Preprint: final version of paper to appear as:

BUCHANAN, G., CUNNINGHAM, S. J., BLANDFORD, A., RIMMER, J. & WARWICK, C. (forthcoming) Information Seeking by Humanities Scholars. To appear in *Proc. ECDL 2005*. LNCS.

Information Seeking by Humanities Scholars

George Buchanan, Sally Jo Cunningham, Ann Blandford, Jon Rimmer, Claire Warwick

University College London, United Kingdom
{g.buchanan, a.blandford, c.warwick}@ucl.ac.uk
University of Waikato, New Zealand
sallyjo@cs.waikato.ac.nz

Abstract. This paper investigates the information seeking of humanities academics and scholars using digital libraries. It furthers existing work by updating our knowledge of the information seeking techniques used by humanities scholars, where the current work predates the wide availability of the Internet. We also report some of the patterns observed in query and term usage by humanities scholars, and relate this to the patterns they report in their own information seeking and the problems that they encounter. This insight is used to reveal the current gap between the skills of information seekers and the technologies that they use. Searches for 'discipline terms' prove to be particularly problematic.

Keywords: Digital Libraries, Human-Computer Interaction, Information Seeking, Humanities.

1 Introduction

Information retrieval research has made significant improvements in the quality of results from interactive searches. Simple approaches such as the log rule [18] and query term expansion [9] give users much better results than earlier techniques. However, in the complex arena of digital libraries, where users have to choose from a rich variety of search criteria - e.g. metadata field, Boolean logic - problems arise in the users' choice of effective criteria. In particular, users seem to have problems interpreting causal relationships between their inputs and the returned results.

This gap of interpretation requires further study. Our research is founded on the approach of information seeking research. In contrast to information retrieval, which focuses on improving search performance by technical means, information seeking

describes human behaviour. Improvement is achieved through a better understanding of user abilities and expectations, leading to changes in underlying system mechanics or the human-computer interaction as appropriate. The objective of our research is to reveal more about users' intentions in their information seeking behaviour, and the strategies and tactics they choose that they believe will improve their outcomes. In understanding their anticipations, we hope to provide interactions that better bridge the gap between the system's underlying operation and the users' conception of the system.

This study focuses on humanities scholars – intellectually able seekers who are not technical in orientation. The body of the paper commences with an examination of information seeking research - both general material and findings particular to the humanities. Having reviewed the current state of knowledge, we introduce our user study in Section 3. The findings of that study are presented in Section 4 and discussed in Section 5, before we conclude with a summary of our contribution and future plans.

2 Information Seeking

A key theme in user-centred information retrieval research has been the investigation of the choices users make when searching: what terms they choose, how many, and which other features (e.g. phrase search or Boolean logic) they naturally use.

One immediate distinction that can be made between users is their level of experience, in terms of either interactive search or the subject domain. Lucas and Topi found that experienced, trained searchers use more query terms and exploit Boolean logic more frequently [10]. Hsieh-Yee [7] and Wildemuth [17] studied the impact of domain knowledge; they found that underlying search skills have a more significant effect on query formation than domain knowledge.

The overall picture from studies of choice patterns within user queries, particularly on the web (e.g. [8]), has been that users use few terms and seldom employ Boolean logic or other advanced search criteria. The consistent picture of interactive search is that few expert searchers exist, and most seekers use simple two or three term queries.

In addition to studies of users' behaviour when engaged in interactive search, other researchers have investigated the broader patterns of users' information seeking strategies. One key models is that of Ellis [5,6], who identifies actions such as *Starting*, where the seeker identifies initial sources of information, and *Chaining*, where references are followed forwards or backwards to extend the scope of the covered area. In Ellis' model, users may move from one action to another and no order is assumed. Below, we relate our findings to the work of Ellis.

2.1 Information Seeking in the Humanities

The available literature on information seeking in the Humanities is relatively limited. Wiberley et al's [15] early work disputed the then well-accepted view that humanities search terms were frequently general and imprecise. He identified the frequent use of items such as the names of persons and places. This work was later

extended by Bates [2]. This body of evidence has revealed that humanities scholars in fact frequently use specific and highly selective query terms.

Recently, comparisons have emerged with other academic disciplines. Whitmire [14] reports that humanities seekers demonstrate a significantly higher use of library facilities than other academics. They more often use catalogues, turn to librarians for assistance, browse, use reserve collections and journal indexes, etc. The collaborative aspect of information seeking has frequently proven significant; Watson-Boone [13] identifies the importance of the professional network of neighbouring and distant colleagues in the information seeking of humanities scholars. Thus, humanities information seeking demonstrates a strong use of human support, as well as a more intensive use of printed or mechanised seeking tools. However, though humanities scholars do turn more frequently to librarians, they do so with some reluctance [16].

The impact of the digital information seeking environment in the workplace is particularly poorly understood. Tibbo, [12] in her investigation of the information seeking of historians, states that "we have no idea if they succeeded in finding these materials based on web searches", and similarly questions electronic catalogues. Wiberley's later report [16] itself faces the problem of being a retrospective over the previous ten years. The availability of electronic resources on the researchers' work table was only becoming a factor towards the end of that period.

Though this material is helpful in informing the design of library services for humanities scholars, the picture is far from complete. For example, little has emerged about the use of new electronic sources of information such as the Web, and other recent developments such as the widespread introduction of online journals also have yet to be systematically studied. For the technical development of digital library (DL) systems, the distinctions made in [15] and [2] between different query terms do not fit well with search indexes for text, where distinguishing the role of a particular word is extremely difficult. In other words, there is a gap between our information seeking insights and our information retrieval technologies.

2.2 Information Seeking: Summary

This paper reports our findings which are helping to narrow this gap, and bringing information retrieval technology and interaction closer to the information seeking of humanities scholars. In the next section, we describe a user study that we undertook to update our picture of the information seeking strategies of humanities scholars, and to reveal the parts where the misfit between technology and user skill is greatest. In [3], we used information retrieval evaluation techniques to identify patterns in the textual properties of documents gathered by users in the course of their information seeking. In this paper, we again use information retrieval measures to scrutinise the information seeking of users - in this case, analysing the types of terms that they use to search for documents. Through this, we identify the relationship between the types of search terms identified by Bates [2,11] and the frequency of occurrence of words associated with each type of term.

3 User Study

To improve our understanding of the information seeking strategies of humanities academics we conducted interviews with eighteen members of the Faculty of Arts and Social Sciences at the University of Waikato in New Zealand.

The interviews were conducted over a six week period, in the participants' own offices. We wished to elicit their experiences of using digital libraries and electronic library catalogues, and their perceptions of the problems and successes they experience in using them. Participants were encouraged to demonstrate their strategies using the University's electronic catalogue, which provides direct access to a number of electronic resources as well as the books physically in the library. The use of a familiar system was intended to reduce the effects of learning, and to focus their explanation of their information seeking in a context where they were able to demonstrate their own expertise. In addition to the traditional interview approach, the search terms used were recorded and later analysed for their effect.

Our goals were to identify areas where strategies are already well formed, areas where seeking was commonly perceived as difficult, and the types of query choices made (terms, options, etc.). In the latter case, we wished to identify any correlation between problematic areas of information seeking and the sorts of query choices made.

3.1 Participants

In this paper a two-letter subject code and number identify each participant. Participants in our study were academic faculty, ranging in age from 28 to 65; their backgrounds are outlined in Table 1. We compared experiences across different levels of experience, background, seniority and age. The participants were recruited through direct contact and a circular through the faculty newsletter.

Due to the relatively difficult access to large reference libraries in New Zealand, we might expect a higher adoption rate of technology compared to academics in Europe or North America.

3.2 Study Method

The interviews occurred in the participants' own offices and they used their own computer equipment when interacting with the library systems. After a short briefing on the objectives of our research, participants engaged in semi-structured interviews concluding with a demonstration by them of their use of the information seeking tools available to them. During the interview, only the investigator and the participant were present, the investigator noting the participant's responses and explanations. Their interactions with the library system and/or Internet search engines were also noted, and search terms recorded for later analysis. In addition to library systems, the academics were also encouraged to describe their use of the web in general as a research and information resource.

After the interview, the university catalogue index was used to identify the frequency of the terms that the participants had used. Where another source was used, the result list for each individual term used was inspected to deduce the effect of each word and the default search criteria recorded. The university catalogue system produces results in a number of forms, including ranked list (default) and alphabetic title order. Search criteria include "Keyword anywhere" and "Title begins with". The participant's explanation of the search - i.e. their intention - was then compared against the achieved effect. Terms were also compared against the search term taxonomy of Bates [2]. Thus, for each term we identified its semantic form (e.g. geographical name) and its rate of occurrence (document frequency).

Query (or browsing) intentions were related to Ellis' information seeking strategies [5,6]. This applied to information seeking activity described by the interviewee during both the interview and the demonstration parts of the study.

Table 1. Brief details of participants

Identity	Subject	Post	Years in Field	Gender	Use
AN1	Anthropological History	Lecturer	15	Female	Low
AN2	Anthropological History	Assoc, Professor	21	Female	Low
AN3	Anthropological History	Professor	20	Male	Low
ED1	Education	Dean	26	Male	High
ED2	Education	Research fellow	20	Female	Med
EN1	English	Senior Lecturer	12	Male	High
EN2	English	Assoc. Professor	20	Male	Low
EN3	English	Senior Lecturer	20	Female	Med
HS1	History	Lecturer	10	Female	Med
HS2	History	Senior Lecturer	35	Male	Med
HS3	History	Senior Lecturer	27	Female	Low
LI1	Linguistics	Assoc. Professor	25	Male	Med
LI2	Applied Linguistics	Senior Lecturer	10	Female	High
MS1	Media Studies	Professor	24	Male	High
PL1	Philosophy	Lecturer	8	Female	High
PL2	Philosophy	Professor	30	Male	Med
PO1	Politics	Professor	25	Male	Low
EV1	Environmental Mgmt	PhD Student	3	Female	Med

4 Findings

Our participants reported a wide variety of experiences with digital libraries, electronic library catalogues and the web generally. In this section, we discuss the role of each of these types of resources in turn. Subsequently, we identify common strategies described by our readers, and then finally discuss the insights gained through a detailed examination of their searches.

4.1 Digital Libraries

Our readers were not explicitly aware of the concept of digital libraries. Commonly, they referred to DLs as "online databases", or referred to the electronic library catalogues (e.g. the university's own catalogue) that themselves linked to the content of actual DLs. The use of electronic journals also confused this issue. However, all the participants used digital libraries in one form or other. For these users, DLs as a concept were subsumed by digital forms of familiar paper-based services.

Actual use of digital libraries varied. Participant AN1 described herself as "a technophobe", seldom using digital resources although she described how she "repeatedly tried to use online journals". In contrast, participant ED1 described himself as an "enthusiastic user of online material". Each of these participants saw themselves as being at an extreme of the digital use continuum.

These differences were not, however, simply personal. Both education users (ED1, ED2) reported high levels of computer use, and stated that colleagues in other institutions were similar to them. ED1 and ED2 both argued that the strong uptake of computer technology in schools had driven them to be early adopters of online information systems: ED1 describing his use of a dial-up connection to a document database in San Diego in the mid 1980's, and ED2 her use of online material from the UK towards the end of the same decade. Conversely, the anthropological history users (AN1/2/3) identified themselves as low-frequency users and also reported that their counterparts elsewhere shared their own preferences.

One marked difference in technology arose around the issue of access. Once material was found online, educationalists reported that they seldom had difficulty retrieving it. On the other hand, Anthropological History and English users regularly found themselves "barred" by access controls - i.e. they were expected to pay for access. A few individual users - e.g. LI1 - were directly involved in projects that gave them privileged access to material of high value to them, whilst others such as EN1 personally paid for valued resources. There was a strong correlation between access difficulties and take-up: all readers who reported access problems as being acute were infrequent users of online libraries and journals.

4.2 Internet Use

Across all participants and degrees of digital library use, the Web was seen as a useful substitute for a traditional encyclopaedia. However, identifying the source was seen as an important step in verifying the trustworthiness of the information. For example, AN1 said "I use the Internet to look up definitions of words, rather than turn to a dictionary like the OED. I know that we've got access to the OED online, but I prefer to search for it instead." Similar views were reported by the other readers.

The Internet was also a good starting point for more specific strategies. For example, it could provide initial sources for chaining, quotations, and checking bibliographic data.

4.3 Career and Community effects

Most participants reported that when they were new to an area of research or teaching, their DL use soared, especially when they were at an early stage of their career. This pattern was repeated on the smaller scale; e.g. when a new project was starting, or a new sub-field of study was explored, participants reported a peak of digital library use. For instance, EV1 reported that the start of her PhD studies resulted in an eighteen month period of extensive literature seeking, followed by another eighteen months of fieldwork, in which all she did was track for later publications. LI1 also reported the same pattern in regard to a book he was writing and, in common with EV1, was now returning to online sources to check for literature that had appeared in the meanwhile to ensure that all the references were up to date.

Findings vary with the age group and experience of the researchers interviewed. More experienced researchers relied more on personal contact and domain knowledge; they would know of many developments before they are formally reported [1]. For these users, personal contact was a more important source of new data than direct information seeking in any medium. Literature was invoked more to back up arguments than to develop them. One particular Internet strategy did emerge in this area: fifteen participants reported 'Googling' for the homepages of known researchers to either update their list of publications or check specific citation details.

4.4 Information Seeking Strategies and Tactics

In this subsection, we report on the higher-level information seeking strategies and tactics used by the participants in our study. As shall be seen, there was generally a strong relationship between the behaviours we observed and those found in information seeking models such as those of Ellis mentioned earlier.

Choosing to Search or Browse: "Browsing" in digital environments was noted as a particular problem - highly effective in physical libraries, but difficult in an electronic environment. EN1 suggested a reason for this: "with an online search you have to be more targeted as the structure is linear. On the other hand, in a real library I go to a physical place and find stuff alongside what I am looking for under the same call number; I think I can do that in the DL but I am less likely to do that." During search, subject classifications were even more rarely used; indeed, any mention was negatively critical. For example, PL2 commented: "I've encountered these things in the past... I don't think of subjects that way. I'm always surprised where books are classified." Our readers demonstrated varying awareness of the existence of browsing in a DL, and none reported doing it regularly. This correlates with findings from our earlier studies. Search also proved an ineffective replacement for browsing, participants observing that it is easy to have lots of hits or no hits. Thus, there was little satisfaction when searching broadly in a field understood by the humanities academic as well-defined. One alternative is the use of chaining instead of browsing in the digital domain.

Choosing where to Seek: Participants reported difficulties early in their use of digital resources in identifying appropriate sources for their areas of interest. Later, having selected sites or journals that met their needs, they would return to these

regularly. However, they often did so directly, not interacting with the institutional catalogue that indexed the external material. ED2 explained: "When I started, I took a long time to find what was there. Now, I just go direct - you know, not using the uni catalogue - straight to the journal or database itself." A different approach was taken by PL1, who said "When I am searching, I just use Google; it usually finds the article I want and I go straight from there to the database - when I locate a paper often I have online access through the library... it will come up and when I click on it and access if good it will just say 'You are logged on at the University of Waikato' or something like that."

Ellis' Patterns: Ellis' principle of chaining - using references and citations in known works to find unknown ones - was noted by all our participants. This was a dominant form of behaviour, which was often reported as the common context within which effective search occurred. Thirteen participants also reported their tracking of particular journals - described by Ellis as 'monitoring', and eleven participants reported the same behaviour in regard to known researchers in their field. Twelve reported behaviours Ellis described as 'browsing'.[1] 'Verifying' previously found information was reported by ten participants.

When returning to journals, a serial reading of the latest issues was common. For example EV1 reported: "There are certain journals that do publish a lot of useful things and I sometimes trawl through those issue by issue - I thought it would be a waste of time, but it wasn't: I found things I wouldn't have by search."

4.5 Query Term Effects

In our study, we observed forty six different searches with 121 query terms. Previous research [2,15] has identified key query term types that appear in the searches of humanities users: names of individuals, geographical names, chronological terms and discipline terms (understood terminology within the field – e.g. in computer science, 'recall and precision'). It is argued that these query term types have a specific meaning within a discipline and therefore should be "good" query terms. However, these term types have not been studied from the perspective of information retrieval, to understand their usefulness when input to a search engine. A specific term may carry strong semantic cues, but information retrieval techniques focus instead on the rate of occurrence of a word or phrase. For example, 'precision' can occur in many more contexts than 'recall and precision' and mean different things even in a computer science corpus. Thus, there is a gap between the human understanding of the term and the treatment of the same term by the computer. In addition, we wished to see how these types of terms are used by humanities academics when performing actual queries. The effect of single words is strongly influenced by the criteria with which they are used. For example, a word occurring in a phrase plays a different role to a word used on its own.

Bates [2] and Wiberley [16] identify a range of term types that are particular forms of 'proper' words; e.g. names of places and people. Proper names were reported by all

[1] N.B. Ellis defines this as scanning known sources of information - this is not quite the common usage in information seeking.

our participants as being a commonplace form of term in their searching – e.g. when searching for a journal article by a particular author. In addition, fifteen of our participants used one or more proper term in their description and demonstration of their own searching. Thus, this particular type of query term is of particular interest.

When a person's name was used, the number of document matches in the university library was typically under 100, and these were grouped by name, so the number of displayed results was usually under 25 (one page of results). Only in two cases was this number exceeded: when the name was matched in any field. In one of these cases, the results for 'Caxton' soared to just over 3000 matches from a body of over 1,000,000 works.

Geographical names were not as selective, particularly when they referred to New Zealand locations. The University of Waikato specialises in New Zealand specific research, so even regional name searches on title return hundreds or thousands of hits; for instance, a search for 'Otago' - a region of NZ - in the title of a book would return 1268 entries; when executed on 'keyword anywhere', 3174 items were returned.

What we observed in the case of common proper names was that 'keyword anywhere' represented a poor choice of search criteria. Unfortunately, this was the default used by the library catalogue, and was seldom changed by moderate and low users of DL resources. However, both forms of proper (person and place) name, particularly when used with additional terms, resulted in smaller result sets that our participants reported positively. Combined with a default relevance ranking, a satisfactory match was consistently found near the top of the search result list. Furthermore, in the case of the six participants who used phrase searching when using people's names (all but one reporting themselves as a 'high frequency' user), the selectivity of the term prove even stronger. Each of these users also reported the need to try alternative forms of the same name.

Another form reported by Bates was the chronological period. This occurred on only six occasions within our participants' searching, but was reported in the course of eight interviews. Where chronological periods were used, a search typically returned around 1000 hits on title, with more modern periods being increasingly common (e.g. 1869 hits for "twentieth century").

One further form of term that Bates [2] specified was the 'discipline term'. Again, these terms were both reported and demonstrated by our seekers. Such terms are often small phrases that consist of two or three individual words.

Examining these terms through the online catalogue's index revealed a clear problem. When using the default "keyword anywhere" search of the university library, very high numbers of results were returned. The library indexer automatically restricts searches to returning at most 10,000 matches. This limit was consistently exceeded when such discipline terms were used; the best case was "mercantile economy", and even that resulted in over 4,000 hits. The participants in these cases found little of relevance even at the top of the search result list. This problem seldom occurred for proper name searches unless, again, several terms were used together. Conversely, use of these terms as a phrase, particularly against individual fields, often resulted in few or no hits (less than 25 hits in 18 of 21 examples, 2 being no hits). Such search criteria prove imprecise unless treated as a phrase. Participants, such as ED2, who were satisfied with digital libraries specifically checked that they used

these terms in a phrase. Naive users, on the other hand, performed no checks for this and consistently obtained very large result sets.

This problem with discipline terms seems to correlate with problems reported by seekers when searching more broadly, and also ties with the strategies that they use.

When terms are combined, the effects vary widely depending on the search criteria used. For example, a combination of 'Otago' and 'twentieth century' can range from one hit to 10,000 hits, depending on the options used. The default search option for the library catalogue (keyword anywhere) in fact returns the highest figure.

Participants who reported satisfaction with their experience of online systems (e.g. LI1, ED1, ED2, EV1, PL1) all regularly described not only the search terms that they used, but which fields those terms would be used in. For these users, term and field are closely related. On the other hand, AN1 represents those with a less positive experience: "When I look stuff up, it just is far too much. My colleagues are overwhelmed, two of them anyway, by finding lots of stuff and I don't think they find much of it helpful; I don't. It is much easier just to go to the shelves and browse there". In the case of these users, they reported finding too much, and in our interview only demonstrated "keyword anywhere" searching. Given the example above, it can be seen that effective use of fields in search is critical to consistent satisfaction.

Similarly, satisfied readers articulated both a pattern of changing their search terms flexibly and, even in the case of ED1, an awareness of different strategies in different libraries: "...because I have got a, generally got a clear sense of what I am looking for. My knowledge and fluency in what I am looking for lets me get good descriptors. There's always a learning curve with a new repository. The familiarising process takes time - what works in each library." Again, dissatisfied users did not express this understanding. However, most participants expressed little or no awareness of different strategies being needed in different contexts.

When comparing between our seekers, we sought any relationship between their attitudes and experience with digital library resources and their use of queries. As has been touched on a couple of times in this section, query criteria were used more selectively by more frequent and enthusiastic readers. Naive or unenthusiastic users only used the basic search criteria, and appear to be more frequent users of 'concept terms'. As noted above, concept terms require the use of phrase searching to improve precision in results sets; low frequency users AN1 and SS1 used one or more concept terms in each search and never used phrase search. In contrast, highly satisfied intensive users (e.g. PL1, EN1) used concept terms in only one case, and immediately applied phrase search, citing the problems that they found if they forgot to use this. Experienced and keen users of digital resources also exploited fielded querying, and proper names more frequently occurred in their search criteria. All five users who identified themselves as high intensity users described using Boolean search; of the moderate and low frequency users, only two noted an awareness of this option, and both had previously been high intensity users. However, our samples are too small to determine any statistical significance.

Strangely, where naive strategies were applied, readers would often report positive experiences with searching on the Internet as opposed to disappointment with specialised catalogues and libraries. Google was consistently identified as the Internet search engine used by our readers. Given the design choices of Google that maximise precision at the expense of recall (e.g. all search terms are required in matched

documents) and positively weight query term proximity in matching documents (providing a simile to phrase searching), there may be a relationship between the design choices of the search engines, strengthening poorly focussed searches, and the users' positive remarks.

5 Discussion and Conclusions

We studied the information seeking skills and strategies of eighteen humanities academics. Across our participants, there was a correspondence between high usage (present or historic), strong search skills (e.g. use of fielded search) and a greater degree of satisfaction with DL systems. Unfortunately, it is difficult to ascertain what empowers a user to move from being a low-skill, low or moderate frequency user to a more flexible, satisfied user. Our high-frequency users could pinpoint a specific event, e.g the beginning of a project, as a turning point in their use of DLs and this, in turn, suggests that the change can occur over short periods of time. As with our observation of career effects, Wiberley [16] draws similar conclusions. This means that observing this change in a given user study will probably prove extremely difficult.

Through comparing participants' behaviour to Ellis' information seeking model, we discovered that a few basic strategies had a central role in their use of DLs. Citation chaining was one key strategy that formed the common approach to finding contemporary academic research literature and, often, initial leads into archives.

Humanities academics, particularly those established in their field, used the academic community as an important source of recommendations. They also tracked the work of individual known researchers, e.g. by monitoring the researchers' personal home pages – a form of chaining. Where skills of information seeking on computers were limited, it was often because the academic network was strongly developed and minimised the need for active, independent information seeking.

Our participants believed that they were successfully locating information that answered their needs. However, it appeared that on occasion the effort required to achieve success was high. For example, user AN3 succeeded in exploring a new area of interest through an extensive use of chaining, built on an initial set of documents obtained through personal contact. However, even he described this strategy as "time consuming and exhausting".

When information seeking moved from a strongly-defined goal – e.g. 'author and title' search – into more uncertain areas such as conceptual searches related to a discipline term, problems rapidly emerged. Precise searches required the careful selection of search criteria that we only observed in a few users. Classification structures of the library could help locate information about a particular topic. However, our participants seldom used classifications to browse in DLs, even though they often reported browsing in physical libraries. Furthermore, only two mentioned using classifications when searching.

Future work is needed to identify techniques to better support less experienced users by assisting them in selecting appropriate search criteria (e.g. use of fields and

phrases). Similarly, setting default search criteria to emphasise precision over recall should improve the results when naïve search criteria are used with discipline term queries (Google takes this approach). Support for chaining in many DL systems (including the ones studied) is poor. Given the significance of this approach for humanities academics, better citation chaining tools should considerably improve their experience of DL systems.

6 Acknowledgements

This work was supported by EPSRC Grant (GR/S84798). We would also like to thank the study participants and the University of Waikato for its support of the study.

References

1. A. Adams and A. Blandford. Digital libraries in a clinical setting: Friend or foe? Procs. European Conference on Digital Libraries, LNCS 2163, pages 213–224, 2001.
2. M. J. Bates. The Getty end-user online searching project in the humanities: Report no. 6: Overview and conclusions. College & Research Libraries, 57:514–523, November 1996.
3. G. Buchanan, A. Blandford, M. Jones, and H. W. Thimbleby. Spatial hypertext as a reader tool in digital libraries. In Visual Interfaces to Digital Libraries, pages 13–24. Springer-Verlag, 2002.
4. G. Buchanan, M. Jones, and G. Marsden. Exploring small screen digital library access with the greenstone digital library. In ECDL '02: Proceedings of the 6th European Conference on Digital Libraries, pages 583–596. Springer-Verlag, 2002.
5. D. Ellis. A behavioural model for information retrieval system design. Journal of Information Science, 15(4/5):237–247, 1989.
6. D. Ellis and M. Haugan. Modelling the information seeking patterns of engineers and research scientists in an industrial environment. J. of Documentation, 53(4):384–403, 1997.
7. I. Hsieh-Yee. Effects of search experience and subject knowledge on the search tactics of novice and experienced searchers. JASIS, 44(3):161–174, 1993.
8. B. J. Jansen, A. Spink, and T. Saracevic. Real life, real users, and real needs: a study and analysis of user queries on the web. Inf. Process. Manage., 36(2):207– 227, 2000.
9. J. Koenemann and N. J. Belkin. A case for interaction: a study of interactive information retrieval behavior and effectiveness. In CHI '96: Proceedings of the SIGCHI conference on Human factors in computing systems, pages 205–212. ACM Press, 1996.
10. W. T. Lucas and H. Topi. Form and function: The impact of query term and operator usage on web search results. JASIST, 53(2):95–108, 2002.
11. S. Siegfried, M. J. Bates, D. Wilde. A profile of end-user searching behavior by humanities scholars: The Getty Online Searching Project Report #2. JASIS, 44(5):273–291, 1993.
12. H. R. Tibbo. Primarily history: historians and the search for primary source materials. In JCDL '02: Procs. Joint Conference on Digital libraries, pages 1–10. ACM Press, 2002.
13. R. Watson-Boone. The information needs and habits of humanities scholars. RQ, 34(2):203–216, 1994.
14. E. Whitmire. Disciplinary differences and undergraduates' information-seeking behavior. J. Am. Soc. Inf. Sci. Technol., 53(8):631–638, 2002.
15. S. E. Wiberley and W. G. Jones. Patterns of information seeking in the humanities. College and Research Libraries, 50:638–645, 1989.

16. S. E. Wiberley and W. G. Jones. Time and technology: A decade-long look at humanists' use of electronic information technology. College & Research Libraries, 61:421–431, 2000.
17. B. M. Wildemuth. The effects of domain knowledge on search tactic formulation. Journal of the American Society for Information Science and Technology, 55(3):246–258, 2004.
18. I. H. Witten, A. Moffat, and T. C. Bell. Managing gigabytes: compressing and indexing documents and images. (second edition). Morgan Kaufmann, San Francisco, CA., 1999.