

# Patterns of interactions: user behaviour in response to search results

**Hanna Stelmaszewska**

UCL Interaction Centre (UCLIC)  
26 Bedford Way  
London, WC1H 0AB, U.K  
+44 20 7679 5315  
h.stelmaszewska@ucl.ac.uk

**Ann Blandford**

UCL Interaction Centre (UCLIC)  
26 Bedford Way  
London, WC1H 0AB, U.K  
+44 20 7679 7557  
a.blandford@ucl.ac.uk

## ABSTRACT

This paper presents patterns of users' interaction when working with digital libraries. It focuses on strategies developed and applied by users over time to achieve their goals. Results show that users choose different patterns of interaction depending on their evaluation of results, particularly in terms of the number of results returned from a search. This study gives indications about how the user interface could better support users in developing different search strategies.

## Keywords

Interaction, Patterns, Information Seeking, Information Searching, Digital Library, HCI, Video Protocols

## INTRODUCTION

Designing usable digital libraries is difficult and requires effective tools that are based on proven knowledge of design. One key aspect of use is the patterns of interaction between user and system. Patterns, using that term in a specialized sense, have emerged as a possible solution to some of the problems within many disciplines including software engineering and Human-Computer Interaction. Alexander [1] first introduced patterns in the 1970's in the world of architecture. He used them to explain how a collection of architectural design patterns can be identified so buildings and an urban environment could be more usable and pleasurable for their inhabitants. This idea of using patterns has since been incorporated in many other disciplines. Within software engineering, the pattern concept has been applied to facilitate software reuse [7] while Bayle *et al* [2] adopted the idea of patterns for interaction design within HCI. Van Welie and Traetteberg [13] discuss successful solutions in user interface design. These patterns come from understanding existing, established patterns that are successful in order to codify patterns for future design. Digital library usage is not yet well enough understood for us to propose design patterns in this sense, so we are starting by studying the emergent patterns of behaviour; understanding these is a precondition for developing future successful design patterns for digital libraries.

An earlier study [3] analysed how a particular interaction between a user and digital libraries evolves over an extended period of time; here, we focus on the detailed strategies users apply when looking for specific material and also (at a higher level) how they build up their search strategies.

The patterns of users' interactions presented in this paper concentrate on phases of interaction where users change strategy when seeking information; this is a central task

for many activities in the electronic environment [10]. This knowledge could help designers to provide users with appropriate feedback and support them during the information seeking strategy development process. Existing systems do not give users much support during the decision making process when results are evaluated; a better understanding of this problem can be used to guide development of solutions for users that better support their information seeking.

## METHOD

The study aimed to achieve a better understanding of what users do and how they develop information searching strategies when looking for a specific topic within different digital libraries.

A video-based observational study with a think aloud commentary was used to monitor seven users working with several digital libraries in order to achieve their personal objectives. Six of them concentrated on finding material on one topic, while the seventh looked for articles on four different topics. Detailed information about the users is presented by Blandford *et al.* [3], but in summary, three were first year PhD students, one a final year PhD student, one an administrator with a particular specialist interest (in an aspect of healthcare), one a research assistant and one an academic; although all users were experienced at searching the Web, most had limited experience of using digital libraries and none had strong information retrieval skills. The three first year PhD students and administrator were relative novices in their domains of study (compared to the other three users). Users were videoed as they searched in various digital libraries (accessible via 'bookmarks' in the browser window) for material of their own choosing. Transcribed video data was analysed by applying high level concepts of themes that represent activities of the users and systems, comparing the data to the models of information seeking proposed by Kuhlthau [9], Sutcliffe & Ennis [12] and Marchionini [10], and also identifying key patterns within the interactions exposed within the data.

## RESULTS

As has been found by others [4, 8], at a low level, there are two distinct patterns of users' behaviour when information seeking using digital libraries: searching and browsing. The choice of information seeking strategy depends on users' familiarity with the collection and their knowledge of the domain in question. Searching is the most frequently applied strategy for finding interesting articles in digital libraries. It does not appear as a 'one-

shot query' [6], but instead is an iterative process, during which users reformulate a query, change user preferences or change query terms looking for satisfactory results.

Online searching not only involves input of various terms, but also depends on the ability and experience of the individual person performing a search [5]. The digital library user needs to learn how to use the query language, and which strategies to use in a specific online environment. A good search strategy develops over time and it requires not only an understanding of a searching paradigm, but also knowledge of the task domain. Our data shows that users start their search from a very simple query using keywords. Later, they apply more sophisticated searches which involve changing preferences such as: limit a source (publication), date of publication, or use of a 'search within field' option (full text, abstract, title or author). In this paper, we present changes of users' searching strategies triggered by their evaluation of results.

We present the discussion according to the stages of Kuhlthau's model of the information search process [9], which is of students working on an assignment in a physical library. Of the models known to us, this one most closely matched the stages of activity our users went through – probably because they, like Kuhlthau's subjects, were relative novices. However, to expand to the electronic environment of this study, one additional stage is taken from the models of Marchionini [10] and Sutcliffe & Ennis [12] (the stage is common to both models) – namely examining results. Thus, we consider seven stages: initiation; selection; exploration; query formulation; results examination; document collection and results presentation.

### **Initiation**

Problem definition is a crucial point in the information searching activities. In this early stage of information searching, much of the user's work evolves around identifying the initial goal or information need. In our study the problem was partly defined by the study – in that users were asked to find information relevant to their research – and partly by the user, who specified their own topic of interest.

### **Selection**

Working in a familiar environment can significantly contribute towards successful search [12]. As noted above, users were given easy access (through bookmarks) to several digital libraries; the order of these was modified for each user (so that if users just worked through libraries from the top of the list, they would naturally access them in different orders). Some of the users in this study clearly showed their preference for working with familiar resources. For example, user D reported: "I've used this library in the past so more or less I know that I can find relevant papers." Other users were not familiar with any of the resources, and therefore appeared to choose at random initially.

### **Exploration**

Although searching is a common activity within digital libraries, users generally started their interaction with a phase of system familiarisation. They viewed the content and features of every new collection they used, looking for some help and clues on creating an efficient strategy

for information seeking. As user B commented: "It's a new search engine, I am looking to understand it". Six out of seven users went through this phase of looking for information about how to search the library, how to use it effectively and the ways of viewing the collection. One user, an experienced academic, was already familiar with the digital libraries she chose to interact with, and skipped this phase.

### **Query formulation**

The users' primary activity involved formulating queries using vocabulary that articulated their topics of interest. As noted by Marchionini [10] and Sutcliffe and Ennis [12], the query formulation activity depends on the user's domain expertise, the user's skills in generating queries and the system's information retrieval paradigm. Our findings support this view: the users with greater domain expertise used their knowledge to create and extend more complex query terms. The number of terms used varied depending on the user's domain knowledge and their objectives. Experienced users (in the domain) used more specific terms and added more of them whereas novices (in the domain) used more general ones to obtain any material. Moreover, our study showed the four novice (domain) users repeated the initial query terms when working with consecutive digital libraries, whereas the users with greater domain expertise worked in a more variable way. One user changed query terms every time he started interacting with a new digital library; one reformulated query terms as she worked with different collections; one had four different objectives to achieve so she changed query terms every time she worked with a new collection (although, broadly, she searched for the same four topics in each library).

### **Results examination**

Once results are retrieved the user needs to decide whether to view them or continue searching for more appropriate documents. An evaluation of results strongly influences which search strategy will be applied next to obtain more satisfactory results [11, 12]. In our study, three kinds of results were found to affect the users' interaction and their decision about what steps to take next: too many results, 'no matches' and a manageable number of matches. In the discussion that follows, we use the term 'query reformulation' to describe any changes made to the original query terms; that is: keep some original terms, add some or delete some. The 'change of user preferences' refers to 'search within field', restricting source or date, and setting Boolean OR and AND operators. The phrase 'change query terms' describes a search using new query terms.

We compare our findings to Sutcliffe and Ennis' [12] framework; although their framework covers a wide spectrum of users' information searching behaviour, our study discovered some discrepancies between their model and the behaviour of our users.

#### *Results of 'no matches'*

The results of 'no matches' led users to change query terms, reformulate a query, change their preferences or abandon the library; users had low persistence: the maximum number of no matches ever obtained was three, before users moved on to work with a different library.

According to Sutcliffe and Ennis [12], users' response to 'too few' results is (or should be) reformulation of the query. They define query reformulation, in the case of receiving too few results, as reducing the number of search terms, introducing disjunctions (ORs), applying stemming to search terms, adding synonyms and removing related terms, or substituting general for more specific terms. They also suggest other strategies such as use of controlled vocabulary, or thesauri. In practice, users from our study preferred to change query terms, and would not persist for many reformulations. Without system support to guide the user on how to interpret results and how to use these results to improve the search strategy, users are left very much on their own trying to find out what they did 'wrong'. Was it an insufficient query term vocabulary or perhaps a limitation of the collection? Some of the users blamed themselves for obtaining 'no matches' results; for example, user C stated: "I haven't found any matches. Perhaps I shouldn't put trans... publications. So I will search in 'all journals and proceedings'. No, 'journals only'".

Changing user preferences to try to improve results was a common activity among our users. The example presented above shows user applying a different search strategy – in this case, extending the source from a very limited search to a less limited one. In other cases users changed the 'search within field' option or changed the restriction on the date of publication, looking for some improvements.

However, the most telling point is that users frequently gave up the search very quickly. Since digital libraries are discretionary resources (users typically do not *have* to use any particular library), there is an urgent need for users to be able to learn quickly how to conduct effective searches, and have the confidence to persist.

#### *Results of 'too many'*

In the case of too many results, the emerging pattern was to perform three different activities: query reformulation, changing preference settings and changing query terms, in that order. Sutcliffe and Ennis [12] propose that in the same situation, the user's strategy should be adding more search terms, replacing ORs with ANDs, negations (NOTs), reducing synonyms, adding more specific and related terms to improve search results. It appears that Sutcliffe and Ennis are describing what users should do, rather than what novice users actually do.

Although too many results would not appear to pose a problem to users, they tried to understand why the number of results was so high and expressed doubts about their abilities to construct effective queries. One user reported: "So I am confused. Why I got this results while searching for e-commerce. Maybe I have chosen or ...I wrote a misleading keyword." Another user commented: "You search something as a keyword, it always find article, but if you search in the full text it should find too many things". The same user also reported: "I've just try one keyword 'classification' and found too many things. This time I'll try to find something about 'growing grid'". Here again our users were facing a dilemma about what search strategy to use, and changing query terms seemed like a good idea after trying to change user preferences.

As mentioned earlier, changing user preferences was another strategy users applied. The choice of preferences did not depend only on users' understanding of this strategy, but also on library features that were offered. For example user D changed Boolean operator 'OR' for 'AND': "Here I found one thousand seven hundred twenty two commerce and two thousand...OK, I will have to go back and say 'all'." Another user changed the 'Search in Field' option to reduce the number of results.

Some users worked systematically through up to three pages of results from a large results set, whereas others appeared to be overwhelmed, and quickly moved on to a further search.

#### *Results of 'OK number'*

When a manageable number of results were returned, the emerging pattern was to scan results and examine selected results in detail. This was consistent with Sutcliffe and Ennis's [12] model. The *scanning* involves viewing and making decisions about the relevance of search results, while *examining in detail* involves looking into individual results in more depth, trying to establish how relevant the document is for the topic in question. Two emergent patterns were observed when establishing the relevance of the results. The novice users (in the domain) relied on the system's assessment of the relevance; for example, one novice user commented: "The relevance is already too low so I will stop looking at this list.". The more experienced users (in the domain) made their judgement based on the information included in the title, abstract, index terms or keywords. For example one of the experienced user commented: "Actually, I'm not paying the attention on the relevance presented by ACM DL. I'm trying to see the relevance from the title and where I can find them".

#### *General issues*

As outlined above, the users' search strategies frequently unsuccessful, and users commonly felt that the system did not support them with building an understanding that would help them apply an appropriate strategy for further searching. Instead, it provided users with more 'puzzles'; for example, one of the users commented: "If I input several words, keywords, so what's the relation between these? I mean, is it 'or' or 'and'?". Another user commented while searching for three keywords: "I don't know what is the meaning of the word 'all'. It should include three words but it seems...". These examples indicate the lack of constructive system feedback, which would indicate alternative ways to perform a search.

In summary, none of the digital libraries used in the study (all of which were widely available, representing the current state of the art) gave adequate feedback, either about system state or about future possibilities, such that users could develop skills and confidence in query formulation.

#### **Document collection**

The next stage of the search process revolved around extracting and storing information. Within our study there were two emergent patterns users followed: saving or printing. Each of the users managed to save at least one document. Two users decided to print the relevant documents. The reason for doing this for one user was that he had difficulties finding a 'save' option within

Acrobat Reader and after 79 second searching he reported: "I give up and print". The other user did not want to wait for a long document to be downloaded, so she printed it. As she said: "Gosh it's big. If I'm going to read it then I'm going to need it printed so... let's just print it and be done."

### Results presentation

This stage of Kuhlthau's model [9] was irrelevant to this study as our users simply took documents (as saved files or printouts) away with them.

### DISCUSSION

The previous section focuses on various stages of activity, adopted from Kuhlthau's [9], Sutcliffe & Ennis's [12] and Marchionini's [10] models of information searching. Although Kuhlthau's model discusses important patterns of students' activities, our data is not always consistent with them. We found that users are active participants in the information search process and they develop their search strategy, which is consistent with Kuhlthau's model. However, she puts great emphasis on exploring and identifying a topic of interest within the stages of initiation, selection and exploration, whereas our users concentrated on these issues only during the first, initiation, stage. The remaining two stages they dedicated to make decisions about which resources to use. Furthermore, Kuhlthau's model does not incorporate evaluating and analysing the results: it focuses on the search process, which aims to find information rather than evaluating it. These activities were the core ones observed in our study, as they defined for the user what strategy to apply next in order to find interesting results. This contrast arises largely out of the different circumstances in which the two studies were conducted.

### CONCLUSIONS

The cognitive theory proposed by Sutcliffe & Ennis [12] provides predictions of search strategies for experts and novice users. However, our study shows that there are some discrepancies between activities described in their model and patterns of users' behaviour when too many results or 'no matches' are returned. These results led to different patterns being applied by users in the hope of improving their results. Although these results might not be replicated in different user groups and different environments, this view of interaction patterns can aid in researching usability of digital libraries by providing a basis for analysing users' interactions with digital libraries environments.

Three emergent patterns of users behaviour in relation to their evaluation of results were observed. These patterns refer to the sequence of activities users perform when obtaining 'no matches', 'too many' and 'OK number' of results; that is: in the first instance users changed query terms, then reformulated a query, changed preference settings and lastly abandon the library. In the case of 'too many' results, they reformulated a query, then changed preference settings and finally changed query terms. The 'OK number' of matches triggered two activities: scanning results and examining selected results in details. Due to the lack of system support in the form of appropriate and explicit feedback, which could direct users to apply more efficient search strategies, users struggle to develop such strategies.

Patterns identified in this study hold crucial knowledge about users behaviour, which provides a starting point for understanding what users do when looking for specific information. In addition, this supplies designers with a richer, and more thorough understanding of what and when users need to progress effectively. This could help designers to build systems that would allow less experienced users to be more successful when interacting with digital libraries.

### ACKNOWLEDGMENTS

This work is funded by EPSRC Grant GR/M81748. We are grateful to the anonymous participants who took part in this study.

### REFERENCES

1. Alexander C., Ishikawa S., Silverstein M., Jacobson M., Fiksdahl-King I., and Angel S. (1977). *A Pattern Language: Towns, Buildings, Construction*. Oxford University Press, New York.
2. Bayle, E., Bellamy R., Casaday, G., Erickson, T., Fincher, S., Grinter, B., Lehder, D., Marmolin, H., Moore, B., Potts, C., Skousen, G., & Thomas, J. (1998). Putting it All Together: Towards a pattern Language for Interaction Design. *SIGCHI Bulletin*. Vol. 30, No.1(1998) pp. 17 - 24
3. Blandford, A., Stelmaszewska, H. & Brayan-Kinns, N. (2001). Use of multiple digital libraries: a case study. In *Proc. JCDL 2001*, pp. 179-188. ACM Press
4. Borgman, C. L.(2000). *From Gutenberg to the Global Information Infrastructure: Access to Information in the Networked World*. The MIT Press.
5. Efthimiadis, E. N. (1996). Williams, Martha E., ed. *Annual Review of Information Systems and Technology (ARIST), Query Expansion.V31*, 121-187
6. Furnas, G. W., & Rauch, S. J. (1998). Considerations for Information Environments and the NaviQue Workspace. *Proceedings of ACM DL '98*, pp. 79-88
7. Gamma, E., Helm, R., Johnson, J., & Vlissides, J. (1995). *Design Patterns: Elements of Reusable Object-Oriented Software*. Addison-Wesley.
8. Jones, S., McInnes, S., & Stavely, M. (1999). A Graphical User Interface For Boolean Query Specification. *International Journal on Digital Libraries 2(2-3): 207-223*
9. Kuhlthau, C. (1988). Longitudinal case studies of the information search process of users in libraries. *Library and Information Science Research*. 10.3. pp. 257-304.
10. Marchionini, G (1995). *Information seeking in electronic environments*. Cambridge University Press
11. Shneiderman B., Byrd, D. & Croft W.B. (1998). Sorting out Searching. *Communications of ACM*. 41.4.
12. Sutcliffe A., & Ennis M. (1998). Towards a cognitive theory of information retrieval. *Interacting with Computers*, 10 pp. 321-351
13. Van Welie, M. & Traetteberg, H. (2000). Interaction Patterns in User Interfaces. In *Proc. PloP 2000*.

