



**Middlesex
University**



Designing for Expert Information Finding Strategies

Bob Fields, Suzette Keith, Ann Blandford

Technical Report: IDC-TR-2004-001
January 2004

Interaction Design Centre

Interaction Design Centre
School of Computing Science
Middlesex University
Trent Park Campus
Bramley Road
N14 4YZ

For further details of this technical report series contact:
Paul Curzon (p.curzon@mdx.ac.uk)

Designing for Expert Information Finding Strategies

This paper reports on a study of evaluating and generating requirements for the user interface of a digital library. The study involved observation of librarians using the digital library, working on information finding problems on behalf of clients of the library. The study showed that librarians, familiar with the particular digital library system and with information retrieval work in general, possess a repertoire of relatively simple, yet effective, strategies for carrying out searches, and that non-librarians tend not to deploy the same strategies. After describing the study and the most commonly observed strategies, this paper makes some suggestions for how an understanding of how the librarians organise their activities may generate design ideas for user interfaces that aid ‘ordinary’ users in making use of the strategies that help librarians to be effective users.

Keywords: Digital Libraries, Empirical Study, Usability, Expertise

1 Introduction

Digital libraries are large repositories of electronic documents, generally gathered together according to defined collection criteria and providing some assurance of quality. Compared to the web, they generally have a clear thematic organisation and offer sophisticated document searching and management capabilities. From an HCI perspective, digital libraries are an interesting, and

potentially highly fruitful, object of study because they are becoming pervasive (at least within professions such as education and healthcare) and yet they pose substantial usability challenges. They are treated by many as ‘walk up and use’ systems, and yet learning to use them effectively can take weeks, months or even years. They raise a plethora of use and user experience issues, as well as design issues, many of which have not emerged so starkly in less information-rich applications.

This paper reports on an evaluation and empirical study of the use of a digital library that allowed the researchers to investigate new requirements for the support of searching, browsing and information retrieval tasks. The study began with the unremarkable observation that experienced digital library users are generally more effective at using digital libraries than non experts. The purpose of the study was therefore to investigate in more detail the nature of experienced users’ expertise in order to inform the design of user interfaces that will better support non-experts.

The primary goal of this study was to find out what makes experts better at finding information in the library, with a view to re-designing the user interface so as to better aid non-experts. The model of Sutcliffe and Ennis (1998) allows us to make some initial hypotheses about where the nature of expertise might lie. These authors identify four categories of knowledge that are essential for information finding: knowledge about the domain, the available resources the resources and the device or computer system, and also information retrieval skills. In our study, therefore, pertinent research questions relate to the nature of these classes of knowledge possessed and used by experts, and the relative importance of these knowledge domains.

In the following sections we describe the digital library at the centre of our study, the study itself, and some of the resulting observations about the strategies that help librarians to be successful. We are then in a position to make some conjectures about the implications for user interface design. The findings of the observational study will be set in the context of the development of expertise in the use of digital library systems.

2 The Digital Library and its context

The BT Digital Library has been developed over a number of years by BT to meet the needs of a large and diverse user population within the company. The library gives electronic access to a wide selection of databases of journals, magazines, books, reports, and other digital and online resources. The content of the databases ranges from technical subjects of relevance to engineers and scientists in the company, to business and market oriented material of interest to managers, those charged with following market trends and so forth.

The digital library was developed to largely replace an existing paper based library. Existing library staff undertook the task of making the transition from being librarians in a fairly traditional sense, to being digital librarians. Inter alia this meant that the librarians acquired many new skills and became intimately involved in the design, implementation and maintenance of the digital library system itself, in addition to the ongoing requirements of librarianship.

The digital library offers keyword and phrase searching, full-text access to selected journals and a number of browsable links. An ‘information spaces’ feature provides a monitoring service for new documents on specific themes or categories. The keyword search feature provides a simultaneous search and summary of a number of resources and databases using abstract, index and catalogue data. Two such resources are INSPEC, a scientific and technical resource, and ABI/Inform, which is more business orientated. Each resource can subsequently be searched in more detail, and a number of powerful facilities are provided for specifying searches, viewing search results and how they relate to the underlying indexing mechanisms, and so on.

3 The Trouble with Digital Libraries

While a digital library offers the promise of “access for all” to a rich variety of information resources, this promise often does not become a reality (Blandford et al., 2001; Stelmaszewska and Blandford, 2002) There are several reasons why digital library use has not yet had a deep impact on the information finding practices of potential users, but among the key reasons is the usability of such systems. Users typically find it difficult to arrive at a set of search terms or a query phrase that captures the essence of the information need and produces a sensible set of search results. The problem of information retrieval is sometimes conceived of as one of making a translation between an information need and the language of the search interface. Carrying out such a translation effectively requires knowledge of not only the domain and the syntax for writing queries, but also a knowledge of the indexing and categorisation schemes used within the digital library’s databases.

Against this kind of background, Dillon (2001) argues that the future of both digital libraries as a technology and HCI as a discipline are bound up together. On the one hand, the promise that digital libraries offer needs a user centred approach in order to become reality. On the other hand, digital libraries provide challenges to HCI that will induce the discipline to develop further.

The goal of the study was to gain a better understanding of the behaviour of expert users (i.e., librarians), or in other words, to determine what our experts are experts in. Is it simply the case that experts have a better knowledge of a difficult-to-use user interface? Or is it that they are better at translating an information need, or what Belkin (1980) calls an “anomalous state of knowledge”, into a query that can be entered into a search box to yield an effective set of results from a search engine? The appropriate design response will be different depending on the answers to these questions. Mapping the boundaries of experts’ expertise holds the promise of a strategy for designing and formulating requirements for user interfaces that put the experts’ knowledge “in the world” in such a way that it can materially affect the practices of non-expert users. The study, therefore sought to explicate a little more clearly what experts do and how they are able to do it.

4 Development of expertise in information seeking

Various studies (e.g., Marchionini, 1995; Borgman, 1996; Blandford et al., 2001; Stelmaszewska and Blandford, 2002), demonstrate that many typical users of electronic information resources lack both the knowledge and experience to construct organised and disciplined search strategies across individual queries. Inexperienced users often attempt failed queries repeatedly, while partially successful ones are frequently abandoned without development. When such users work across multiple collections of documents, effective combinations of query criteria are often used in one source, but not with others, whereas failed searches are often retried. In contrast, expert librarians rely heavily on systematic approaches to the evolution of their searches and the selection of search criteria. Ellis and Haugan (1997) present a behavioural model of information seeking that lays out a set of user strategies; when compared to this model, experts demonstrate comprehensible strategies, whereas novices can seem perplexingly illogical. In addition, novices seem confused about how to make effective decisions at the tactical level: Sutcliffe and Ennis (1998) outline a set of decision points and effective tactics in their cognitive model, but real users make such decisions poorly. This seems to make the exploitation of such models in interactive system design difficult. Inexperienced users apparently lack some of the decision-making tactics identified by Sutcliffe and Ennis and the coherent strategies suggested by Ellis and others.

Considering the nature of expertise, recent studies have compared some of the search criteria used by expert librarians against those used by less knowledgeable searchers. Two key variables that have been studied are information seeking expertise and subject knowledge.

Some researchers (e.g., Smith et al., 1989) focus on the importance of domain knowledge as a component of expertise. Vakkari (2001) studied development within one extended searching episode (that took place over several weeks). He reports that users' searches become more focused, as the nature of the information problem becomes better understood. As the search progresses, so the problem definition and query terms become more clearly defined, and users are more easily able to make relevance judgements. In this way, we see a development of searching *with respect to a particular information problem*. Vakkari argues strongly that the effects that are observed are primarily due to development in subject expertise, and not in searching expertise.

Others (e.g., Dillon and Song, 1997; Lucas and Topi, 2002) treat expertise as a binary state – either novice or expert. A detailed study of the roles of searching expertise and subject knowledge was conducted by Hsieh-Yee (1993). One important finding of her study is that experts tend to explore synonyms, to establish what effects these have on search results, whereas novices (even subject experts) do not. However, the effects of differences when they were searching within their area of subject knowledge were relatively small. Outside their area of subject knowledge, differences were much greater: experts were able to use on-line tools such as thesauri to assist in generating alternative search terms, whereas novices relied on their own intuition in selecting terms. In the study reported here, we focus on novice searchers who were knowledgeable within their domain, and on

librarians who searched across domains, and acquired limited and focused domain knowledge through the searching activity.

5 The Study

To investigate the expertise displayed by expert digital library users we observed librarians acting as “intermediaries”. That is, the librarians were acting on behalf of clients of the library who had a real information need and expertise in the problem domain, but lack specific training or skills either in information retrieval in general or in the use of the digital library itself. In each case, the client had attempted to satisfy their information need using the digital library, but had failed to achieve results they were happy with. Problems were elicited from three users in interviews, and two librarians were observed and video recorded working independently on each of the information finding tasks. Each librarian worked on each search problem for around an hour before arriving at a point where an acceptable set of search results had been found, or where more detailed information was needed from the client to improve the search.

The librarians were asked to think aloud while carrying out the search tasks. Actions and verbal commentary on the resulting video recordings were transcribed for later analysis. In analysing the data we were looking for recurrent patterns in the activity of the librarians that seem to characterise aspects of their expertise. As noted above, clients who generated the problems had previously attempted to search themselves, so that log data of their own search attempts was available. This log data reveals something about how they approached the problem before judging their attempts unsuccessful and turning to the librarians for help. The non-expert clients provide a valuable contrast with the librarians, allowing us more clearly to distinguish the expression of expertise from relatively less competent performances.

5.1 Findings

One striking feature of what the librarians did was the complexity and sophistication of their searching activities. For instance, in all the searches we observed, librarians went through a process of re-formulating a query many times before arriving at a set of search results that they were happy to pass back to the client. Table 1 shows one series of re-formulations undertaken as part of an attempt to find documents relating “fault diagnosis and ADSL (Asynchronous Digital Subscriber Line)”. This series is a subset of the librarian’s total search activity while working on this one problem. In the query terms shown, ‘+’ means that the term *must* be included; ‘*’ denotes a ‘wild card’ (so ‘diagnos*’ will match with ‘diagnosis’, ‘diagnose’, ‘diagnosing’, etc.); ‘-’ means that the term must *not* appear in the document; and ‘de=’ means that the term must appear in the list of descriptor terms associated with a document. In this library, the default search is an ‘or’ search; the effects of this are graphically illustrated in the jump in number of hits from query 6 to query 7 (Table 1).

	Query term	Number of hits
1	fault, diagnosis, ADSL	>7000
2	diagnos*, fault, ADSL	>20000
3	diagnos*, fault, +ADSL	644
4	+fault, +ADSL	6
5	+fault, +de=subscriber	76
6	+de=fault, +de=subscriber	32
7	de=fault, de=subscriber,-optical	>20000

Table 1: Typical series of query reformulations

In such a series of queries, and the accompanying verbal protocol, the librarian displays a deep knowledge of the syntax of the query language and the various different resources and databases that the library provides access to. As the sequence progresses, they are also developing a growing awareness of the specific problem domain and how it is represented within the library. In particular, the librarians develop their understanding of which keywords and terms are going to provide the discriminatory power needed.

At numerous stages in the interaction, the librarians display a sensitivity and orientation towards the size of the result set. The size of this set is often used in making the decision as to whether they are happy with the result set. If the set is 'too small' then the search may be too narrow; if it is too large, it is unmanageable and the useful results within it will be hard to find.

A further observation is that librarians in this study appeared to be adept at recognising deficiencies in the knowledge they possessed (e.g. about the field in which the search is taking place) or the query they had produced. Furthermore, they have ways of working that allow them to recognise and remedy such deficiencies.

However, in addition to such "low level" knowledge, librarians appear to possess an array of information finding *strategies*, or higher-level ways of organising their activity. In a sense the root of expertise here seems not to reside in an ability to formulate better queries, nor in possessing a deep understanding of the structure and properties of the underlying databases, though these things are important. On the contrary, the librarians' effectiveness seems to rely on being able to strategically organise a series of queries that allow the expert to arrive at an acceptable result by a process of exploration. The librarian does not form any obvious plan in advance – indeed, neither librarian was able to articulate any plan beyond the current step – but has an effective overall strategy for interacting with the available databases that is tightly coupled with, and dependent on, both the information currently displayed and the history of the interaction so far.

A look at the pattern of re-formulations by librarians and clients supports this. Table 2 shows the number of different queries issued by each librarian for each of the problems.

Problem:	1	2	3	Mean
Librarian A	20	12	14	17
Librarian B	21	15	18	

Table 2: Numbers of query reformulations per search.

Contrast this with the fact that the non-librarian users produced far fewer queries (mean < 10) for these same problems. Furthermore, a detailed look at the data shows that what some client users do is issue a number of separate queries to follow several relatively disconnected hunches or lines of inquiry that they suspect may lead them to useful material. For instance, one client, A, who was interested in the area of emotions and design, issued only three queries, as shown in Table 3.

Query term	
1	games
2	+pleasurable, +design, de=interface
3	jordan (<i>the name of a key researcher in the field</i>)

Table 3: Series of queries produced by client A

It is interesting to note not only that A produced only a small number of queries before giving up and turning to the librarians for help, but also that each query is apparently unrelated to the others. This client appeared reluctant or unable to reformulate a query, preferring instead to try a completely different line of attack in the face of unacceptable results.

Another client, M, a frequent user of the library, was both more persistent and more sophisticated in producing a series of queries in order to find material related to specific forms of fraud involving credit cards, as shown in Table 4. M, it appears, is much happier to reformulate or re-phrase a query on the basis of the result set returned. However, there still appears to be a tendency to persist with an approach for only a relatively short time before trying a different approach. Coupled to this, M appears to have only a relatively limited set of strategies for modifying searches, for example substituting one term for a similar term on synonym (for example, replacing 'cnp' by 'card not present' and then 'credit card').

Query term	
1	ecommerce statistics
2	internet fraud
3	internet_fraud +de=fraud +de=electronic+commerce
4	internet_fraud +de=fraud +de=credit+transactions
5	cnp fraud
6	card not present fraud
7	credit card fraud

Table 4: Series of queries produced by M.

Librarians, on the other hand, rarely produce an optimal and acceptable set of query terms (and therefore result documents) first time around. As can be seen (e.g. in Table 1), while they might have a number of separate ways of approaching the problem, each approach is likely to be exploratory and will involve carrying out a series of manipulations of a set of search terms. This will progressively modify the search set so as to improve the results obtained, test out hypotheses, or to gain a better understanding of the domain in which the search is taking place.

The skill of librarians, then, seems not (only) to be in accurately translating the needs of the user into an effective set of keywords. Indeed, the figures above suggest that librarians possess ways of going about the tasks of searching, strategies that are different from those used by non-experts.

5.2 *Strategies*

In interacting with the library, the librarians were able to call upon a number of 'strategies' or 'template' ways of organising their activity. These strategies allowed the librarians to make use of patterns of action that had proved successful in the past, and which can be flexibly and contingently deployed in future situations. A number of such recurrent patterns of activity were evident in this study in the way that librarians made artful use of the features of the user interface. Three commonly observed strategies, discussed further in the remainder of this section, are:

1. to systematically reformulate queries (rather than just abandon them and try again) in order to improve a search;
2. to manipulate a search query in order to increase or decrease the size of the set of search results found; and
3. to carry out searches and explore the result sets so as to learn more about the domain of inquiry and the discriminatory power of search terms within it.

5.2.1 *Reformulate to improve search*

As has been noted already, non-expert users were frequently observed to issue a series of queries, each reflecting a different approach to finding the required information. If one query fails to return a useful set of results (either because the set contains very few items or, as is more frequently the case, the set contains far too many items) then there is a tendency for the non-librarian to abandon that line of enquiry and attempt another one. As we have seen above, it is relatively rare to see non-librarian users issuing a series of tightly connected queries where each is a reformulation of an earlier one.

Librarians, on the other hand were observed to issue a series of queries, each building on earlier queries and modified in the light of the results obtained. For instance, one librarian accounted for one particular reformulation in a series of queries, where part of the query is left untouched and part is altered, thus:

Librarian B: So I am fairly happy about getting towards ADSL words and now I am looking for the diagnosis words

Within the strategy of systematically reformulating, librarians displayed a number of tactics that included replacing search terms (e.g. replacing “ADSL” with “subscriber” or “CNP” with “card not present”) and manipulating the query term syntax (e.g. replacing “fault” by “+fault” and later by “+de=fault”).

5.2.2 *Expand and contract*

Both librarians and other clients displayed a sensitivity to the size of the set of results returned from a search, but in rather different ways. Put simply, for many of the clients a very large result set was a poor outcome that often engendered a complete change of approach. For the librarians, on the other hand, a large result set was often seen as an opportunity: a set possibly containing useful material that must be “whittled down” to the useful core.

Librarian B: This is a complex search, it's not clear from the outset, what is the core and what you are going to use to narrow it down.....you've got fraud and you've got credit cards and the internet

Therefore, it was not unusual for librarians to carry out their searching in two relatively distinct phases. The first, exemplified in the fragment of talk below, is concerned with generating a large set of results that will contain the desired material. Sometimes this will involve “broadening” the search to include more hits, by replacing a specific term or phrase with a more general one.

Librarian B: I know there are those two terms for ADSL. I'm broadening out ADSL a bit now into subscriber lines

The second, embarked upon when the librarian was reasonably certain that the search terms they had found so far resulted in a reasonable selection of useful material, involves narrowing the search once more to eliminate as many as possible of the irrelevant hits.

Librarian B: it is very difficult to get a tight search that is exactly what you want

5.2.3 *Improve knowledge of the domain*

A number of episodes were observed in which the librarian carried out searches that did not appear to get them closer to the information-seeking goal. Many such interactions were exploratory sidelines in which the terminology and keywords of the domain were being investigated. In some cases, this reflects the fact that the librarians tend not to be experts in the domain of the search, as reflected in the following account:

Librarian A: I'm wanting to see what will turn up in general, because I'm looking for a bit of help with the terminology

In some cases, the problem wasn't simply that the librarian was unable to produce terminology relevant to the search. Rather, the librarian was concerned to find out how the domain was represented in the keywords used by the database:

Librarian A: I think the problem is we are playing with stuff that is a bit....they are soft terms

Librarian B made similar comments on the same search task, remarking that the research area (“pleasurable design”) was ‘new’ and the vocabulary ‘soft’. Therefore, it could be some time before agreed key terms emerged and became part of the controlled vocabulary used by the indexers.

Finally on this point, it is worth noting that the librarians’ tactics for improving their knowledge of the important keywords in a domain are sophisticated and involve the artful use of the library’s facilities. For instance, on more than one occasion, a librarian would conduct a search in one database, Inspec, which is well known for its robust system of keyword classification, but not with the expectation of finding useful results. On the contrary, they were seeking to improve their knowledge of keywords in the domain, which would then be used to search in another database, ABI, with its less sophisticated classification scheme.

6 User Interface Requirements

6.1 *Strategies at the user interface*

An obvious but important conclusion of the above discussion is that one characteristic of non-experts is that they do not possess ‘expert’ strategies such as those identified above. Among other things, this prevents them from making as effective use of the DL as may be possible. A goal for design, therefore could be to give users of the digital library the information and resources they need in order to be able to carry out some of the strategies that prove so successful what carried out by the librarians.

As an illustration of this, we turn briefly to a simpler case where superficial aspects of a user interface can markedly affect the way users carry out a task. In a study of the use of a simple game, Cockayne et al. (1999) use the concept of strategy that is further developed by Wright et al. (2000). The point of these papers (especially the former) is to say that we can identify strategies that the design of a user interface makes possible or that the design encourages particular strategies.

Cockayne et al. describe a study of the familiar game, the 8-puzzle, in which a simple strategy exists that will allow players to complete the puzzle quickly and more reliably than most other strategies. However, most players very rarely guess the strategy in the conventional form of the game. If, however, the visual presentation of the game is slightly re-designed, players rapidly acquire and deploy the strategy, allowing them to complete the puzzle more quickly and with fewer moves.

A conclusion that follows from the above-mentioned work is that in order to be able to employ a strategy, or engage in some pattern of behaviour, two conditions must hold. Firstly, the user must be aware of the strategy itself, and secondly, she must be provided, by the user interface, with any information resources that are required for its performance. It is worth noting that apparently trivial changes to a digital library’s user interface can produce significant changes in the ways users

formulate queries. Belkin and his co-workers (2003) show that the simple addition of a text instruction (“Information problem description (the more you say, the better the results are likely to be”) adjacent to a digital library search box leads users to enter significantly longer query phrases. This, in turn, leads to significantly better search results than queries produced without the additional instruction.

6.2 Supporting information finding strategies

In the current paper it will only be possible to begin to explore ways that librarians’ searching strategies may be supported or “externalised” in a way that makes them easier for non-librarian users to discover and adopt. Let us illustrate how this might work by taking one of the strategies identified above: that of selectively and deliberately expanding and contracting the set of search results.

A standard search interface, as found on most search engines and digital libraries, permits this kind of strategy, as the user is presented with a “search box” where arbitrary search terms may be entered. However, such an interface does nothing to encourage or suggest strategies for manipulating the search or result set.

Some well-known internet search engines (such as Google) do appear to have this kind of behaviour in mind and provide some encouragement in the form of a “search within these results” feature. The purpose here is, presumably, to give users a way to progress their search when it initially yields a large set of results. Thus the user is prompted for additional keywords that will narrow down the search space.

Of course this isn’t the only way of affecting the size of the results set, and the user interface of a digital library could offer other possibilities. In the remainder of this section we offer some possible ways that user interface designs could provide direct support to assist users to engage in the strategies described in Section 5.2.

6.2.1 Reformulate to improve search

Provide explicit hints for kinds of reformulation other than simply adding new keywords. For instance, many search mechanisms distinguish between separate keywords and multi-word phrases, and the user interface could suggest the conversion of one to the other to either expand or contract the result set. Recall the librarians’ knowledge of search term syntax that allowed them to change the status of particular search terms within the query. Providing hints may remove the need to be familiar with the syntax, while allowing the non-librarian to see the potential offered by particular kinds of query reformulation.

Librarians in this study made extensive use of a feature known as the ‘keyword browser’. This gave access through a separate page of the user interface to the set of index terms according to which documents in the library are organised. It provided a highly effective means of affecting the selectiveness of searches. However, non-expert users in our study made almost no use of this feature. The reasons for this are likely to include poor visibility on the display, lack of obviously indicated purpose, and poor integration with the more obvious search features. Clearer presentation and better integration of the keyword browser feature

with the core searching system could empower less sophisticated searchers by giving them easier access to powerful features.

6.2.2 *Expand and contract*

Provide context-sensitive help for improving searches. Part of what librarians are able to do is to make highly tailored responses to search results sets that depend on the size of the results set produced. As noted above, they will explicitly try to expand the set to a large one that includes relevant material, and then will try to reduce that large set to one that has fewer irrelevant hits. Help or hints that offer advice based on the result sets produced could support such a strategy. Search engines such as Google go some way to providing this kind of support when they offer alternative or corrected spellings of search terms.

6.2.3 *Improve knowledge of the domain*

Show how discriminating each of the search terms is. This could be a numeric presentation indicating how many ‘hits’ each term returns, or could be a graphical representation, such as a Venn-diagram, that shows how the result set relates to the search terms. For instance, Jones and others (Jones, 1998; Jones et al., 1999) describe such a graphical presentation of search queries and results. Recall the librarians’ use of the number of ‘hits’ produced in response to a query; providing additional information in the interfaces may endow non-librarians with some of this sensitivity to search results. It might even help librarians.

7 **Conclusions and Further Work**

We have described a study in which it was observed that “expert” users of a digital library are, not surprisingly, able to be more successful than “non-experts” at finding information. In addition to the possession a better and more detailed knowledge of the library, its user interface, and its internal structure, there are other differences between “expert” librarians and “non-expert” library clients that allow the former to be very effective at conducting searches and finding information. In particular, it was noted that librarians are able to draw upon an extensive repertoire of strategies for finding information in the library. It has been suggested here that user interfaces should be constructed that will aid non-librarians to engage in some of the strategies that allow librarians to be successful.

The study yielded many insights about the nature of our experts’ expertise, including the range of strategies to which they appear to have access. This in turn has led to a number of possible design innovations that seek to improve the effectiveness of non-experts. This, of course, leaves open a number of questions that may be explored further. For example, precisely how do librarians make use of the available information in enacting their strategies? Do they direct their attention to the same things as non-librarians? How effective are the design suggestions outlined above?

In addition to showing how an understanding of expert behaviour can guide design, the work carried out here highlights the value of using the difference between user groups with varying forms of expertise as a useful focus in design. A valid goal for design is to accommodate variation and change in expertise by understanding the range of strategies that are employed, and designing so that “expert strategies” are more readily available to non-experts too.

8 Acknowledgements

We would like to thank the Librarians and Clients who participated in this study as well as members of the Interaction Design Centre at [anonymous] who commented on earlier versions of this work. This work was supported by EPSRC grant [anonymous].

References

- Belkin, N. J. (1980). Anomalous states of knowledge as a basis for information retrieval. *Canadian Journal of Information Science* **5**: 133-134.
- Belkin, N. J., C. Cool, D. Kelly, G. Kim, J.-Y. Kim, H.-J. Lee, G. Muresan, M.-C. Tang and X.-J. Yuan (2003). *Query Length in Interactive Information Retrieval*. In 26th Annual ACM SIGIR Conference on Research and Development in Information Retrieval, ACM Press.
- Blandford, A., H. Stelmaszewska and N. Bryan-Kinns (2001). *Use of multiple digital libraries: a case study*. In Proc. JCDL 2001, ACM Press.
- Borgman, C. L. (1996). Why are online catalogs still hard to use? *Journal of the American Society for Information Science* **4**(7): 493-503.
- Cockayne, A., P. Wright and B. Fields (1999). *Supporting interaction strategies through the externalisation of strategy concepts*. In Proceedings of Interact'99.
- Dillon, A. (2001). Technologies of Information: HCI and the Digital Library. In *Human Computer Interaction in the New Millennium*. J. M. Carroll ed., Addison Wesley: 457-474.
- Dillon, A. and M. Song (1997). An empirical comparison of the usability for novice and expert searchers of a textual and a graphic interface to an art-resource database. *Journal of Digital Information*. **1**(1).
- Ellis, D. and M. Haugan (1997). Modelling the information seeking patterns of engineers and research scientists in an industrial environment. *Journal of Documentation* **53**(4): 384-403.
- Hsieh-Yee, I. (1993). Effects of Search Experience and Subject Knowledge on the Search Tactics of Novice and Experienced Searchers. *Journal of the American Society for Information Science* **44**(3): 161-174.
- Jones, S. (1998). VQuery: A graphical user interface for Boolean query specification and dynamic result preview. Department of Computer Science, University of Waikato. Working Paper. 98/3.

- Jones, S., S. McInnes and M. S. Staveley (1999). A graphical user interface for Boolean query specification. *International Journal on Digital Libraries* **2**: 207-223.
- Lucas, W. and H. Topi (2002). Form and Function: The Impact of Query Term and Operator Usage on Web Search Results. *Journal of the American Society for Information Science and Technology* **53**(2): 95-108.
- Marchionini, G. (1995). *Information Seeking in Electronic Environments*, Cambridge University Press.
- Smith, P. J., S. J. Shute, D. Galdes and M. H. Chignell (1989). Knowledge-based search tactics for an intelligent intermediary system. *ACM Transactions on Information Systems* **7**(3): 246-270.
- Stelmaszewska, H. and A. Blandford (2002). *Patterns of interactions: user behaviour in response to search results*. In Proc. JCDL Workshop on Usability, A. Blandford and G. Buchanan eds.
- Sutcliffe, A. and M. Ennis (1998). Towards a cognitive theory of information retrieval. *Interacting With Computers* **10**: 321-351.
- Vakkari, P. (2001). A theory of the task-based information retrieval process: a summary and generalisation of a longitudinal study. *Journal of Documentation* **57**(1): 44-60.
- Wright, P. C., R. E. Fields and M. D. Harrison (2000). Analyzing Human-Computer Interaction as Distributed Cognition: The Resources Model. *Human-Computer Interaction* **15**(1): 1-41.