LUCE LIBRARY BULLETIN

New Times, New Tools

Clifford B. Anderson

*T*he practice of librarianship has been changing rapidly during the past decade. Looking back, it is hard to believe that Wikipedia—the most widely used reference source in the world—is less than ten years old. Google is a mere twelve years old, but the company's search engine has transformed the information landscape. The settlement of the Google Books lawsuit promises another leap forward. Finally, the Internet Archive, a non-profit library that has been archiving web sites since 1996 and building a free alternative to Google Books, is taking on the challenges of preserving our digital heritage for the future.

How can librarians working in academic libraries keep up? The pace of change seems so rapid that it threatens to become overwhelming. In this issue, we talk about some of the tools we are using to keep up with the rapid pace of development. Many of these tools were originally developed by software engineers—not surprising, given that programmers work in a far more unstable environment than librarians.

The basic attitude that we've learned to adopt is to become continuous learners. In that spirit, we invite you to write to us about practices that you've developed in your archives, or your professional life, to keep up with the times.

On a different note, we are sad to announce that Sarah A. Seraphin, manuscript librarian, will be leaving our employ this June to become the special collections librarian (part time) for the Department of Special Collections at La Salle University's Connelly Library, in Philadelphia, Pennsylvania. Sarah has done a tremendous job on multiple projects in Special Collections and we wish her the very best in her new position.

Collaboration Tools for a Digital Archive

Clifford B. Anderson

*A*t a library conference a few years ago, Christine Schwartz (our metadata librarian) and I gave a presentation on our digital library program. Among the comments we received was a message that said something like, "What you're talking about may work for Princeton, but it's too complicated for the average archivist."

Well, that's true to some degree. Developing a digital program requires archivists to learn new skills. Up to recently, the so-called "average archivist" has not needed to think much about data formats and programming languages. If archives used software at all (beyond an integrated library system), it came as a "black box" supplied by a vendor.

The demand to digitize archival collections affects every archivist. These days, patrons expect to find information about our holdings online and, increasingly, are questioning why our collections themselves are not available electronically. Archivists have been learning new techniques and tricks to meet this demand. The "average archivist" has become more tech-savvy in the years since Christine and I made that presentationin 2007.

Special Collections serves as the official repository for the archives of Princeton Theological Seminary and houses rare books, personal papers, manuscripts, and special research collections that document the Presbyterian and Reformed tradition and ecumenical Christianity in the United States and abroad.

Supported by the vision and resources of Princeton Theological Seminary, the collections are located in the Henry Luce III Library, joined to Speer Library on the Seminary campus.

For more information about the facilities and collections of Luce Library, see the Special Collections web page on the Seminary Library home page: http://library.ptsem.edu/collections.

PRINCETON
THEOLOGICAL SEMINARY

Collaboration Tools for a Digital Archive

In this newsletter, I share some thoughts about five things we've found helpful when developing a digital program for our archives. Actually, most of these are not really technical. If anything, the ideas described below might be classified as means for collaboration rather than as technological tools.

Continuous Training

Learning a programming language is like learning a spoken language. While some companies promise to teach you French in twenty-four hours, we all know that it takes longer to become fluent. Mastering a programming language also requires sustained commitment. While our staff did spend a few days in intensive training with a professional consultant, that was only the beginning of our journey. For more than a year now, we've run a programming seminar. We meet every Friday for roughly an hour to study and learn together. Leadership for the seminar rotates. Typically, someone will show off a new feature in development. Or I will introduce a new language construct. We try to keep the tone light and to make it enjoyable. The collaborative setting is crucial, since programming can be frustrating, especially when you just cannot find that "bug" lurking in your code. The opportunity to talk through our work and to get feedback from our peers has really helped to move us toward programming "fluency."

Source Control

Programming is really just another form of writing. And, like any writing project, you produce drafts before you turn in the final version. We all have informal systems for identifying our drafts. For example, we save the filename with the last revision date. But we've all mixed up files from time to time and either overwritten the last draft or mixed it up with previous drafts. Now think about this problem when writing a document collectively. The chances of overwriting good copy with bad greatly increase with multiple authors. And what happens when you or a colleague has written something that just isn't very good? You'd like to be able to roll back to the last good version, right? In a nutshell, this is the problem solved by source control. Source control software maintains every draft and keeps record of who made any change. With source control software, you can always determine the latest version of any document, and you can always undo any change that you come to dislike. Using source control is essential for maintaining harmony in a multi-staff environment, but it's also pretty useful even for maintaining order with individual projects.

We use a commercial system called SourceGear Vault. However, there are excellent open sources alternatives such as Apache Subversion and Git.

Wikis

Documentation is a funny thing: we know that we need it, but never seem to have time to write it. A problem with developing a traditional "office manual" for digital collections is that with information changing so quickly, it's easy for a printed manual to go out of date. To solve this problem, we maintain an internal wiki of protocols, procedures, and helpful tips. A wiki is basically a content management system with source control built in. We're all familiar with Wikipedia. Wikipedia democratizes the production of encyclopedic knowledge by allowing anyone to edit its pages. Similarly, our wiki democratizes our internal procedures. I almost never write articles on our wiki unless I am directly involved in a project. Procedures and protocols emerge from the experience of staff members who are directly working on projects. For this reason, our staff wiki tends to have more "buy in" than traditional office manuals. Moreover, it tends to stay up to date. That's important when working on digital projects, since best practices may shift frequently, especially as staff members discover better and more efficient ways to do their work.

We use an open source system called Screwturn Wiki. I'm proud to say that my colleague, Bill French, director of telecommunications, network, and support services, is among the elite contributors to this software.

Outsourcing

Controversial, yes. But outsourcing is critical for the success of major digital initiatives. An outsourcing partner can not only help you to work more cheaply, but also work more efficiently and smartly. That is why I classify outsourcing as a collaboration tool. A good outsourcing company becomes a trusted partner committed to furthering the success of your projects.

In our case, the outsourcing calculus was simple. We just compared the staff time (and hence cost) of asking staff members to markup finding aids in XML versus the cost of outsourcing that project. The cost figures were not even close. By freeing up our staff members to work on more complex tasks, we were able to accomplish much more than we would have if we had asked them to type in XML tags by rote.

Of course, outsourcing is not like shifting an office onto autopilot. We've learned to set up quality assurance procedures and to provide feedback whenever a course correction is needed. Of all the tools on this list, though, outsourcing has by far provided the greatest

benefit for our digital program.

Our outsourcing partner is codeMantra, a company with front offices in Plymouth Meeting, Pennsylvania, and back offices in Chennai, India.

Kanban

It is a truism that digital projects require different methods of organization than traditional library workflows. Digital projects are inherently interdisciplinary, involving numerous library departments as well as the IT department and external partners. Top-down administration does not work very well in such situations. We prefer to work as a team, with team members reporting to one another on their progress, sticking points, and plans.

We keep track of our projects using a kanban board. The idea of a kanban board originated in Japan, where factory workers maintained their workflows by using simple bulletin boards with columns for marking tasks as "ready," "working," and "review." The idea is not only to capture the status of the project at a glance, but also to coordinate the activities of many workers with responsibilities for different aspects of a task. We're still feeling our way into this project management philosophy, but it's already helped us to gain broader perspective on all the "moving parts" required to digitize archival documents, from selection to quality review.

We use a digital kanban system called AgileZen for our projects. However, a simple whiteboard or corkboard can readily serve as a substitute for a software program.

I am just scratching the surface here. A tremendous amount has been written about all of these tools. And there are many others. I hope to have made the point that developing digital library programs is not just a matter of purchasing new software. To develop an innovative digital program, we need to learn and adopt some of the "best practices" from software engineering. Strangely, many of these best practices are not really about software development at all, but about working as a team to accomplish innovative projects in fast-paced and quickly changing environments. ∎

Presenting the EAD Search Engine

Sarah Seraphin

In the spring 2009 *Luce Library Bulletin* we announced that more than 200 finding aids to our archival collections were available online. That project converted nearly 2,000 pages of information previously only available to our researchers in print into machine-readable XML format that could be searched and displayed online. The method of delivery that we employed to deliver this information was to generate HTML links to the documents and add those links to the online cataloging record. The catalog record itself lists brief bibliographic information about the collection, including creator, title, dates, number of boxes, and subject headings. The finding aids list descriptions of contents of the collection, along with other biographical and historical information.

Using the catalog as a platform for the finding aids relied on a couple of presuppositions: first, that researchers seeking information about our archival collections would access the Seminary Libraries' online catalog (as a local or remote user); and second, that in their search they could locate the catalog record for a given collection. Since any given search is restricted to the bibliographic information of the catalog record, a successful search is contingent upon a user's ability to narrow a concept into search terms. Terms coded in the online aids, however, are not indexed in the catalog. In effect, one could search and locate the link to view a finding aid, but could not search the finding aids themselves.

Providing an online link was a step in the right direction for accessing finding aid documents, but we were aware that we were not delivering them with the full-text searching capability of their XML format. In 2007 the library switched to a MarkLogic server, which is an extremely sophisticated XML content server. With this server we can store our digital files, build search engines, and

Recent Acquisitions

Kenneth Woodrow Henke

"Dancing Scotchmen" Build New Seminary Library

Did you know that when the Seminary's original Lenox Library, the first free-standing library in Princeton, was being constructed in 1842, the Scottish workers occasionally indulged in a little recreation to lighten their toil? Among the pastimes we now know about was a Sunday climbing contest, up and down the scaffolding. Another was the day they invited a wandering blind French clarinet player and his German fiddling sidekick over to the work site. While the musicians played, the workers took a little time off from the job to enjoy dancing with each other on the ladder and scaffolding. These and other facts about life at Princeton Seminary in the nineteenth century come from three newly acquired manuscripts in the Seminary archives. One volume is a journal kept by Professor Joseph Addison Alexander for 1838, with selections also dated 1839. The other two volumes are a kind of short diary called "The Twopenny Book" started by two of the children of Archibald Alexander and kept up (though in some periods more fully than others) by a number of hands in the Alexander household between the years 1832 and 1844.

Volume Two of *Paul Tillich, His Life and Thought*

Marion Pauck writes that she is currently preparing the second volume of *Paul Tillich, His Life and Thought* (mentioned in the spring 2008 issue of *Luce Library Bulletin*) for publication. The second volume will be based on the drafts and notes left by her husband, Wilhelm Pauck. She also clarifies that she was the primary author of the first volume. ∎

Presenting the EAD Search Engine

deploy applications online in a relatively simple way. With the goal of increasing the depth of access to our finding aids, we set out to design a search engine that would perform full-text searching across all of the finding aids. Building on the dynamic search features of our server, we were able to design this search engine in house and host it online as a new digital feature.

The new feature can be found online at: http://digital.library.ptsem.edu/ead. The page displays welcome graphics that link to guides to some of our most well known collections; however, the main feature is the search box. Entering a query into this search box will perform a "Google™-like" search over all of the archival finding aids in our database. Searching any keyword or phrase will return a set of results, listing up to ten archival collections that include the search term per page. The results display the title of the collection (as a hyperlink), a brief summary of the collection, the scope and content (or abstract) of the collection, and a list of highlighted results for where the term appears in that collection's list of contents. Clicking the hyperlinked title will direct the user to the full finding aid display, and the search term will appear throughout the document with a yellow highlight.

We hope that this search engine will enhance the information-seeking process in a number of ways. Providing the finding aids their own searchable database in itself narrows the scope of the search, and this full-text searching feature enables the user to browse through results in multiple documents in minutes. As librarians and archivists, we are aware that the archival collections in our library have many complex and interesting relationships that could take years to index in the traditional sense, but now through a simple search, those connections can be discovered by our researchers. Descriptive information, including letters from notable theologians, photographs, scrapbooks, and rare publications, are described in the finding aids. We invite you to explore this new feature and discover these conduits in the history that we preserve. ∎

Book Launch: *John A. Mackay Biography*

Kenneth Woodrow Henke

The Hand and the Road: The Life and Times of John A. Mackay, the first English-language biography of former Princeton Theological Seminary President John A. Mackay, appeared in early 2010. It was researched and written by John Mackay Metzger, grandson of John A. Mackay and son of the late Princeton Seminary George L. Collord Professor of New Testament Language and Literature Emeritus, Bruce Manning Metzger.

To celebrate the occasion of the appearance of this important biography, the Special Collections Department hosted a book launch evening in February. A group of invited guests was treated to a dinner talk by Dean of Academic Affairs and Henry Winters Luce Professor of Missional and Ecumenical Theology Darrell Guder. Guder spoke of Mackay's influence on his own life and thought, beginning with hearing him speak while Guder was still a young student. He went on to point to the significance of John Mackay's life and thought for exploring such important topics as church-parachurch relationships, the theology and practice of evangelism, the role of the seminary and theological education in the life of the church, the public witness of the church (as related, for example, to Mackay's response to the politicized Roman Catholicism he encountered in Latin America or to his challenge to McCarthyism during the early Cold War period), the emergence of ecumenical theology
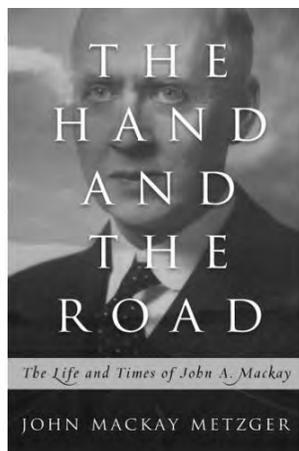
The Hand and the Road, by John Mackay Metzger

in the 20th century, and approaches to contextualization of the Christian faith in a given culture, such as the work of John Mackay as a missionary to Latin America in the first part of his career, and his very important book, *The Other Spanish Christ.*

After dinner John Metzger gave a public talk on John Mackay's life and thought. The illustrated talk touched on topics such as John Mackay's record of his own early religious experiences; his educational background in Scotland, Princeton Seminary, Spain, and Latin America; his approach to Christian faith and theology, particularly his understanding of the church; his service to Latin American and world Christian missions and to the Presbyterian Church in particular; and his understanding of and contributions to Christian ecumenism. Metzger also spoke individually with those who attended the event and signed copies of his book.

The Hand and the Road has been beautifully printed by Westminster/John Knox Press. The narrative flows smoothly, making it easy to read, and the book contains a generous selection of photographs. There is a comprehensive index, 82 pages of carefully detailed notes, and a 22-page selected bibliography of Mackay's published works in both English and other languages, making the book an excellent starting point for those wishing to do further research into John Mackay's thought and legacy. For those wishing to order their own copy from their local bookstore, the ISBN-13 number is 978-0-664-23524-6. ∎