

User Centred Interactive Search: a Study of Humanities Researchers in a Digital Library Environment

Claire Warwick (*c.warwick@ucl.ac.uk*)

*School of Library, Archive and Information
Studies, UCL*

Ann Blandford (*a.blandford@ucl.ac.uk*)

UCL Interaction Centre

George Buchanan (*g.buchanan@ucl.ac.uk*)

UCL Interaction Centre

This poster proposal describes research on humanities users of a digital library (DL). It seeks to understand their needs and behaviour both in digital and more traditional information environments, in order to develop and refine a digital library system, the better to support use in the humanities. This study of humanities users forms part of the larger *User Centred Interactive Search (UCIS)* project.

Background

Large, structured information repositories such as digital libraries (DLs) are becoming commonplace. To realise their potential, they need to be usable and useful - by a range of users, in different situations, supporting a variety of information tasks. The current generation of DLs still poses substantial user difficulties: searches are often time-consuming and frequently unsuccessful (Blandford et al.), and the reasons for success or failure remain mysterious to most users. Within the broader information task, the information requirements are often poorly defined, as users are often trying to refine the information problem by using available information to understand what is possible, so that information acquisition is an evolving, highly interactive activity.

It is widely recognised that creating effective search criteria to achieve a particular information goal is a demanding and difficult task, particularly for less experienced users, and particularly when the goal is as yet under-defined. Shneiderman et al. observe the challenges of selecting a variety of search attributes, such as the words to be used in a query and the syntactic peculiarities of the system at hand. In addition, the

mapping of an information need to the use of metadata fields or full text search can prove difficult (Blandford et al.). Unlike the web, where the document text is the only possible target for a search, DLs provide a rich environment for information seeking: the user has a much wider potential range of selections (classification, author, publication date, etc.) to make. Effective searching relies on the careful selection and use not only of words or syntactic commands, but also of fields and information sources.

Use in Context

Surprisingly little work on information seeking has set it within the context of the broader information work of which it is a component (Attfield et al.). While this divorce from the context may be valid when considering work in physical libraries, where the information seeking task is often a bounded activity delineated by arrival at and departure from the library building, it is less so for DLs that can be accessed from the user's normal place of work, removing the marked transitions between information seeking and other activities. One hypothesis this study will test is that users expect information seeking to flow more naturally into their broader information task when searching from their normal place of work.

Humanities Users

Humanities researchers are the focus for studying use in context for several reasons: they typically have little technical or mathematical knowledge (e.g. for immediately understanding the designs of complex interactive systems or intuitively being able to construct the Boolean queries that are often key to successful query formulation); they often do not have a clear idea of what they are looking for, but will usually recognise it when they find it; and they have not been extensively studied, although they have substantial and sophisticated information requirements. In summary, humanities researchers are a particularly challenging population to design for, and many solutions that work for this user population are likely to also suit less demanding users. Studies of humanities researchers have tended to concentrate on needs or the types of resources used (*Library Trends; Open University*). Many of these are now relatively dated, and although their conclusions were important at the time, both the types of resources available and the technology used to find them have changed. Studies by Stone and Watson-Boone established that humanities users need a much wider range of resources than those in other disciplines; for example, they may need to refer to material which is much older than that used by researchers in the sciences and social sciences. They may still need to use historical material in the form of manuscripts or early printed books even if digital

surrogates are available (Warwick and Carty, Duff and Johnson). Relatively low levels of use of digital resources have been blamed on these particular needs, and on a lack of knowledge about their capabilities (Corns), but this has yet to be verified by an empirical survey.

Warwick's previous work suggests that humanities users may find it difficult to adapt search behaviour from a traditional to a digital library setting, and thus become discouraged by failed attempts to locate appropriate resources (Warwick, 1999a; 1999b) She has found that the patterns of use of digital resources in English literature have changed little since Corns' work in 1991, and argued that this may be because of lack of fit between the searching tasks users wish to carry out and the present capabilities of DLs. This was based on only a small sample of users, and on theoretical data. It is therefore important to test these hypotheses by studying a meaningful sample of humanities users in both a traditional library setting and a digital environment. This is one of the tasks that the present research is engaged in.

Aims of the project

Overall, there are four strands of work in the UCIS project:

1. studying use of information in context, focusing on humanities researchers;
2. studying the development of expertise in searching (focusing on information management students);
3. identifying requirements on the design of digital libraries; and
4. developing and testing system modules for a digital library.

The proposed poster will describe the first strand of work, briefly outlining how it fits within the rest of the project. We believe that this work is important since very little work has studied use in context - particularly in the humanities - and translated findings into testable design requirements.

Methods

Qualitative data (from interviews, observations, diary studies, transaction logs, etc.) will be gathered from academics and other researchers in the humanities regarding their activities with DLs and similar information resources. Two sub-issues will direct this work: how humanities researchers work with digital resources and how they integrate use of electronic and paper resources - both within the broader task context.

The first approach to data collection will be by user diaries, in which humanities users record their use of information resources (both traditional and digital) to support their research. This will provide base-line data to inform the use of techniques for subsequent study (depending on the patterns of resource use). The main approach to data collection will be contextual inquiry interviews, observing users as they work with digital libraries of their own choice and interviewing them on their perceptions of the usability of such electronic information sources. The focus of this data collection will cover what users currently do, their perceptions of the strengths and limitations of current technologies (including traditional resources), and requirements for future systems.

Data will be analysed in two different but complementary ways: first, using a Grounded Theory-style approach (Strauss and Corbin) to develop theory on the use of digital resources in context by humanities users; second, using design-oriented techniques to draw out requirements for design.

To enable us to focus on new technical challenges rather than needing to replicate work already done by others, technical developments will be based on the *NZDL Greenstone* software, for which a test collection material specific to the humanities has been developed. By grounding the work in empirical studies, we will be able to identify and present further requirements on the design of such systems. By basing system development on an established DL platform, we will be able to test candidate design solutions, deliver working components as part of the *Greenstone* system and provide examples for developers of other DL systems that illustrate tested approaches to improving user experience.

Findings

The *UCIS* project began in August 2004 and the humanities phase will begin in early 2005. We therefore propose to use this poster to report on early findings of the research. It is for this reason that we have proposed a poster session, since this will be a report or work in progress.

Acknowledgement

This work is funded by *EPSRC Grant GR/S84798*.

Bibliography

Attfield. S.J., A. E. Blandford, and J. Dowell. "Information seeking in the context of writing: a design psychology

interpretation of the 'problematic situation'." *Journal of Documentation* 59.4 (2003): 430 - 453.

Blandford, A.E., H. Stelmaszewska, and N. Bryan-Kinns. "Use of multiple digital libraries: a case study." *Proc. JCDL 2001*. Roanoke, VA, 2001. 179-188.

Corns, T.N. "Computers in the humanities: methods and applications in the study of English Literature." *Literary and Linguistic Computing* 6.2 (1991): 127-131.

Duff, W.M., and C.A. Johnson. "Accidentally found on purpose: Information-seeking behavior of historians in archives." *Library Quarterly* 72 (2002): 472-496.

1992.

Open University Library. *Safari: Skills in Accessing, Finding and Reviewing Information*. 2001. Accessed 2005-04-15. <<http://www.open.ac.uk/safari/>>

Shneiderman, B., D. Byrd, and B. Croft. "Sorting out searching." *Communications of the ACM* 41.4 (1998): 95-98.

Stone, S. "Humanities Scholars-Information needs and uses." *Journal of Documentation* 38.4 (1982): 292-313.

Strauss, A., and J. Corbin. *Basics of qualitative research: grounded theory procedures and techniques*. Newbury Park, CA: Sage Publications, 1990.

Warwick C., and C. Carty. "Only Connect, a Study of the Problems caused by platform specificity and researcher isolation in humanities computing." *Electronic Publishing 01, 2001 in the digital Publishing Odyssey. Proceedings of the 5th International ICC/IFIP Conference on Electronic Publishing*. Ed. Arved Hubler, Peter Linde and John W.T. Smith. 2001. 36-47.

Warwick, C. "English Literature, electronic text and computer analysis: an unlikely combination?" *Proceedings of the Association for Computers and the Humanities- Association for Literary and Linguistic Computing, Conference, June 9-13*. University of Virginia, 1999a. 71-74.

Warwick, C. "The lowest canonical denominator: Electronic literary texts, and their publication, collection and preservation." *New Fields for Research in the 21st Century, Proceedings of the Anglo Nordic Conference 1999*. Ed. M. Klasson, B. Loughridge and S. Loof. Boras, Sweden: Swedish School of Library and Information Studies, 1999b. 133-141.

Watson-Boone, R. "The Information Needs and Habits of Humanities Scholars." *Reference Quarterly* 34 (1994): 203-216.