

Learning by Building Digital Libraries

David M. Nichols and David Bainbridge
University of Waikato
Hillcrest Road, Hamilton,
New Zealand
+64 7 838 4021

{dmn, davidb}@cs.waikato.ac.nz

J. Stephen Downie and Michael B. Twidale
University of Illinois at Urbana-Champaign
Champaign, IL 61820
USA
+1 217 333-3280

{jdownie, twidale}@uiuc.edu

ABSTRACT

The implications of using digital library software in educational contexts, for both students and software developers, are discussed using two case studies of students building digital libraries.

Categories and Subject Descriptors

H.3.7 [Information storage and retrieval]: Digital Libraries – user issues. K.3.2 [Computers and Education]: Computer and Information Science Education – information systems education.

General Terms

Human Factors

Keywords

Digital librarian, education, software, digital library.

1. INTRODUCTION

Typically we think of digital libraries in education as resources that students use for searching, browsing and accessing their contents. However in certain circumstances, it can be valuable for students to actually create a digital library (DL) rather than to simply 'consume' one. This naturally applies to students wishing to learn how to build DLs or those who will need familiarity with their underlying nature. Some of these students may be described as future 'digital librarians' (in Library and Information Science (LIS) Schools or Information Schools) whereas others may simply be acquiring competencies for life in the information society [1].

In this paper we examine two aspects of students *creating* DLs:

1. Effects of the DL software on the students' experiences
2. Feedback from student use to DL software development

We discuss two case studies, in a computer science department and an LIS school, that use Greenstone digital library software.

2. CASE STUDY A: SERVICE TEACHING

At Waikato University—home of Greenstone [5]—DL creation is taught as one of six lab-based modules in a second year undergraduate course. In this service course, open to all students, a range of applications are explored: from video editing to

desktop publishing, animated 3D-modeling to web design. The pedagogic value of teaching DL construction is to expose students to principles of practical information organization; hopefully as a transferable skill across many domains.

The Digital Libraries module has been included seven times since 2004. In its first inception, installation of the software was an aspect covered, necessitating the use of a non-networked lab of PCs so students could be granted the necessary administration privileges to install it. This requirement was at odds with latter parts of the lab, which expected an Internet connection so the students could go out and locate sample documents of their own. The solution used was for students to go into an adjacent (networked) lab, where they could access Web resources burning them to CD-ROM before returning to the original lab.

In response to this less than ideal workaround (and the resource implications of using two labs), the Greenstone team developed an applet version of the main Graphical Interface. This means students no longer need to install the software, and so can use a standard lab (with a Java-enabled web browser) for all work. This client-server arrangement generally performed satisfactorily; but as students tend to leave their work to just before deadlines the pattern of usage was unusual. When a significant number of students tried to simultaneously upload data the server software was stressed in a previously unrecognized manner. The local presence of Greenstone programmers is useful in dealing with these situations and, as with usability testing, there is no substitute for seeing your software fail with your own eyes. The original simplistic data transfer process has since been re-engineered but it is noteworthy that typical server or single-user use of Greenstone is unlikely to have revealed these issues.

3. CASE STUDY B: TEACHING DIGITAL LIBRARIANS

In 2005, the Graduate School of Library and Information Science at UIUC launched its post-Masters Certificate of Advanced Study in Digital Libraries (CAS-DL). The Digital Libraries: Research and Practice (DIL) course prepares students for a major implementation project. Small groups are required to construct fully functional DL prototypes consisting of text, graphic and audio documents. The DLs created were relatively small in size (limited to 100-150 documents), but were broad in that students were responsible for creating complete metadata records for each document, and generating appropriate index sets for searching, browsing mechanisms and interfaces for their potential users.

A reading of the 2945 postings made on the course's web discussion board reveals a set of seven positive and three negative themes in students' opinions about the use of the Greenstone DL (GSDL) software for the course. Positive Themes:

1. Near instant interface redesigns to quickly test out ideas
2. Useful and flexible metadata creation tools (especially in creating and augmenting standards such as Dublin Core)
3. Greenstone's Librarian Interface (GLI) Tool providing a step-by-step 'image' of the DL creation workflow
4. Large, varied install base for examples
5. Proven flexibility on a variety of media types
6. Platform independence: Linux/PC/Mac versions that students could install on home machines for 'fooling around', particularly important to the distance students
7. Tie-in with *How to build a digital library* textbook [5].

Negative themes:

1. Problems understanding the difference between real-time tasks (e.g., 'tweaking the interface') and rebuilding tasks (e.g., re-indexing, hierarchy changes)
2. Problems comprehending the interplay between GSDL's sub-components (e.g., config scripts, Perl macros, etc.).
3. Problems and fears (and bad experiences) concerning the lack of a robust 'rollback' feature in GSDL to accommodate 'failed' design experiments (i.e., losing previous work by incorrectly implementing a new or unknown feature).

The instructor's opinions about these themes also happen to align with the students'. From these 10 themes, two larger themes emerge. First, the students valued highly the ability to 'experiment' and 'tweak'. They especially desired and appreciated the positive reinforcement of seeing the results of their experimentations in real-time to clearly establish 'cause and effect'. Catastrophic failures (from which they could not recover) likewise became an inhibiting, thus negative, factor. This is a classic case of the power and desirability of creating a supportive microworld in an educational setting [3] (and negative theme #3 notes the problems when this does not occur). Second, and related to the establishment of clear cause and effect understandings, the unrecoverable catastrophic failures demonstrated that students did not fully comprehend the complicated interplay between GSDL's computational sub-components (e.g., not truly understanding how a particular Perl module passes metadata to another Perl script that controls the indexing process, etc.). This analysis leads us to recommend further Greenstone design effort to enhance the establishment of 'cause and effect' and increase the transparency of the underlying operations of the software.

4. THE PEDAGOGICAL IMPLICATIONS OF DIGITAL LIBRARY TOOLKITS

The history of professional training in librarianship shows recurrent themes of core values of access and the role of intermediary services, manifested in a variety of technologies and media from card index files and books to OPACS, DVDs and MP3 files [4]. Librarians have always been open to the challenges of incorporating new technologies into their traditional mission.

There is an ongoing discussion on what it means to be a digital librarian [1,2]. It is similar to being a traditional librarian in terms of ethos and applicability of core guiding theories, including access, cataloguing, collection development and teaching people search skills. However it also requires additional technical skills, extending the librarian's role into new areas. For example, DL creation can be more akin to publishing than collection development, involving aspects of editing, revision and aligning to data and metadata standards. Previous work on education for

digital librarians has focused on curriculum design issues (e.g. [1,2]) and has not considered the influences educational use can or should have on the development of the DL *software* itself.

The topic of digital librarianship [2] is subject to rapid change as it is partially *defined* by the availability and functionality of appropriate software. The ease of building a digital library or an institutional repository has been radically changed over the past decade by the emergence of software such as Greenstone and DSpace. For example, the Waikato lab exercises are derived from material presented in tutorials at digital library conferences.

Using digital library software in an educational context reveals different elements of its usability and functionality to its use in server environments or personal collections. From an inspection of the interface alone some things, such as the editing of format statements, are easily identifiable as difficult to use. This needs a significant commitment of programming time to improve. For others, answering questions in the lab is an opportunity for software developers to see what causes difficulty for students. The design goal for GLI is that 'easy things should be easy, difficult things should be possible'. After one training course it was clear that having the interface 'littered' with the features necessary to support 'difficult things should be possible' led some students to be overwhelmed. The result was the introduction of different interface modes. GLI starts in 'Librarian' mode which provides access to commonly used features. Progressing through different modes, up to 'Expert', reveals more powerful features that require greater user knowledge (e.g. regular expressions). In this way the software can be used at different levels (e.g. undergraduate and masters) and can also support progress for individual students as they gain familiarity with the concepts and practice of DL creation. As with the applet extensions (section 2), the modes are now part of the software distribution available to all.

5. CONCLUSION

Although the majority of educational uses of Digital Libraries will involve using resources constructed by others, the act of constructing a digital library can also be an important learning experience. The progress of the field can be seen in activities that were research a decade ago but now occur in self-paced undergraduate courses. In this paper we have started to outline the issue surrounding the educational potential of DL construction and also its implications for the design of DL software.

6. REFERENCES

- [1] Coleman, A. Interdisciplinarity, interactivity, and interoperability for educating the digerati. *Education for Information* 23, 4 (2005) 233-243.
- [2] Mostafa, J., Brancolini, K., Smith, L.C. & Mischo, W. Developing a digital library education program. In *Proceedings of JCDL2005*, 427. ACM Press, 2005.
- [3] Papert, S. *Mindstorms: Children, Computers, and Powerful Ideas*, Harvester, Brighton 1980
- [4] Wilson, A. M., & Hermanson, R. Educating and training library practitioners: A comparative history with trends and recommendations. *Library Trends* 46,3 (1998) 467-499.
- [5] Witten, I.H. & Bainbridge, D. *How to build a digital library*. Morgan Kaufmann, San Francisco, CA, 2003.