An Approach to Search for the Digital Library

Elaine G. Toms  
Faculty of Information Studies  
University of Toronto  
416-978-7802  
toms@fis.utoronto.ca

Joan C. Bartlett  
Faculty of Information Studies  
University of Toronto  
416-978-4715  
bartlett@fis.utoronto.ca

ABSTRACT
The chief form of accessing the content of a digital library (DL) is its search interface. While a DL needs an interface that integrates a range of options from search to browse to serendipity, in this work we focus on analytical search. We propose using Bates' search tactics as a basis for the re-design of search interfaces. We believe this approach will help to identify the types of tools that need to be supported by a DL interface.

Categories and Subject Descriptors
H.3.3 Information Search and Retrieval

General Terms
Design; Human Factors.

Keywords
Searching; Digital Libraries; Search Interface; Search Tactics.

1. INTRODUCTION
The public face of a digital library (DL) is its interface – the part that users see and manipulate, the part that handles the communication process between a user and the DL components and ultimately the part that contributes to the success or failure of a user’s interaction with a DL. The user is both rewarded by the wealth of options and limited to what the interface allows. Although a DL may provide a range of services, from analytical search to browsing and serendipitous processes, we focus specifically on the search service in this aspect of our research.

At present searching is the preferred and most popular means of finding specific information on the WWW and sometimes the only means of accessing DL content. But, despite their popularity, search engines seem to be no more proficient in aiding the user to find what they need than online catalogues or text databases. The problem is partially concerned with query formulation, and with users’ understanding of the task and the system’s ‘understanding’ of how the user performs that task. The problem is best expressed in a playful definition: “search engines can be quite tiresome and not very fruitful if you don’t know how to use them correctly.”[8] We prefer however to approach the problem from the other perspective: systems do not know what to provide for users.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.

JCDL ’01, June 24-28, 2001, Roanoke, Virginia, USA.
Copyright 2001 ACM 1-58113-345-6/01/0006...$5.00.

In DLs, searching is served by a WWW page that contains a pane(s) for entering a search expression(s), options for selecting search constraints on that expression, and sometimes user directions. However, the design of these pages follows conventional electronic form design practice, rather than being empirically based on how people search for information. This paper will propose an approach to the design of search interfaces that is based on prior research into search tasks and processes, and grounded in Bates [1]’s information search tactics.

2. SEARCH TASK
In general, a quest in a DL starts with an information problem that is conceptualized as a question. The user transforms that question into a query; the system returns a set of references from which the user assesses relevance and may revise the query and re-submit.

In this research we are concerned with the process from expression of need through to query generation and system response. The difficulty is in expressing in words what is in one’s head. The process requires the user to translate need into words and format that are acceptable to the system. This process requires not only the selection of appropriate words, but also knowledge of system specific attributes, e.g., truncation, phrase specification.

Over the past four decades of information retrieval, users have interacted with systems using a variety of styles, most of which were dictated by the level of systems development of the day. From command line interfaces, to simple menu selection, to the graphical user interface, control has shifted from the system to the user, and the requisite skill level from expert-only to include the novice. The search interface in vogue today relies on an interface design technique – ‘form filling’ and DL search interfaces provide little beyond a single box and a search command button to support this process.

3. MODELS OF THE SEARCH PROCESS
To date multiple approaches to the search processes have been suggested [5, 7, 10, 11]. The search process tends to be discussed at a macro level, e.g., Kulthau [6] who provides only guidance and orientation on how that task can be implemented at the interface level. But, at the ‘keystroke’ level, the search task is procedural; commands are entered and responses received. We need an understanding of component steps used to perform the search task at that micro level [9].

Four procedural models offer insights into the search process from different disciplinary perspectives. Three of these are: 1) how students locate information in textbooks [5]; 2) a task analysis of information-seeking activities [7]; and, 3) an information systems framework [11]. These have comparable stages but at differing levels of granularity and have some elements of both problem-solving and analogical reasoning; that, is they include both the
encoding of a goal and stimulus, to compare features and to infer similarities.

The fourth is Norman [10]’s seven stages of human-computer interaction (HCI), considered one of the most influential models in HCI. It too is grounded in problem-solving theory, but describes a generic process that is applicable to each selection at the interface. Thus, for any single stage in the first three, there will be multiple renditions of the fourth, one for each human input that takes place within that stage. Norman’s model is useful for assessing micro stages in the search models, e.g., what intention is taking place? what action sequence must be specified?

4. SEARCH AT THE MICRO LEVEL

While information search models provide generic views of the search process, none are sufficiently detailed to prescribe the interface components needed to support aspects of the search task. Another approach to the search process is that of Bates [1, 2, 3] who proposed 29 information search tactics which she defined as “move[s] made to further a search” – a heuristic meant to assist the meeting of short term goals and the procedures taken during the search process. She grouped these tactics into four categories:

a) Monitoring: “tactics to keep the search on track and efficient.” These range from evaluating search formulation to spell checking.
b) File structure: “techniques for threading one’s way through the file structure … to the desired file, source, or information within source.” These include tactics such as breaking a complex search down into smaller components, or using an indirect route to access information.
c) Formulation: “tactics to aid in the process of designing and re-designing the search formulation.” These involve broadening and narrowing the search with search terms, and Boolean operators.
d) Term: “tactics to aid in the selection and revision of specific terms within the search formulation.” These tactics relate to the selection of search terms, by exploiting the relationships between terms, such as broad and narrow concepts, or spelling variations.

Bates’ tactics were derived from her analysis of search strategies for bibliographic systems and at that time were meant to aid the search process in both manual and online systems. Each tactic might include one or more moves. A move is more fine-grained and can be equated with a cycle of Norman [9]’s seven stage model. Despite the attention that these tactics have received, they remain a theoretical perspective, having never been tested. However, they can prescribe the sorts of tools that one needs at the interface level to support steps in the search process.

In addition to search tactics, Bates also defined idea tactics [2], which are meant to aid the thinking/creative process. These are most useful at the early stages of the search process during problem understanding. She also devised stratagems [3], which are broader, providing guidance for a specific class of search task, e.g., chasing footnotes. These tend to handle particular contexts. Approaching the search process from the perspective of Bates’ tactics replaces the emphasis on search functionality with one that focuses on the user’s task: tools are needed to do the job? We are not alone in taking this approach. Fuhr [4] too suggests that Bates’ four-level plan could be used to implement “strategic support of the information seeking process.” We are merging Bates’ tactics with a procedural model of the search process that is emerging from HCI and assessing the appropriateness of each tactic for each stage of that process. For each tactic we are identifying the type of interface tool that is needed to support that tactic. At the same time we are updating the tactics. For example, no tactic suggests the need for an overview of the digital library content, which has long been recognized as essential information for the user in resource selection [7].

5. CONCLUSIONS

We believe that Bates’ tactics provide a useful approach with which to conceptualize and model the search interface. The tactics need updating and amplification to conform to systems development of today and they need testing. In addition, the tools that are derived from those tactics must be integrated into the DL interface so that they support not only the analytical search function but also all information exploration processes.

6. ACKNOWLEDGMENTS

This research is supported by a Connaught Fund Grant and NSERC grant to the first author.

7. REFERENCES

[8] Netlingo (URL: http://www.netlingo.com)