Subject searching requirements: the HILT II experience

Emma McCulloch
Ali Shiri and
Dennis Nicholson

The authors

Emma McCulloch is a Researcher and **Dennis Nicholson** is the Director, both at the Centre for Digital Library Research, University of Strathclyde, Glasgow, UK.
Ali Shiri is Assistant Professor at the School of Library and Information Studies, University of Alberta, Canada.

Abstract

The HILT Phase II project aimed to develop a pilot terminologies server with a view to improving cross-sectoral information retrieval. In order to inform this process, it was first necessary to examine how a representative group of users approached a range of information-related tasks. This paper focuses on exploratory interviews conducted to investigate the proposed ideal and actual strategies of a group of 30 users in relation to eight separate information tasks. In addition, users were asked to give examples of search terms they may employ and to describe how they would formulate search queries in each scenario. The interview process undertaken and the results compiled are outlined, and associated implications for the development of a pilot terminologies server are discussed.

Introduction

The HILT (High-Level Thesaurus) Phase II project (HILT, 2003) run by the Centre for Digital Library Research at the University of Strathclyde in Glasgow was funded by the Joint Information Systems Committee (JISC) to set up a pilot terminologies server for the JISC Information Environment (IE) (JISC, 2003), aiming to:

- provide a practical experimental focus within which to investigate and establish subject terminology service requirements for the JISC IE; and
- make recommendations as regards a possible future service.

In order to inform the HILT model in general terms and also to pick up on specific points relevant to the design of the pilot terminologies server, HILT sought to investigate users' information-seeking behaviour through a series of one-to-one interviews.

This paper reports on and analyses the results of these interviews and considers their significance. The interviews were designed to provide information on the nature of users' subject search requirements, their willingness to consult a range of collections to find information for particular tasks, the level of specificity of the search terms they tend to use, and the mix of search strategies they are likely to employ.

The following questions were considered in studying users' general search behaviour:

- How do users formulate search queries?
- Are they aware of different search techniques – phrase searching, Boolean, truncation and so on?
- Do they know how to use these techniques effectively?

It was considered of great importance to investigate the needs of users with varying levels of search experience. As such, HILT studied a range of users with different degrees of expertise; students of various levels (HND, undergraduate degree, MSc and PhD), lecturers, and intermediaries such as librarians and electronic information service staff. It was thought that different types of user would employ different search strategies and techniques and that they would also have distinct aims at the outset of the exercise.

For example, it was thought that students would try to retrieve material of direct relevance to the task, perhaps by taking a broad sweep of available

resources in order to pass a piece of coursework. In contrast, it was thought that intermediaries would conduct a more in-depth search to retrieve a wide variety of sources to provide a more balanced view in relation to the task.

The selection of participants also attempted to achieve a mix of users from different subject backgrounds. It has been claimed that "as the search specificity increases the need for effective search strategy becomes more critical" (Debowski, 2001). Bilal (2000) has also reported differences in search strategies according to type of task. Within the context of HILT, it was thought that search techniques employed by users studying general subjects at a broad level such as HND would differ from those undertaking, for example, a PhD in an extremely specific area. To assess this further, interviewees were presented with eight tasks covering different subject areas with varying levels of specificity.

Data was gathered on the subject terms users were likely to employ in relation to specific information tasks. The main purpose here was to investigate the range of terms that users with different searching ability and subject backgrounds may select in relation to particular topics. Past research (Debowski, 2001) highlights the errors of novice searchers whom she claims commonly display "incorrect choice of search terms" and "inappropriate mixing of concepts in one inquiry". In addition, it has been claimed that less advanced users typically employ single terms rather than co-linked or compound terms (Brown, 1995). It was hoped that this data would provide an insight into the variation of terms favoured by users and give some indication of whether or not users' preferred terms tend to appear in standard subject schemes. This part of the exercise also served to inform the mapping element of the pilot terminologies server, providing guidance on the level of granularity at which terms from different subject schemes should ideally be mapped.

It was of interest to HILT to investigate the types of search strategy employed by users and their levels of competency in doing so. This was an important aspect of the study as the pilot terminologies server must be designed to cope with commonly used search techniques. Debowski (2001) has looked at specific techniques employed by users suggesting that "the use of Boolean connectors underpins successful database searching, with extensive use of conventions such as 'and', 'or' and 'not' to help refine the search". In contrast, Ford *et al.* (2002) found, in relation to Web-based information, that "retrieval effectiveness was associated positively with best-match searching and negatively with Boolean searching". Supporting this view of Boolean

operators, Jansen (2000) claims that "people did not feel comfortable using them". Further work by Holscher and Strube (2000) uncovered differences in techniques according to type of user claiming that "Web experts make use of advanced search options like Boolean operators, modifiers, phrase search etc., much more frequently than the average user". They also claim that "A noteworthy exception is the "+" operator. It is equally popular among the general public, making it the most important query formatting tool for non-expert users".

The HILT study hoped to consider which techniques specific users are aware of and which they tend to employ in relation to different types of task. The claims that different techniques were more effective in specific environments have implications for HILT in terms of the way a terminologies server is built into the JISC IE. Will it be Web-based or part of a collections database, for example?

Methodology

The HILT team investigated these issues within a one-to-one interview setting, guided by the use of a structured questionnaire, as it seemed unlikely that user behaviour such as decision making would be captured in a more practical hands-on setting. It was first thought that observation or the use of screen capture software may have been required to capture users' online search behaviour but in the context of HILT, it was essential to tap into users' thought processes. Transaction logging techniques are valuable for collecting data on user interaction with a system but do not provide any data on users' reasoning. For example, they do not address the complex processes users undergo when considering terms and deciding on how best to formulate their search strategies.

The methodology adopted within the HILT project ensured that the interviewer was able to question what types of information users would look for in relation to a specific task, why the user would choose particular terms, how he/she would combine such terms, and where users typically look for information within their specific subject discipline(s).

The interview approach enabled HILT researchers "to adjust the pace and style of asking questions so as to bring out the best in the respondents" (Hannabuss, 1997). This was crucial due to the varying levels of search ability and experience within the user group. Findings were enhanced by the fact that interviewers were

able to follow up ideas and probe users' reactions further (Baines, 1997).

Recruitment of users

Users were recruited primarily through e-mail. Since the study required specific types of user many were targeted directly. Contact details were located through university departmental Web pages and staff lists.

In addition to this, advertisements were posted in local college libraries to attract HE students and lecturers. Payment was offered to undergraduate students as an incentive to participate.

Tables I-III illustrate the distribution of users recruited.

Table I Distribution of users by institution

Institution	
Caledonian University	6
Glasgow College of Building and Printing	4
Glasgow College of Commerce	1
Napier University	5
Public Records Office	1
Scottish Library and Information Council	1
Strathclyde University	9
University of Sheffield	3
Total	*30*

Table II Distribution of users by status

Status	
BA student	2
HND student	4
MSc student	7
PhD student	4
Lecturer	2
Researcher	5
Intermediary/information professional	6
Total	*30*

Table III Distribution of users by subject area

Subject area	
Medical/biological/applied sciences	9
Library intermediary/information professional	6
Construction and architecture	4
Library and information science	3
Finance	3
Business administration	1
English	1
Fashion marketing	1
Psychology	1
Publishing	1
Total	*30*

Data gathering techniques

Questionnaire design
The questionnaire set out eight information tasks requiring varying degrees of search activity and different methods of query formulation.

Interviews
Interviews were carried out on a one-to-one basis using a structured questionnaire and response recording form. The interviewer briefed each user with a standard scenario prior to any questioning:

> You have been asked to find information for [tasks 1-8, in turn, to be inserted here]. You are told that the library has paid so that you can have free access over the Web to six services with different content that each have relevant information.

Task 1 asked users to find information for writing an essay on the current UK status of their own individual subject area. For example "write an essay on the current status of fashion marketing in the UK".

Question 1 then asked users to state which of the strategies 1-6 (below) they would adopt in an ideal situation i.e. one devoid of time constraints, conflicting deadlines and so on, to find material of value to this task.

(1) Choose one at random or the one you are most familiar with and study only the material from that service.
(2) Look at all of the services but study some material from a couple of them in depth.
(3) Use all of them, identify all relevant resources and study all in depth.
(4) The minimum required to ensure a reasonable grade.
(5) Some other variation of the above.
(6) Something else? Please specify.

Question 2 asked which of the strategies (again, selecting from 1-6) they would actually adopt in practice when undertaking the task.

Question 3 required users to give some examples of subject terms they would use when searching for the information.

Finally, question 4 asked how they would enter their search, if they would enter a single term or if they would combine terms in any way (and if so, how?).

Below is listed the first task, followed by a further seven tasks:

(1) Finding information for writing an essay on the current UK status of their own individual subject area.
(2) Compiling a bibliography on publishing techniques.
(3) Finding a specific book about Robert Burns.
(4) Identifying key articles on the history of architectural conservation.

(5) General study of journalism software.
(6) Preparing for a test on statistical methods/tests.
(7) Preparing for a discussion-based tutorial on article writing.
(8) Planning a presentation to your tutorial group on poster design.

Questions 1 and 2 were designed to work in tandem. It was hoped that if participants were asked both what they felt they should do in the circumstances described, and what they thought they would actually do in practice, they would be more likely to give an honest assessment of probable actual behaviour when answering question 2 (while at the same time indicating what they thought the ideal approach would be).

Analysis of results

Data gathered from interviewees was loaded into an Access database to facilitate comparisons between ideal and actual scenarios as quoted by different user groups, subject areas, and search terms given.

Results and discussion

Results proved inconclusive in terms of statistical significance between user groups, subject areas and search terms given. However, the following results are of interest and serve to highlight a number of problems relating to user behaviour and associated implications for the subsequent design of the terminologies server.

Table IV illustrates that the highest proportion of participants in all user groups claimed the ideal strategy for retrieving material relevant to writing an essay on the status of their own subject area in the UK was strategy 2 – look at all of the services but study some material from a couple of them in depth (selected by seven students, three researchers, one lecturer and four intermediaries; a total of 15 users out of 30). Strategy 2 was also the most popular actual strategy selected by intermediaries. Of the intermediaries, four of the

six viewed strategy 2 as ideal with three claiming they would adopt this approach in practice. This suggests that intermediaries are willing to undertake a broad search of resources and consult more than one source.

However, in practice, students claimed they would behave quite differently to their ideal situation with eight opting to do something different (strategy 6), explaining that they would normally consult a single known source only. This suggests that, in general, students are less willing than intermediaries to conduct a thorough search of resources. This result seems to support Debowski's (2001) claim that users of different levels tend to employ different search strategies. Although seven out of 17 students viewed strategy 2 as the most effective, only four claimed they would do this in practice. Eight out of 17 claimed they would opt for strategy 6 in practice; an option only viewed as ideal by one. In contrast, none of the intermediaries interviewed saw strategy 6 as a viable strategy in either the ideal or actual scenario.

Looking more widely at the data collected, there is clear variation in the ideal strategies chosen for different tasks. For task 1, 24 of 30 users (80 per cent) claimed strategies 2 or 3 were the ideal approaches. This dropped to 60 per cent for tasks 2 and 4 (both 18/30). For tasks 6 and 8 this fell to 16/30. For tasks 5 and 7 13/30 users claimed strategies 2 and 3 were ideal. In the case of task 3 (finding a specific book about Robert Burns) only 7/30 users claimed strategies 2 and 3 were ideal. In contrast, the majority (12/30) thought strategy 6 was the ideal approach in this particular scenario. This variation in strategy selection between tasks supports Bilal's (2000) claim that different search strategies are evident according to the type of task presented.

When asked what they would actually do in practice in relation to each of the tasks, a similar pattern of variation emerged. While 12 users said they would adopt strategy 1 to tackle task 4, only one person selected this option for tasks 1 and 8 with the majority opting for strategy 6 in these cases.

Table IV Number of users quoting strategies 1-6 as their ideal and actual strategies, by group, for task 1

Group Strategy	Students Ideal	Actual	Researchers Ideal	Actual	Lecturers Ideal	Actual	Intermediaries Ideal	Actual	Total Ideal	Actual
1	3	1	0	1	0	0	0	0	3	2
2	7	4	3	3	1	1	4	3	15	11
3	4	3	2	1	1	0	2	2	9	6
4	1	1	0	0	0	0	0	1	1	2
5	1	0	0	0	0	1	0	0	1	1
6	1	8	0	0	0	0	0	0	1	8
Total	17	17	5	5	2	2	6	6	30	30

Table V shows the number of participants, by group, who selected the same strategy for both their ideal and actual approaches in relation to each of the tasks. It also shows the number of users who claimed to do something different in practice, compared with their view of what the ideal approach would be in each case.

Data shows that a total of 16 users (eight students, three researchers, one lecturer and four intermediaries) gave the same strategy for their ideal and actual strategies in relation to the tasks. A similar number, 14 (nine students, two researchers, one lecturer and two intermediaries) selected different strategies as their ideal and actual approaches. Some of the reasons quoted for inconsistency between ideal and actual strategies include:

- user would automatically go to services he is familiar with (intermediary);
- user would adopt a wide mix of methods depending on his timescale (student);
- user claimed he wouldn't use such services in practice; he would use the library OPAC or Google (student); and
- user claimed she would go directly to her library OPAC in practice.

The most well-defined task presented to users was to find a specific book about Robert Burns (task 3). In this example, 17 of the 30 users claimed they would enter "Robert Burns" as their search string. One further user suggested "Burns, Robert" and three users said they would simply enter "Burns". The remaining six users claimed they would employ a more complex strategy. For example, one toxicology researcher claimed she would enter "Robert Burns *and* Scotland *and* poet". Another participant claimed he would use the author, title or ISBN, if known, while two users were unable to provide specific terms for this task.

The less well-defined tasks elicited a broader range of subject terms from users. That is, there was a greater degree of variation between terms given by participants where key terms were not obvious from the task itself. For example, the

question asking which terms users would employ to find information to prepare for a discussion-based tutorial on article writing resulted in a wide variety of terms. In sharp contrast to the previous example where 17 users quoted the same terms, only six users out of 30 quoted the same search string: "article writing". Other suggestions included "business writing", "good article writing; proper English", "how to write articles", "report writing", and "writing style". Since the level of effectiveness of search terms and strategies was not assessed within HILT – that is, terms were not physically searched for nor subsequent results evaluated due to funding and time constraints – we have no evidence to support or refute Debowski's (2001) or Ford *et al.*'s (2002) claims on the effectiveness of search techniques in different search environments.

Throughout the entire study, it was found that phrase searching was employed most frequently (78/208) followed by the use of synonyms and alternative variants (42/208). Boolean techniques were evident in 37 of the 208 responses obtained, but the distribution pattern of these responses meant that at least half of the members of three of the four user groups employed Boolean to some extent in their searching (50 per cent of lecturers, 80 per cent of researchers, 67 per cent of intermediaries and 29 per cent of students displayed Boolean techniques). Data indicated, therefore, that lecturers, researchers and intermediaries tended to use Boolean techniques more frequently than students, who were the one group of the four in which a minority of searchers used Boolean. Thus, Jansen's (2000) claim that users are uncomfortable employing Boolean operators was not supported by HILT data.

The study did provide some evidence in favour of Holscher and Strube's (2000) findings that experts employed Boolean techniques more often than inexperienced users but did not support their claim that the inexperienced users tend to rely primarily on the "and" operator on occasions where they do use Boolean. In fact, across all tasks

Table V Number of users quoting the same/different strategies 1-6 as their ideal and actual strategies, by group, across all tasks.

Group	Students		Researchers		Lecturers		Intermediaries		Total	
Task	Same	Different	Same	Different	Same	Different	Same	Different	Same	Different
1	8	9	3	2	1	1	4	2	16	14
2	11	6	2	3	1	1	3	3	17	13
3	12	4	3	2	2	0	6	0	23	6
4	9	7	1	4	1	1	3	3	14	15
5	11	6	3	2	2	0	3	3	19	11
6	10	7	2	3	2	0	3	3	17	13
7	9	8	1	4	2	0	4	2	16	14
8	10	7	3	1	2	0	4	2	19	10
Total	80	54	18	21	13	3	30	18	141	96

within the HILT study, each instance of Boolean searching used the "and" operator with two exceptions, both of which were proposed by students. The "or" operator was used in two single instances; once by an information management MSc student and once by a BA student in applied graphics and technology. On 34 occasions throughout the study, no response was given, that is, no terms were suggested. Only 12/208 responses comprised single term searches. Although Brown (1995) claimed that inexperienced searchers tend to rely on single rather than compound terms, the HILT results do not provide support for this. Although students opted for single term searches on a greater number of occasions than the other user groups, the figures were not conclusive. Finally, five responses demonstrated the use of inverted search terms such as "Burns, Robert".

A number of caveats in the HILT methodology should be noted and considered by those undertaking future research in the area of user search behaviour. Specifically:

- The study involved a small sample with unequal numbers in different user groups. It was felt that the study would have been more effective if equal numbers of student, teacher and professional user groups had been recruited. However, a poor response was received from lecturers and information professionals, making this impossible within the project's timescale.
- An assumption was made that students were inexperienced/novice searchers compared with the other user groups recruited (researchers, lecturers and intermediaries). Although the student group is likely to be less experienced than the intermediary group, no assessment of existing search skills was made.
- Tasks were randomly selected. Findings may have been more conclusive had established tasks from previous research been used.
- Interviews were conducted in a non-practical setting. Search refinement and learning effects would be likely to occur in a hands-on setting.
- No attempt was made to assess the effectiveness of search terms proposed by users or the value of search techniques they claimed they would employ.
- No formal statistical analysis was possible due to unequal numbers of user groups and disproportionate numbers involved in different subject areas.

Conclusion

Interviews conducted within the HILT project served to provide useful information about user search behaviour in general. The following summary of findings can be reported:

- strategies vary according to user group;
- strategies vary according to task;
- greater overlap in subject term selection is evident for well-defined tasks; and
- a variety of search techniques are adopted by users including Boolean, combination, single term and free text searching.

In addition, results from the user interviews informed the technical element of the project: the construction of a pilot terminologies server. Several conclusions regarding its design were reached. An effective terminologies server needs to cater for all types of user incorporating:

- broad subject coverage;
- simple and advanced search facilities;
- a range of subject schemes to account for variation in user terms;
- non-standard terms – many user terms proposed throughout the interviews do not appear within standard schemes – along with a mechanism of mapping these to existing standard terms; and
- both general and detailed levels of granularity in term mapping.

User behaviour, as outlined here, should be noted by those developing Web-based search systems, particularly those with a focus on terminology research. To develop findings from the present study it would be useful to actually search for terms proposed by participants within a controlled environment and assess the value of these terms for retrieving material of relevance to each task. In this way, a more informed evaluation could be made relating to term effectiveness in terms of precision and recall.

However, for the purposes of the HILT study, the exploratory interviews described here provided researchers with an insight into how users typically approach a range of search-based tasks. Differences in the strategies adopted were evident between user groups and type of task undertaken. Valuable information was also gleaned on the types of terms users employ and how they manipulate these into search strings. Such findings have proven invaluable in the design of the pilot terminologies server.

References

Baines, S. (1997), "End-user searching behavior: considering methodologies", *The Katharine Sharp Review*, No. 4, available at: http://alexia.lis.uiuc.edu/review/winter1997/bains.html (accessed 27 May, 2004).

Bilal, D. (2000), "Children's use of the Yahooligans! Web search engine: 1 Cognitive, physical and affective behaviors on

fact-based search tasks", *Journal of the American Society for Information Science*, Vol. 51 No. 7, pp. 646-65.

Brown, M.E. (1995), "By any other name: accounting for failure in the naming of subject categories", *Library and Information Science Research*, Vol. 17 No. 4, pp. 325-429.

Debowski, S. (2001), "Wrong way: go back! An exploration of novice search behaviours while conducting an information search", *The Electronic Library*, Vol. 19 No. 6, pp. 371-82.

Ford, N., Miller, D. and Moss, N. (2002), "Web search strategies and retrieval effectiveness: an empirical study", *Journal of Documentation*, Vol. 58 No. 1, pp. 30-48.

Hannabuss, S. (1996), "Research interviews", *New Library World*, Vol. 97 No. 5, pp. 22-30.

HILT (2003), available at: http://hilt.cdlr.strath.ac.uk/ (accessed 26 May 2004).

Holscher, C. and Strube, G. (2000), "Web search behavior of Internet experts and newbies", paper presented at the 9th International World Wide Web Conference, Amsterdam, 15-19 May, available at: http://www9.org/w9cdrom/81/81.html (accessed 27 May 2004).

Jansen, B.J. (2000), "The effect of query complexity on Web searching results", *Information Research*, Vol. 6 No. 1, available at: http://InformationR.net/ir/6-1/paper87.html (accessed 27 May 2004).

JISC (2003), available at: http://www.jisc.ac.uk/ (accessed 27 May 2004).