

Creators, Composers and Consumers: experiences of designing a digital library

Ann Blandford¹, Jeremy Gow¹, George Buchanan²,
Clare Warwick³ and Jon Rimmer³

¹ UCL Interaction Centre, UCL, 31-32 Alfred Place, London WC1E 7DP, U.K.

A.Blandford@ucl.ac.uk
<http://www.ucl.ac.uk/ucl-icc/annb/>

² Department of Computer Science, Swansea University

³ School of Library, Archive & Information Studies, UCL, London

Abstract. Many systems form ‘chains’ whereby developers use one system (or ‘tool’) to create another system, for use by other people. Little work within Human–Computer Interaction (HCI) has considered how usability considerations propagate through such chains and yet in many situations the usability of systems is determined by design decisions made at one or more steps removed from the immediate developers of the system in question. In this paper, we relate our experiences of developing digital library components and collections to this notion of “design chains”. This case study illustrates the necessity of looking beyond the immediate users to try to anticipate the needs of stakeholders elsewhere in the design chain.

Keywords: digital libraries; system development; design chains

1 Introduction

It is widely recognised that there are many stakeholder groups in any design project. Approaches such as Soft Systems Methodology [5] encourage explicit consideration of these different stakeholder groups in design. However, when it comes to considering usability, the focus narrows immediately to the end users of the system under consideration. Blandford et al. [2] present a framework for considering ‘chains’ of developers and users of products, which they term ‘creators’ (commonly referred to as designers or toolsmiths), ‘composers’ (users of the tool who compose artefacts for other users) and ‘consumers’ (end users of artefacts). In this paper, we use those concepts to frame our experience of developing and testing digital library components for creating collections to be used by Humanities scholars.

The idea of ‘design chains’ is widely used in supply chain management (e.g. [7]), but the core concern there is typically with ensuring that components from different suppliers can be integrated seamlessly into a product. In contrast, Blandford et al. consider how concepts are propagated through chains of system development, as decisions made by the creator constrain or influence the behaviour of the composer.

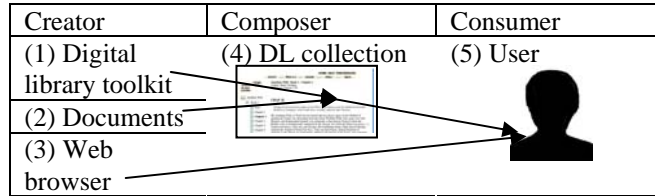


Figure 1. Creators, composers and consumers for a Digital Library interaction.

Chains are not always linear: the end user’s experience of web browsing will depend on both the design of a site and the environment in which pages are viewed. As well as joining, chains may also stretch further in both directions.

A digital library system is an example of a tool within a development chain: developers work with software development environments to create a layer of tools, such as Greenstone [8], with which librarians can create collections of documents to be made available to end users. Reflecting on their work as the creators of Greenstone, Witten et al. [9] note that *“The ease of acquiring and installing a software project has a direct impact on the users it attracts, and consequently—in the open source world—on the extent and nature of contributions that users make to the project”*. Here, it is clear that the users of concern are the users of the toolkit.

However, in a study [3] evaluating Greenstone, another developer suggested that the creators of the toolkit were not recognising their potential role in making it easy for composers (who typically have little HCI expertise) to construct usable digital libraries for consumers: *“[There is a] difficulty with the way Greenstone is perceived by different parties. [The developers] see Greenstone very much as a toolset which other folks should ‘finish off’ to make a good collection. Their conception is that it would be very hard to take Greenstone to a level where a librarian could make a few choices [with a] GUI and have a reasonable (not to say actively excellent) interface for the library.”*

The user’s experience depends on the design of both the web browser (which we do not consider further) and the DL collection (items (3) and (4) in Figure 1); the design of the collection, in turn, depends on the design of the DL toolkit (item 1) and that of the data (item 2), i.e. the set of documents. The creators’ design decisions influence the work of the composer, and hence the experience of the consumer. This can happen directly or because of interference between the decisions of multiple creators.

2. Case Study: Developing a Humanities Digital Library

Many Humanities documents are being digitised in order to make them more accessible to scholars and enthusiasts (e.g. [6]). To investigate alternative interfaces and interaction styles for a Humanities digital library, we created collections that were accessible via the Greenstone DL interface. The designs were informed by empirical studies of the use of existing digital resources (e.g. [1]); they are now being evaluated to develop a richer understanding of Humanities users’ requirements and behaviours.

To be suitable for studies with Humanities scholars, collections had to resemble real-world resources and not ‘toy’ ones. Hence we used large document collections containing both large and complex documents, such as the text of Early Modern bibles. The document creators had encoded these in a variety of bespoke SGML formats. Meanwhile, the toolkit creators had designed a modular plug-in architecture specifically to handle such non-standard formats, and so we wrote our own import code to map each SGML document to a corresponding Greenstone document. This preserved information about sections, which was needed for navigation given the size and complexity of some of the texts.

While implementing a basic collection was relatively straightforward, our work as composers was dictated by the decisions of the document and toolkit creators, as we attempted to shape their products into a form we judged acceptable for the consumers. The document creators had clearly aimed to faithfully encode the source documents in SGML, but some of the markup was not suitable for presentation to the consumer. For example, many section titles were excessively long or uninformative. As composers, we had to extend our import code to cope with unsatisfactory cases – in effect taking on the role of document creators.

Some of the most problematic aspects of composing the collections were due to interference between the two creators’ decisions about what constituted a document section. Firstly, the SGML documents were divided into ‘sections’ ranging from the very large (e.g. ‘book’) to the very small (‘verse’, ‘sentence’), with a great deal in between. In contrast, Greenstone sections are units of text which the reader can choose to browse individually. As composers, we had to decide which SGML tags would make sensible units of browsing – an easy task for extreme cases, but more difficult for others, particularly if poorly documented.

A second interference issue was that the SGML ‘sections’ could contain *mixed content*, i.e. plain ‘unsectioned’ text interleaved with subsections. Greenstone’s presentation of documents assumes that a section may contain only *introductory text* and subsections, so all higher-level sections are placed before any subsections, putting material out of order relative to the narrative flow of the document. We had to force the documents to conform to Greenstone’s view of sections by introducing special ‘wrapper’ sections around the problematic text.

The consumers can also negatively experience the creators’ decisions in ways that the composer can do little about. For example, even between documents in the same collection, the date metadata sometimes reflected the creation date of the work, sometimes the date of the particular edition: so medieval poetry could be dated anywhere between the Middle Ages and the 20th Century.

The decisions of the toolkit creators had most impact when it came to customizing the interface. Their HTML macro language made changes to individual page components very easy, but the large-scale page redesign we required was more difficult. For instance, the ‘table of contents’ navigation menu was displayed above the document text, which for some large documents meant the consumer would repeatedly have to scroll down past a huge menu before viewing the section text – a particular problem for users in an earlier study [4]. To overcome this we were forced to take on the role of toolkit creator and edit the source code.

Despite our efforts as composers, consumer issues still arise where we have

attempted to overcome interference between the creators. For example, some sections are inappropriate for browsing alone, being either too small or large, or not comprehensible in isolation, while other sections have inappropriate or missing titles; this impacts on the consumer's navigation of the document. Also, the 'wrapper' sections we introduced to overcome Greenstone's problems with mixed content can sometimes result in oversized and incomprehensible navigation menus.

3. Discussion

As composers, our aim was to produce a set of collections for Humanities users. The collections had to look and behave as 'realistic' resources and meet a basic standard of usability. With Greenstone the role of the composer is often played by librarians [9], with a wide variety of technical skills, so the toolkit creators have ensured that very little configuration is required to build a working collection from a set of digitized text documents. However, there are standardization issues that emerge from different creators having different tacit assumptions, e.g. over the role of a 'section' in navigating and analyzing text; this issue mirrors the compatibility concern in traditional 'design chains' from supply chain management. In addition, toolkit creators' decisions have a direct impact on the experience of consumers as they navigate the interface. As composers, we frequently had to switch into the creator's role to facilitate the development of a reliable and usable Humanities digital library for the end users. Thinking in terms of design chains has helped to clarify the responsibilities of some important stakeholder groups in the development of systems that involve the use of tools over datasets to create interactive experiences.

References

1. Bates, M. J. (1995) The Getty End-User Online Searching Project in the Humanities: Report No. 6: Overview and Conclusions. *College & Research Libraries* 57. 514-523.
2. Blandford, A., Green, T. R. G., Connell, I. & Rose, T. (2007) Knowledge Representation Environments: an Investigation of the CASSMs between Creators, Composers and Consumers. To appear *Proc. Engineering Interactive Systems 2007*.
3. Blandford, A., Keith, S., Butterworth, R, Fields, B. & Furniss, D. (2007) Disrupting Digital Library Development with Scenario Informed Design. *Interacting with Computers* 19, 70-82.
4. Blandford, A., Stelmaszewska, H. & Bryan-Kinns, N. (2001) Use of multiple digital libraries: a case study. In *Proc. JCDL 2001*, 179-188. ACM Press.
5. Checkland, P. B. (1981) *Systems Theory, Systems Practice*. Chichester: John Wiley.
6. Jefcoate, G. (2000) Digitization for access: the digitization of rare books at the British Library. *Proc. Kyoto Int. Conf. on Digital Libraries*, 103-109.
7. Twigg, D. (1998) Managing product development within a design chain, *International Journal of Operations and Production Management*. 18 (5) 508-524.
8. Witten, I.H., Bainbridge, D. & Boddie, S. J (2001) Greenstone: open-source digital library software with end-user collection building. *Online Information Review* 25(5), 288-298.
9. Witten, I. H. & Bainbridge, D. (in press). A retrospective look at Greenstone: Lessons from the first decade. To appear in *Proc. JCDL 2007*.