understanding task effects. *Proceedings of the Annual International ACM Conference on Research and Development in Information Retrieval*, 377–384.

Kelly, D., Deepak, V., & Fu, X. (2005). The loquacious user: A document-independent source of terms for query expansion. *Proceedings of the 28th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval*, 457–464.

Kelly, D., Dollu, V. D., & Fu, X. (2004). University of North Carolina's HARD Track experiments at TREC 2004. *Proceedings of the 13th Text REtrieval Conference*. Retrieved January 4, 2007, from trec.nist.gov/pubs/trec13/papers/unorthcarolina.hard.pdf

Kelly, D., & Lin, J. (2007). Overview of the TREC 2006 ci QA task. *SIGIR Forum*, *41*(1), 107–116.

Kelly, D., & Teevan, J. (2003). Implicit feedback for inferring user preference: A bibliography. *SIGIR Forum*, *37*(2), 18–28.

Kim, J. (2006). Task difficulty as a predictor and indicator of Web searching interaction. *Proceedings of the Conference on Human Factors in Computing Systems*, 959–964.

Kim, K.-S., & Allan, B. (2002). Cognitive and task influences on Web searching behavior. *Journal of the American Society for Information Science and Technology*, *43*, 109–119.

Koenemann, J., & Belkin, N. J. (1996). A case for interaction: A study of interactive information retrieval behavior and effectiveness. *Proceedings of the ACM SIGCHI Conference on Human Factors in Computing Systems*, 205–212.

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

Kraft, R., Chang, C. C., Maghoul, F., & Kumar, R. (2006). Searching with context. *Proceedings of the 14th World Wide Web Conference*, 477–486.

Kruschwitz, U., & Al-Bakour, H. (2005). Users want more sophisticated search assistants: Results of a task-based evaluation. *Journal of the American Society for Information Science and Technology*, *56*, 1377–1393.

Kuhlthau, C. C. (1991). Inside the search process: Information seeking from the user's perspective. *Journal of the American Society for Information Science*, *42*, 361–371.

Lee, U., Liu, Z., & Cho, J. (2005). Automatic identification of user goals in Web search. *Proceedings of the 13th World Wide Web Conference*, 391–400.

Legg, C. (2007). Ontologies on the semantic Web. *Annual Review of Information Science and Technology*, *41*, 407–451.

Lin, J., Quan, D., Sinha, V., Bakshi, K., Huynh, D., Katz, B., et al. (2003). What makes a good answer? The role of context in question answering. *Proceedings of the 9th International Federation for Information Processing TC13 International Conference on Human–Computer Interaction*, 25–32.

Lin, S-j. (2005). Internetworking of factors affecting successive searches over multiple episodes. *Journal of the American Society for Information Science and Technology*, *56*(4), 416–436.

Lin, S-j., & Belkin, N. J. (2005). Validation of a model of information seeking over multiple search sessions. *Journal of the American Society for Information Science and Technology*, *56*(4), 393–415.

Liu, B., Zhao, K., & Yi, L. (2002). Visualizing Web site comparisons. *Proceedings of the 11th Annual WWW Conference*, 693–703.

Liu, F., Yu, C., & Meng, W. (2006). Personalized Web search by mapping user queries to categories. *Proceedings of the 11th ACM International Conference on Information and Knowledge Management*, 558–565.

Liu, H., Lieberman, H., & Selker, T. (2002). GOOSE: A goal-oriented search engine with common sense. *Proceedings of the 2nd International Conference on Adaptive Hypermedia and Adaptive Web-Based Systems*, 253–263.

López-Ostenero, F., Gonzalo, J., & Verdejo, F. (2005). Noun phrases as building blocks for cross-language search assistance. *Information Processing & Management*, *41*, 549–568.

Lorigo, L., Pan, B., Hembrooke, H., Joachims, T., Granka, L. & Gay, G. (2006). The influence of task and gender on search and evaluation behavior using Google. *Information Processing & Management*, *42*, 1123–1131.

Lynam, T. R., Buckley, C., Clark, C. L. A., & Cormack, G. V. (2004). A multi-system analysis of document and term selection for blind feedback. *Proceedings of the 13th ACM International Conference on Information and Knowledge Management*, 261–269.

Lynch, C. A. (2001). When documents deceive: Trust and provenance as new factors for information retrieval in a tangled Web. *Journal of the American Society for Information Science and Technology*, *52*, 12–17.

Malik, S., Klas, C.-P., Führ, N., Larsen, B., & Tombros, A. (2006). Designing a user interface for interactive retrieval of structured documents: Lessons learned from the INEX Interactive Track. *Proceedings of the 10th European Conference on Digital Libraries*, 75–86.

Maña-López, M. J., De Buenaga, M., & Gómez-Hidalgo, J. M. (2004). Multidocument summarization: An added value to clustering in interactive retrieval. *ACM Transactions on Information Systems*, *22*, 215–241.

Mani, I. (2001). Recent developments in text summarization. *Proceedings of the 10th International Conference on Information and Knowledge Management*, 529–531.

McDonald, S., Lai, T., & Tait, J. (2001). Evaluating a content based image retrieval system. *Proceedings of the 24th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval*, 232–240.

McDonald, S., & Tait, J. (2003). Search strategies in content-based image retrieval. *Proceedings of the 26th Annual International ACM Conference on Research and Development in Information Retrieval*, 80–87.

McKeown, K., Passonneau, R. J., Elson, D. K., Nenkova, A., & Hirschberg, J. (2005). Do summaries help? A task-based evaluation of multi-document summarization. *Proceedings of the 28th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval*, 210–217.

Milic-Frayling, N., Jones, R., Rodden, K., Smyth, G., Blackwell, A., & Sommerer, R. (2004). SmartBack: Supporting users in back navigation. *Proceedings of the 13th Annual WWW Conference*, 63–71.

Muramatsu, J., & Pratt, W. (2001). Transparent queries: Investigating users' mental models of search engines. *Proceedings of the 24th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval*, 217–224.

Navigli, R., & Velardi, P. (2003). An analysis of ontology-based query expansion strategies. *Workshop on Adaptive Text Extraction and Mining*, 42–49.

Niemi, T., Junkkari, M., Järvelin, K., & Viita, S. (2004). Advanced query language for manipulating complex entities. *Information Processing & Management*, *40*, 869–889.

Norman, D. A. (2004). *Emotional design: Why we love (or hate) everyday things*. New York: Basic Books.

Oard, D., & Kim, J. (2001). Modeling information content using observable behavior. *Proceedings of the 64th Annual Meeting of the American Society for Information Science and Technology*, 38–45.

O'Keefe, R. A., & Trotman, A. (2004). The simplest query language that could possibly work. *Proceedings of the 2nd INEX Workshop*, 167–174.

Over, P. (2001). The TREC interactive track: An annotated bibliography. *Information Processing & Management*, 37, 369–381.

Ozmutlu, S., Ozmutlu, H. C., & Spink, A. (2003). Multitasking Web searching: Implications for design. *Proceedings of the 66th Annual Meeting of the American Society for Information Science and Technology*, 416–421.

Pirkola, A., Puolamäki, D., & Järvelin, K. (2003). Applying query structuring in cross-language retrieval. *Information Processing & Management*, *39*, 391–402.

Pu, H-T., Chuang, S.-L., & Yang, C. (2002). Subject categorization of query terms for exploring Web users' search interests. *Journal of the American Society for Information Science and Technology*, *53*, 617–630.

Radev, D. R., Jing, H., Styś, M., & Tam, D. (2004). Centroid-based summarization of multiple documents. *Information Processing & Management*, *40*, 919–938.

Reid, J., & Dunlop, M. D. (2003). Evaluation of a prototype interface for structured document retrieval. *Proceedings of the 17th Annual Human–Computer Interaction Conference*, 73–86.

Reid, J., Lalmas, M., Finesilver, K., & Hertzum, M. (2006a). Best entry points for structured document retrieval, part I: Characteristics. *Information Processing & Management*, *42*, 74–88.

Reid, J., Lalmas, M., Finesilver, K., & Hertzum, M. (2006b). Best entry points for structured document retrieval, part II: Characteristics. *Information Processing & Management*, *42*, 89–105.

Riedl, J., & Dourish, P. (2005). Introduction to the special section on recommender systems. *ACM Transactions on Computer–Human Interaction*, *12*, 371–373.

Rieh, S. Y., & Xie, H. (2006). Analysis of multiple query reformulations on the Web: The interactive information retrieval context. *Information Processing & Management*, *42*, 751–768.

Robertson, S., & Callan, J. (2005). Routing and filtering. In E. M. Voorhees & D. K. Harman (Eds.), *TREC: Experiments and evaluation in information retrieval* (pp. 99–121). Boston: MIT Press.

Rodden, K., & Wood, K. R. (2003). How do people manage their digital photographs? *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 409–416.

Rose, D. E., & Levinson, D. (2004). Understanding user goals in Web search. *Proceedings of the 12th World Wide Web Conference*, 13–19.

Rosenfeld, L., & Morville, P. (2002). *Information architecture for the World Wide Web: Designing large-scale Web sites*. Sebastopol, CA: O'Reilly.

Roussinov, D. G., & Chen, H. (2001). Information navigation on the Web by clustering and summarizing query results. *Information Processing & Management*, *37*, 789–816.

Ruthven, I. (2002). On the use of explanations as a mediating device for relevance feedback. *Proceedings of the 6th European Conference on Digital Libraries*, 338–345.

Ruthven, I. (2003). Re-examining the potential effectiveness of interactive query expansion. *Proceedings of the 26th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval*, 213–220.

Ruthven, I., Baillie, M., & Elsweiler, D. (in press). The relative effects of knowledge, interest and confidence in assessing relevance. *Journal of Documentation*.

Ruthven, I., & Lalmas, M. (2003). A survey on the use of relevance feedback for information access systems. *Knowledge Engineering Review*, *18*, 95–145.

Ruthven, I., Lalmas, M., & van Rijsbergen, C. J. (2003). Incorporating user search behavior into relevance feedback. *Journal of the American Society for Information Science and Technology*, *54*, 528–548.

Ruthven, I., Tombros, A., & Jose, J. M. (2001). A study on the use of summaries and summary-based query expansion for a question-answering task. *Proceedings of the 2nd European Conference on Information Retrieval*, 1–14.

Savolainen, R., & Kari, J. (2006). Facing and bridging gaps in Web searching. *Information Processing & Management*, *42*, 519–537.

Shen, X., Tan, B., & Zhai, C. (2005). Context sensitive information retrieval using implicit feedback. *Proceedings of the 28th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval*, 43–50.

Shipman, F. M., Furuta, R., Brenner, D., Chung, C.-C., & Hsieh, H.-w. (2000). Guided paths through Web-based collections: Design, experiences, and adaptations. *Journal of the American Society for Information Science*, *51*, 260–272.

Shneiderman, B., Byrd, D., & Croft, W. B. (1998). Sorting out searching: A user-interface framework for text searches. *Communications of the ACM*, *41*(4), 95–98.

Sihvonen, A., & Vakkari, P. (2004). Subject knowledge improves interactive query expansion assisted by a thesaurus. *Journal of Documentation*, *60*, 673–690.

Slone, D. J. (2002). The influence of mental models and goals on search patterns during Web interaction. *Journal of the American Society for Information Science and Technology*, *53*, 1152–1169.

Smeaton, A. (2004). Indexing, browsing and searching of digital video. *Annual Review of Information Science and Technology*, *38*, 371–407.

Smith, D. A. (2002). Detecting and browsing events in unstructured text. *Proceedings of the 25th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval*, 73–80.

Smucker, M., & Allan, J. (2006). Find-Similar: Similarity browsing as a search tool. *Proceedings of the 29th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval*, 461–468.

Spärck Jones, K. (1979). Search term relevance weighting given little relevance information. *Journal of Documentation*, *35*, 30–48.

Spärck Jones, K. (2005). Epilogue: Metareflections on TREC. In E. M. Voorhees & D. K. Harman (Eds.), *TREC: Experiment and evaluation in information retrieval* (pp. 421–448). Boston: MIT Press.

Spink, A. (1996). Multiple search sessions model of end-user behavior: An exploratory study. *Journal of the American Society for Information Science*, *47*, 603–609.

Spink, A. (2004). Multi-tasking information behavior and information task switching: An exploratory study. *Journal of Documentation*, *60*, 336–351.

Spink, A., & Cole, C. (2005a). *New directions in cognitive information retrieval*. Dordrecht, The Netherlands: Springer.

Spink, A., & Cole, C. (2005b). *New directions in human information behavior*. Dordrecht, The Netherlands: Springer.

Spink, A., & Jansen, B. J. (2004). A study of Web search trends. *Webology*, *1*(2). Retrieved December 22, 2006, from www.webology.ir/2004/v1n2/a4.html

Spink, A., Park, M., Jansen, B. J., & Pederson, J. (2006). Multitasking during Web search sessions. *Information Processing & Management*, *42*, 264–275.

Swan, R., & Allan, J. (2000). Automatic generation of overview timelines. *Proceedings of the 23rd Annual International ACM SIGIR Conference on Research and Development in Information Retrieval*, 49–56.

Sweeney, S., & Crestani, F. (2006). Effective search results summary size and device screen size: Is there a relationship? *Information Processing & Management*, *42*, 1056–1074.

Szlavik, Z., Tombros, A., & Lalmas, M. (2006). Investigating the use of summarisation for interactive XML retrieval. *Proceedings of the 21st ACM Symposium on Applied Computing, Information Access and Retrieval Track*, 1068–1072.

Tan, B., Velivelli, A., Fan, H., & Zhai, C. (2005). Interactive construction of query language models: UIUC TREC 2005 HARD Track experiments. *Proceedings of the 14th Text REtrieval Conference*. Retrieved January 4, 2007, from trec.nist.gov/pubs/trec14/papers/uillinois-uc.hard.pdf

Teevan, J., Dumais, S. T., & Horvitz, E. (2005). Personalizing search via automated analysis of interests and activities. *Proceedings of the 28th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval*, 449–456.

Tombros, A., Malik, S., & Larsen, B. (2005). Report on the INEX 2004 interactive track. *SIGIR Forum*, *39*(1), 43–49.

Tombros, A., & Sanderson, M. (1998). Advantages of query biased summaries in information retrieval. *Proceedings of the 21st Annual International ACM SIGIR Conference on Research and Development in Information Retrieval*, 2–10.

Tombros, A., Villa, R., & van Rijsbergen, C. J. (2002). The effectiveness of query-specific hierarchic clustering in information retrieval. *Information Processing & Management*, *38*, 559–582.

Toms, E. G. (2002). Information interaction: Providing a framework for information architecture. *Journal of the American Society for Information Science and Technology*, *53*, 855–862.

Topi, H., & Lucas., W. (2005a). Mix and match: Combining terms and operators for successful Web searches. *Information Processing & Management*, *41*, 801–817.

Topi, H., & Lucas, W. (2005b). Searching the Web: Operator assistance required. *Information Processing & Management*, *41*, 383–403.

Turpin, A. H., & Hersh, W. (2001). Why batch and user evaluations do not give the same results. *Proceedings of the 24th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval*, 225–231.

Vakkari, P. (2001). A theory of the task-based information retrieval process: A summary and generalisation of a longitudinal study. *Journal of Documentation*, *57*, 44–60.

Vakkari, P. (2002). Task-based information searching. *Annual Review of Information Science and Technology*, *37*, 413–464.

Vakkari, P., & Hakala, N. (2000). Changes in relevance criteria and problem stages in task performance. *Journal of Documentation*, *56*, 540–562.

Vakkari, P., Pennanen, M., & Serola, S. (2003). Changes of search terms and tactics while writing a research proposal: A longitudinal case study. *Information Processing & Management*, *39,* 445–463.

van der Eijk, C. C., van Mulligen, E. M., Kors, J. A., Mons, B., & van den Berg, J. (2004). Constructing an associative concept space for literature-based discovery. *Journal of the American Society for Information Science and Technology*, *55*, 436–444.

Vechtomova, O., & Karamuftuoglu, M. (2005). Experiments for HARD and Enterprise Tracks. *Proceedings of the 14th Text REtrieval Conference*. Retrieved January 4, 2007, from trec.nist.gov/pubs/trec14/papers/uwaterloo-vechtomova.hard.ent.pdf

Vechtomova, O., & Karamuftuoglu, M. (2006). Elicitation and use of relevance feedback information. *Information Processing & Management*, *42*, 191–206.

Voorhees, E. M. (2000). Variations in relevance judgments and the measurement of retrieval effectiveness. *Information Processing & Management*, *36*, 697–778.

White, R. W., Jose, J. M., & Ruthven, I. (2003). A task-oriented study on the influencing effects of query-biased summarisation in Web searching. *Information Processing & Management*, *39*, 707–733.

White, R. W., Kules, B., Drucker, S. M., & Schraefel, M. C. (2006). Supporting exploratory search. *Communications of the ACM*, *49*(4), 36–39.

White, R. W., & Marchionini, G. (2007). Examining the effectiveness of real-time query expansion. *Information Processing & Management*, *43*(3), 685–704.

White, R. W., & Ruthven, I. (2006). A study of interface support mechanisms for interactive information retrieval. *Journal of the American Society for Information Science and Technology*, *57*, 933–948.

White, R. W., Ruthven, I., & Jose, J. M. (2002). Finding relevant documents using top ranking sentences: An evaluation of two alternative schemes. *Proceedings of the 25th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval*, 57–64.

White, R. W., Ruthven, I., & Jose, J. M. (2005). A study of factors affecting the utility of implicit relevance feedback. *Proceedings of the 28th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval*, 35–42.

Whittaker, S., & Sidner, C. (1996). Email overload: Exploring personal information management of email. *Proceedings of the ACM SIGCHI Conference on Human Factors in Computing Systems*, 276–283.

Wiesman, F., van den Herik, H. J., & Hasman, A. (2004). Information retrieval by metabrowsing. *Journal of the American Society for Information Science and Technology*, *55*, 565–578.

Woodruff, A., Faulring, A., Rosenholtz, R., Morrison, J., & Pirolli, P. (2001). Using thumbnails to search the Web. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 583–590.

Wu, M., Fuller, M., & Wilkinson, R. (2001a). Searcher performance in question answering. *Proceedings of the 24th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval*, 375–381.

Wu, M., Fuller, M., & Wilkinson, R. (2001b). Using clustering and classification approaches in interactive retrieval. *Information Processing & Management*, *37*, 459–484.

Wu, M., Muresan, G., McLean, A., Tang, M.-C., Wilkinson, R., Li, Y., et al. (2004). Human versus machine in the topic distillation task. *Proceedings of the 27th Annual International ACM Conference on Research and Development in Information Retrieval*, 385–392.

Xu, C., Shao, X., Maddage, N. C., & Kankanhalli, M. S. (2005). Automatic music video summarization based on audio-visual-text analysis and alignment.

Proceedings of the 28th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval, 361–368.

Xie, H. (2002). Patterns between interactive intentions and information-seeking strategies. *Information Processing & Management*, *38*, 55–77.

Yee, P., Swearingen, K., Li, K., & Hearst, M. (2003). Faceted metadata for image search and browsing. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 401–408.