



CALA Occasional Paper Series

No. 4

June, 2009

Message from the Editor-in-Chief

Welcome to the *CALA Occasional Paper Series* no. 4. In this issue, Jin Xiu Guo discusses the building of an institutional repository at a small library.

I would like to take this opportunity to express my sincere gratitude to two copy editors: Manuel Urrizola and David Hickey for their careful and labor-intensive editing on CALA OPS issues. I would also like to say “thank-you” to our layout editor, Xiaoli Fang, for her work on designing layout of the published pieces.

Judy Jeng
Editor-in-Chief, CALA OPS

CALA Occasional Paper Series Editorial Board

Editor-in-Chief



Judy Jeng, Ph. D.
Head of Collection Services
Congressman Frank J. Guarini Library
New Jersey City University
2039 Kennedy Blvd.
Jersey City, NJ 07305-1597
(201) 200-2372
Fax: (201) 200-2330
jjeng@njcu.edu
<http://web.njcu.edu/sites/faculty/jjeng/>

Assistant Editor for Production



Xiaoli Fang
Distance Learning Librarian
Congressman Frank J. Guarini Library
New Jersey City University
2039 Kennedy Blvd.
Jersey City, NJ 07305-1597
(201) 200-3008
Fax: (201) 200-2330
xfang@njcu.edu

Assistant Editors for Manuscripts



Manuel Urrizola
Head of Cataloging & Metadata Services
University Libraries
University of California, Riverside
P. O. Box 5900
Riverside, CA 92517-5900
(951) 827-5051, Fax: (951) 827-3720
manuelu@ucr.edu

Members



Lois Mai Chan, Ph. D. Professor
School of Library and Information Science
University of Kentucky
Lexington, KY 40506-0039
loischan@uky.edu



David Hickey
Asian Studies Bibliographer
Smathers Libraries
University of Florida
P. O. Box 117007
Gainesville, FL 32611-7007
(352) 392-9075 Ext.312,
Fax: (352) 846-2746 dhickey@uflib.ufl.edu



Chihfeng P. Lin, Ph. D.
Associate Professor and Chair
Dept./Graduate Program of Information and
Communications
Shih Hsin University
No. 1, Lane 17, Muzha Road, Section 1
Wen-Shan District
Taipei 11603, Taiwan
886-2-22364906
Fax: 886-2-22361722
chihfeng@cc.shu.edu.tw

Building an Institutional Repository at a Liberal Arts College

Jin Xiu Guo
Cataloging / Metadata Librarian
Washington College
Chestertown, Maryland

Email: jguo2@washcoll.edu
Phone (office): (410) 778-7740
Fax: (410) 778-7288

Jin Xiu Guo, is a Cataloging / Metadata Librarian at Washington College, Chestertown, Maryland. She is a NITLE Technology Fellow, and her fellowship focuses on DSpace and institutional repository development. Her blog Metadata Librarian Experience <http://metadatalibrarian.blogspot.com/> talks about exploring technology, metadata, and taxonomy in a digital library. Jin can be reached at jguo2@washcoll.edu

Building an Institutional Repository at a Liberal Arts College

Abstract:

Institutional repositories (IRs) are being used increasingly as infrastructures for self-created digital collections at large academic and research libraries. This paper demonstrates how a small liberal arts college can build an institutional repository (IR) in collaboration with an appropriate consortium. It also examines IR emerging issues in metadata creation and preservation, as well as IR promotion, and collaboration across departments in the college community through the Senior Capstone Experience (SCE) project. Lessons learned from the project and possible solutions proposed by the author are highlighted.

Introduction

Since we moved into the digital age in the early 1990s, academic libraries have had more opportunities to serve researchers effectively with diversified and easily accessible electronic resources than ever. Libraries have initiated various digital projects and created open-access resources with emerging technologies for users. The cost of building an institutional repository (IR) has been greatly reduced with the development of open-source software, which makes technology-driven institutional repositories (IRs) possible for some libraries. MIT, for instance, is a pioneer of development and the use of open-source software DSpace to provide access to research and learning resources. MIT's CWSpace project was to integrate the course materials found in MIT's OpenCourseWare (OCW) into MIT's institutional repository DSpace. The project goal was to "harvest and digitally archive OCW learning objects and make them available to learning management systems by using Web Services interfaces on top of DSpace" (Reilly, Wolfe & Smith, 2008). Similarly, the University of Maryland Libraries currently uses Fedora, another open-source software, as its IR to provide public access to the collections of films, images, manuscripts and postcards. Image software, the Madison Digital Image Database (MDID) developed by James Madison University, can integrate the digital image library into teaching and learning process. Not only American libraries are involved in the IR software development and application, other libraries around the world have also explored IRs. The New Zealand Digital Library Project at the University of Waikato developed a multilingual software Greenstone, which possibly attracts more institutions to build digital libraries in non-English speaking countries. Brazil used open-source software TEDE as a national digital repository for scholarly information and provided access to Brazilian electronic theses and dissertations (Marcondes & Sayão, 2003). Cambridge University Library in the United Kingdom initiated an 18-month collaboration project SPECTRa (Submission, Preservation, and Exposure of Chemistry Teaching and Research Data), which enabled chemists routinely to deposit experimental data in DSpace (Morgan, 2007).

As a result of the remarkable development and application of open-source software and IRs, academic libraries have been transformed into publishing and archival institutions through "creating a mechanism whereby information can be collected, organized, preserved, and broadly disseminated outside the confines of the traditional publication format" (Jantz & Wilson, 2008). To break through the traditional roles of collecting and preserving information for communities, academic libraries actively seek opportunities to serve users by exploring their roles in information facilitation, dissemination, and creation. Additionally, an institutional repository could also "increase visibility and prestige, and act as a web promotion to funding sources, potential new faculty, and students" (Prosser, 2003).

It is no longer a dream for academic and research libraries to create customized collections and provide open access to the world researchers. In such case, faculty and students can contribute their publications to the collections in IRs; libraries will make the collections easily accessible to the community; faculty and students are able to know what has been done at their universities and colleges by searching the collections, and they have free access to these research publications. When these collections are shared by large communities, such as OAIster and WorldCat, not only faculty and students benefit from their own research work, but also external researchers can share the same information, which will further promote scholarly publications in IRs. Although open-source software is free of charge, libraries still need professional expertises and facilities to implement and maintain

institutional repositories. It is not cost-effective for a small college with limited resources and IT staff to develop and maintain an institutional repository on its own. Today, small colleges are seeking solutions to implementing a less expensive, but a well-functioning institutional repository. This paper describes how a small liberal arts college works with the college community and a consortium to establish and maintain an institutional repository. The purpose is to share some lessons learned from building such a repository with those who plan to initiate an IR in the future.

IR Implementation at Washington College

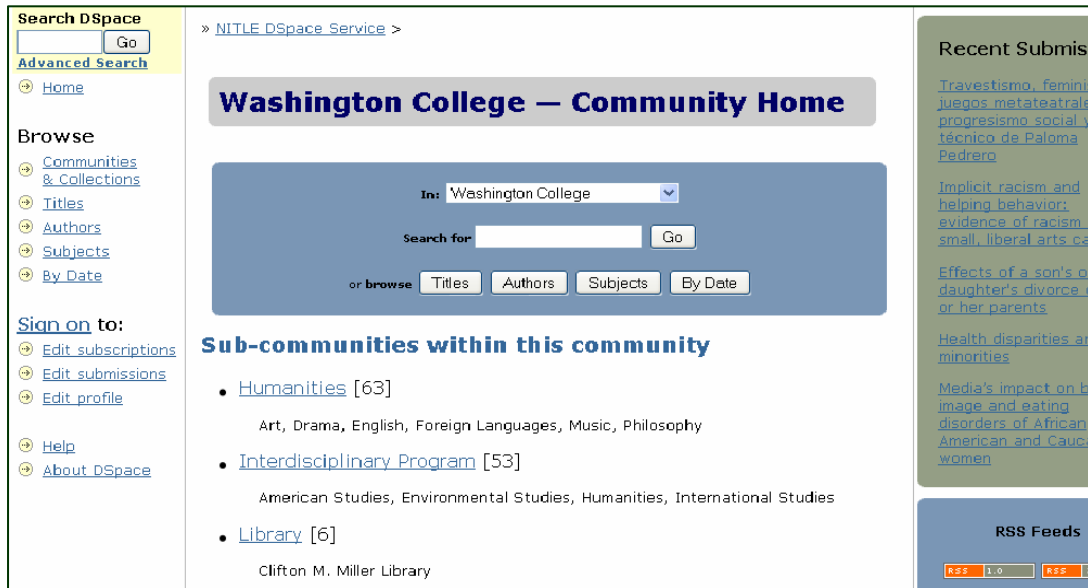
Washington College is a small private liberal arts college on Maryland's Eastern Shore. The college offers around 50 undergraduate and graduate programs in the arts and sciences for 1,400 students. Innovation has been the key to keeping the college dynamic and leading to a memorable learning experience for its students. To adapt to the digital world and serve the twenty-first century students, the library tries to maximize the use of college resources to support teaching, create an impressive students learning environment, and meet the needs of students and faculty.

The Clifton M. Miller Library at the college preserves college archives and the work of faculty and students. These collections are not available online, researchers have to visit the library to access the collections. To promote these resources, the library wants to use an institutional repository as a long-term preservation infrastructure to provide online access for users. It is more practical and cost-effective for the library to join an IR consortium rather than build an IR from the scratch. Therefore, the library has made great efforts to select a consortium. This consortium should provide high-quality IR services with effective digital technologies for liberal arts colleges, and help the library move to the digital age. This consortium should also be a place where Miller Library is able to share ideas and communicate with peer institutions. The digital content management system should adopt a widely used metadata schema to provide sustainable access to digital collections, such as manuscripts, images, archives, audio and video media collections. The system should be able to exchange metadata with other information systems to share digital collections and integrate various digital collections into the Integrated Library System. It is also important for the library to be able to afford the IR.

The National Institute for Technology and Liberal Education (NITLE) seems to be a good fit. With the funding from Andrew W. Mellon Foundation, NITLE was created in 2001 and “became the first virtual, networked collaborative of its kind in higher education with eighty-one affiliated liberal arts colleges” (Barrow, 2004). Its DSpace Pilot began with the Liberal Arts Scholarly Repository (LASR) project. NITLE promotes innovation and interinstitutional collaboration. Currently, there are around 140 liberal arts colleges in America and Europe using NITLE information services and products of DSpace, Moodle, and Sakai. NITLE devotes itself in helping liberal arts colleges manage information in well-planned and cost-effective ways. NITLE DSpace adopts qualified Dublin Core metadata schema and supports the Open Archives Initiative’s Protocol for Metadata Harvesting (OAI-PMH) v2.0 as a data provider. NITLE is also an OAIster (a union catalog of digital resources) data contributor. Therefore, participating in NITLE not only increasingly exposes Washington College digital collections to a wider range of audiences, but also helps the library leverage academic resources with technologies and provide a long-term preservation solution for college archives.

The college library joined the NITLE DSpace in November 2006 and the implementation started immediately. Since NITLE was responsible for the server, configuration and customization were the major tasks at the time. The customizations mainly included user authentication, RSS feed, the display of news, DSpace license, Creative Commons License in submission, strength display of items per collection, advanced search operator, browsing thumbnails, the list of search results and submission form. Users can easily find out these features via the NITLE DSpace search interface and submission process.

Figure 1. Washington College DSpace Homepage



This is a typical NITLE DSpace homepage. Browsing collections and advanced search are on the left side of the page. Users can easily navigate in the NITLE DSpace communities and collections by changing the community or collection. The keyword search results are listed by community hits first, then collection hits; item hits are shown in the last and sorted chronologically.

The submission process is an important part of DSpace software. Currently, seniors do not directly submit their work via DSpace, but via email and blackboard. The metadata librarian is responsible for creating descriptive metadata and submitting theses to DSpace. Thus the library can keep consistent and high quality metadata for all collections. The workload has been manageable with the use of item templates created by the metadata librarian. Below is a standard submission form without /with an item template. The submission process is shown on the top of the form. The highlighted icon indicates the current workflow status. Since the metadata librarian designs the item template for each collection, she decides which element could have a pre-filled value. It is an efficient way to create large amounts of descriptive metadata for the same collection with a well-designed item template.

Figure 2. NITLE DSpace Standard Submission Form

Figure 3. Use an Item Template in the NITLE DSpace Submission Form

Submit: Describe this Item

Please fill in the requested information about this submission below. In most browsers, you can use the tab key to move the cursor to the input box or button, to save you having to use the mouse each time. ([More Help...](#))

Enter the names of the authors of this item below.
Last name *First name(s) + "Jr"*
e.g. Smith *e.g. Donald Jr*

Authors

Enter the main title of the item.

Title

Enter the series and number assigned to this item by your community.
Series Name *Report or Paper No.*

Series/Report No.

If the item has any identification numbers or codes associated with it, please enter the types and the actual numbers or codes below.

Identifiers Other | POL/08-01

ISSN

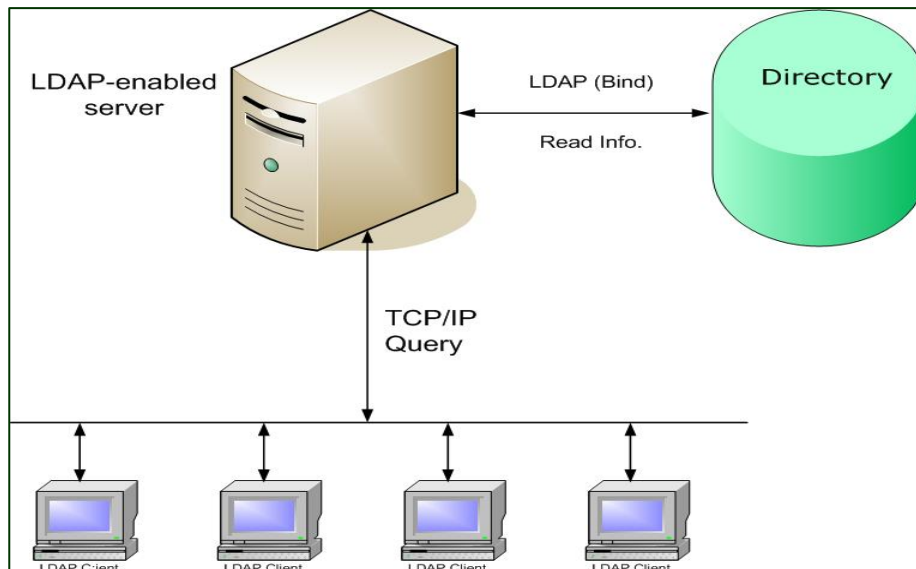
Select the type(s) of content of the item. To select more than one value in the list, you may have to hold down the "CTRL" or "Shift" key.

Type

In addition to users search interface and submission process form, user authentication is crucial to DSpace access restriction. User authentication verifies a person's identity with respect to individually granted access rights. Concerning issues on copyrights for unpublished works, theses excellence, and seniors willing to publish their theses, the library decided to provide access to the theses collection only for the college community. The restricted access gives the library flexibility and time to plan other digital projects in the long run. For instance, collections such as college archives, IR policies, and documentation will be open to the public by assigning anonymous (means everyone) as item and bitstream read group. Some internal reports such as minutes and administration records will have restricted access through limiting item and bitstream read group to the college users with collection authorizations. When users access the restricted collection, user authentication requests user identifications in order to authenticate college users. So how can the library implement the user authentication?

Lightweight Directory Access Protocol (LDAP) is an internet protocol used to access information directories. To be able to set up user authentication service via the Lightweight Directory Access Protocol (LDAP) binding, NITLE needs to run user authentication out of the college LDAP directory to allow DSpace access via the college's campus-based authorization credentials.

Figure 4. Process of Authenticating Clients with LDAP Binding



This is a typical authentication process with LDAP. Clients access information on a protected server, the server binds to the LDAP directory to verify the client information. After the authentication succeeds, the client is authorized to access the information. The LDAP Binding Component is the key to securing all transactions. Through LDAP binding, faculty and students could authenticate against the campus LDAP directory, and be granted access privileges based on their campus profile. The library started to implement the user authentication in February 2007. It took the library around three months to complete the IR implementation. By the end of April 2007, the DSpace repository has been ready for digital collections.

Digital Collection — Senior Capstone Experience (SCE)

Today, most large university libraries deposit electronic theses and dissertations (ETDs), technical reports, archives and grey literature in their institutional repositories. The library wants to start with a highly requested and self-created college collection. The Senior Capstone Experience is a promising one. The library has been preserving the SCE collection for twenty-three years. It is one of the heavily used library collections. When senior students initiate their senior theses, they usually search the collection. Seniors want to know what previous seniors have done or how they can make differences or expand more on the same topics. Without browsing previous senior theses, it is impossible for students to accomplish it. Therefore, creating online access to the Senior Capstone Experience fits the need of both students and the library; students are not only the content users, but also the content contributors.

The Senior Capstone Experience (SCE) has a variety of formats; the SCE could be a thesis, an art production, a design, or a play direction and performance, or a media production. The library has been looking for a permanent home for SCEs, which could provide consistent and sustainable access to the collection for students and faculty.

Creating an easily accessible digital collection will need well-designed information architecture to facilitate users to browse information in the IR. For example, students should be able to browse thumbnails in an image collection by title, subject, or date. Most students like to browse theses titles by department or discipline on the library online catalog. To keep the consistent information structure and web display, the metadata librarian created hierarchical subcommunities for major departments, where the library could help each department determine what would be deposited in the repository. This design not only encourages departments to get more involved in digital projects, but also lets the library macro-manage digital collections throughout the campus.

The library used to collect print copies of senior theses, have an external firm produce a microfiche copy, and add brief bibliographic records to the online catalog. These were time-consuming and labor intensive processes. Even though the library made efforts to preserve theses, none-the-less the collection was not easily accessible. When the library decided to move toward the digital repository and provided online access to SCEs for the college community, the library started to accept electronic senior theses submissions in 2007. Students submit their electronic theses in WORD or PDF via blackboard or email. If a thesis was in WORD, reference librarians would convert it to PDF. To make the best presentation for their research, students expressed their results with images and color photographs; some even used Photoshop software to create illustrations and demonstrate their ideas. Electronic paper gave students more choices to generate creative work. When seniors browsed the theses collection, they saw that the global disclosure of SCEs had brought them some benefits and challenges they ignored before. For example, some graduates even emailed the library and requested to resubmit their theses because they were not satisfied with their paper presentation after they saw papers online. Online access to the SCE collection not only provides convenient services for students as users, but also request students to contribute high quality contents as contributors.

The wider information dissemination also makes faculty members more recognizable as theses advisors on the web than in traditional papers. Users could search the SCE collection by advisors' names through the NITLE search engine, Google or OAIster. More exposures also encourage faculty to improve the SCE course, and students to seek for theses excellence. Although the library restricts the SCE access only to the community, librarians still receive some written requests to download theses from non-college users. In the NITLE community, most participating colleges are more conservative in providing open access to the Electronic Theses, faculty concern about a variety of thesis qualities, and the copyright for unpublished work. The NITLE consortium needs

to work with affiliated colleges to share the best practices of improving theses quality and copyright compliance to facilitate scholarly communication in an IR, and eliminate the worry of open access.

In order to promote open access to research information and utilize the IR to disseminate information, the metadata librarian deposited her conference presentation, technical reports, and manuals in the IR. Librarians know that open-access resources will allow more users to share and disseminate information broadly and promptly, and authors get more feedbacks directly from users as its return. Open-access resources will also help libraries alleviate the pressure of rapidly increased subscription fees. Libraries have been faced with increased journal subscription costs in recent years. Creating open-access resources in academic and research libraries could be a solid solution for libraries to reduce such a burden. The American Chemical Society (ACS) offers authors an option to deposit copies of manuscripts in IRs by paying fees. Some scholars have done so in order to provide open access to their publications. Although an IR as a platform for scholarly communication is not fully recognized in the research community, it is still worth faculty exploring it. Instead of depositing scholarly articles, faculty can begin with small research collections, or presentations and learning objects, to experience the benefit of the IR.

Metadata Creation and Digital Preservation

Without metadata, users cannot access and discover digital collections in the IR. Consistent metadata not only lets users retrieve digital objects easily in an institutional repository, but also helps librarians preserve digital collections. Factors such as types of digital objects, digital tools, and end-user needs determine the adoption of the metadata schemas for digital tools. Nevertheless, the determination of metadata schemas was becoming more of a challenge because it was more difficult for libraries to be the only source of metadata (Llona, 2007). For instance, many publishers use Online Information Exchange (ONIX) as their metadata standard to distribute electronic information about books and media. Libraries can also benefit from sharing metadata between different metadata schemas and digital repositories. NITLE is an OAIster data provider; metadata librarians have to think about how to facilitate data sharing when creating metadata for digital collections.

The college does not have campus-wide standards for senior theses; seniors write theses in different formats and styles. To ensure that seniors submit theses with appropriate and consistent descriptions, some departments prepare SCE guidelines, where information such as keywords, abstracts, table of contents, names of advisors and authors could be easily transferred to descriptive metadata. The metadata librarian uses Collaborative Digitization Program: Dublin Core Metadata Best Practice (CDPDCMBP) as guidelines to create descriptive metadata for the SCE collection. However, NITLE DSpace metadata element labels do not exactly reflect our collection requirements. For instance, users are more interested in the submission date instead of issue date (the date when a publication is issued in the IR). Therefore, the metadata librarian create dateSubmitted element to describe the date when seniors submit theses to their departments. Unfortunately, the dateSubmitted is not indexed and displayed in brief item records. It is not practical for NITLE to customize database indexing and to record web display for various collections held by different colleges; but metadata librarians can articulate this difference in the SCE Metadata Data Dictionary, and map the local metadata element labels to NITLE Dublin Core metadata elements.

Using controlled vocabularies as keywords is another challenge for seniors at the college. The metadata librarian suggests students search thesauri to assign keywords to their theses. Students usually look up Art and Architecture Thesaurus(AAT), Library Congress Thesaurus for Graphic Materials (LCTGM) and Getty Thesaurus for Geographic Names; but it is still difficult for most students to search AAT, LCTGM, and assign controlled vocabularies without additional training. Therefore, the metadata librarian has to review all subjects to ensure seniors use appropriate controlled vocabularies as search terms to enhance the searchability.

It is challenging to create metadata for diversified collections with a single metadata schema in an institutional repository. Dublin Core (DC) metadata schema is the simplest schema; it has a one-to-one principle (which means one metadata description represents only one resource). Therefore, metadata creators cannot use just one Dublin Core metadata record to describe the original analog object and the surrogate digital item. In fact, when researchers browse a digital image, most of the time they are interested in the information of the original work. If the primary metadata schema of NITLE DSpace were Metadata Object Description Schema (MODS), it would be easier to provide related original information for the digital object by inserting MODS relatedItem

element. However, Qualified Dublin Core is the primary metadata schema in NITLE DSpace; the metadata librarian has to find out how to satisfy this requirement for researchers in one metadata description. According to CDP (Collaborative Digitization Program, 2006) Dublin Core Metadata Best Practices version 2.1.1, the metadata librarian uses qualified Dublin Core element refinements to distinguish the metadata between an original analog object and its digital surrogate, and allow the mixed metadata of the two objects in one metadata description. For example, the metadata librarian can use `dateOriginal` element and `dateDigital` element to describe the date of an original analog image and the date of its digital surrogate separately. This is a practical way to enable users to locate interested information in one record.

Metadata preservation is as important as creation. Digital contents are exponentially increasing with user demands and emerging technologies. It is challenging to preserve a huge amount of metadata in an appropriate manner and let users easily discover the information with less clicks. Metadata preservation is not only about updating metadata, identifying errors, and guaranteeing data storage, but also about innovation on how to create an effective method to maintain metadata, and how to create a model to make metadata creation easy and consistent. As the result of creation, librarians can preserve metadata systematically without errors. The end users would be the final beneficiaries of this process. As Hixson (2006) pointed out, digital preservation was a much more proactive process compared to the preservation of analog materials. Librarians should make a greater effort to develop standards and models for digital preservation. A Metadata Data Dictionary is a useful tool to control metadata quality. A Metadata Data Dictionary (MDD) is a table that lists a set of collection-specific local elements, their mapping to the elements in a standard scheme, and specifications of the properties of the elements or fields, including the use of controlled vocabularies (Miller, 2008). It allows metadata creators to follow the same practice for similar collections. Metadata creators and contributors can create consistent metadata with compliance of the data dictionary of the collection, which also facilitates metadata migration, metadata design and sharing. Concise and meaningful metadata releases students from the burden of providing additional and redundant descriptions for their papers, and improves submission efficiency. For example, some seniors have a long list of department members on their cover pages; whereas only an advisor's name is required metadata for the work description. Others use half a page to describe their granted degrees, which could be replaced by degree names. The library could avoid this by standardizing the information on the thesis cover page to eliminate the redundancy. Metadata librarians may also integrate some tools to bring more functionality into the IR. For instance, Google Analytics can track and gather metrics about your institution's DSpace instance. Timely upgrading software will bring new features to the IR and improve IR maintenance. If some users want to perform comprehensive searches across databases, librarians and IT professionals could develop a federated search to enable the function for users.

Cross-Functional Collaboration

In the process of building an institutional repository, cross-functional collaboration is a critical drive to get people interested in an IR, and leads to a successful IR experience. Metadata librarians could not truly implement the digital projects without the involvement of other departments. The collaborative role that metadata librarians played was different in that it was also a consultative role, requiring flexibility and negotiation as all parties designed solutions that were mutually acceptable (Chapman, 2007). Librarians need to gain a broader perspective on an IR from administrators, faculty, IT staff, and students to facilitate the collaboration with different parties. During the DSpace implementation period, the library director took a major role in coordinating communication between administrators, librarians, IT staff, and NITLE. She led meetings to identify implementation difficulties and brainstormed the solutions from all parties. When NITLE implemented the IR access control for the library, NITLE staff had to run user authentication out of the college LDAP directory. The college OIT staff addressed concerns about user privacy. The director initiated a special meeting to let OIT staff explain these concerns, and listened to their suggestions. The system librarian also investigated how NITLE dealt with user privacy with other participating colleges. The library found out that most participating colleges had user privacy agreements to protect users information, and NITLE had not heard any negative feedbacks from users yet. Although it took a while for the library and OIT to implement LDAP binding, the communication and further mutual understanding greatly helped clear up such concern, and finally solved the problem via the director's coordination among NITLE, OIT, and the library.

In certain circumstances, librarians work side by side with faculty. Faculty give their support through class instructions and the SCE course. For example, reference librarians helped the Sociology Department look up

and acquire a thesaurus for students to assign keywords to their papers with controlled vocabularies of sociology. When the metadata librarian set up descriptive metadata template, the Business Management Department sent their feedbacks on indexing and record web display. Some professors thought it was more appropriate to display thesis advisors in description field rather than thesis contributors. This adjustment clears up the ambiguity between content contributors and theses advisors, and eliminates faculty members' doubts about their role in senior theses. Besides faculty members, library staff offers assistance through staffing, supplying software, and digitizing manuscripts. The collaboration inspires more inputs across departments, enables the library to brainstorm solutions and hear different voices from different parties.

Working with NITLE is a learning process. When librarians have questions on IRs, NITLE staff always responds to the questions quickly. For instance, in order to transfer metadata from the DSpace to Integrated Library System, the metadata librarian requested NITLE to generate a Dublin Core metadata sample for a test. NITLE staff provided the sample record right away and allowed the library to initiate the project promptly. Through this project, the library has made the 2007-2008 SCE collection available in both the Integrated Library System (ILS) and DSpace. In addition to problem solving, NITLE also supports the college to improve information management with emerging technologies. For instance, NITLE Technology fellows shared their expertise and experience to improve DSpace interface with Manakin and style sheets. NITLE designed a metadata batch loading form to facilitate metadata transformation between information systems. The collaboration with the NITLE community speeds up the delivery of the benefits of IRs and helps the library gain confidence in this ever-changing digital world.

Marketing an IR

Most academic libraries are aware of the importance of marketing and outreach strategies for an institutional repository. However, these libraries are still slow in IR content recruitment compared to the growth of contents in electronic publications. Although librarians, who could see the long-term benefits of an IR, pioneer in developing an IR, it would be more successful with the participation of scholars across boundaries. Gierveld (2006) emphasized that librarians should notice social marketing because it required behavioral changes on the part of researchers. Faculty members at the college have been active in applying technologies to their teaching, but it might take longer for them to see the benefit of an IR. We find that it is challenging to bring faculty publications to the IR because it is not clear to them how an IR can help faculty with scholarly communication in the same way as traditional publications can. Copyrights and IR credentials are also questionable to them. To eliminate these concerns, librarians need to work together with faculty on these issues and to continue exploring scholarly publishing in the IR.

To market IR collections, librarians need to make use of all possible media to notify the college community. Students used to request print senior theses at the circulation desk. When the SCE collection was available online, students still asked for theses at the counter rather than searching the collection online. To change user behavior, reference librarians posted the news on the library homepage to inform the college community that the Senior Capstone Experience collection was available online and searchable. In addition to the official announcements, information literacy instructions and library brochures provide great opportunities to market DSpace. It is also a good idea to talk about the IR with college public relations, alumni, and the college newspaper writers. Unfortunately, we have not talked about the IR in such an exhaustive manner. From our experience, word-of-mouth is one of the most effective methods of marketing and promoting an IR. The more people know about the IR, the more likely they use it, and the more possible the IR will benefit the college community.

Lessons Learned

Although we started to explore an IR just two years ago, we have learned a great deal from the SCE project. Looking back, we would like to share our experiences with those who plan to initiate a digital project at a small liberal arts colleges. Since the construction of an institutional repository is a long-term, on-going project, sustainable support from the college administration is indispensable. Librarians and information technologists should be aware of emerging issues in the field. Among them, the future of institutional repositories should be in the library strategic plan. Some emerging issues include the following. Where is an IR located? Who will have

access to scholarly publications? How can libraries deal with dark archives and grey literature? Should libraries allow dark archives harvested by data services? Some libraries restrict the access to certain collections; others have open access to all collections. Some colleges request faculty to deposit their work in an IR, while faculty members still hesitate to do so. Digital Initiative Librarians might not be fully aware of all these issues when originating digital projects, but libraries should have the flexibility to work on these on-going issues instead of closing the channel of communication. An institutional repository should never become a fashionable infrastructure in some administrative circle (Lynch, 2003). It should be a new model of scholarly communication and an additional vehicle to bridge the gap between knowledge management and learning. In order to target this objective, librarians need to pay attention to content acquisition, IR documentation, and collaboration.

Content Adoption

The source of a digital collection is content adoption. Besides having a clear collection policy, it is necessary to think about user culture and behavior, which could be a good catalyst leading to a healthy IR environment. When digital initiative librarians recruit scholarly work from faculty and students, they have to understand faculty concerns on authorship, peer-review, tenureship and the authority of the institutional repository as a publisher. It would be more constructive to have an effective means to discuss these issues on campus and invent a new and reliable channel of communication to facilitate electronic publishing in an institutional repository. The top-down model might work well on some campuses, but it is also important to give faculty and students flexibility.

It benefits both users and content creators to provide access to grey literature, such as technical reports, datasets, and preprints. For example, it is impossible for scholars to publish datasets in the same way as scholarly articles published in traditional publications, but scholars can publish datasets in an institutional repository. The purpose is to keep datasets secure and easily accessible. Researchers could share experimental data with more peers in the same domain to get professional perspectives, and further leverage their research value. When information professionals update technologies, actions such as upgrading software and equipment and migrating data are inevitable. Even though the cost of preserving digital materials has come down in recent years with the development of digital technologies, it is wise to keep the most valuable collections in an institutional repository instead of unquestioned content adoption. It is also necessary to educate faculty members to build their own digital collections in an IR. In such cases, online tutorials will be an introduction to digitization; faculty can begin to experience a digital project with tutorials. Once faculty become interested in creating a digital collection in an IR, we no longer worry about getting high quality contents in the IR.

IR Documentation

In most case, librarians easily ignore IR documentation when initiating a digital project in an IR. Generally speaking, there are two types of IR documentation for different users. Metadata Data Dictionaries, Metadata Best Practice Guides, Metadata Application Profiles and Metadata Registries are the documents for librarians, paraprofessionals and IT staff. Collection policy and user manuals are for the public to get a general idea of an IR. We should have created Metadata Best Practice Guidelines and the Metadata Data Dictionary for the SCE collection before we actually worked on the SCE project. Although we used Collaborative Digitization Program: Dublin Core Metadata Best Practice (CDPDCMBP) as metadata creation guidelines, we should have addressed how the library created metadata for the SCE collection and how the metadata librarian chose element labels and assigned the value for each element with controlled vocabularies. The Metadata Data Dictionary should have defined all these issues. For example, in SCE Metadata Data Dictionary, the metadata librarian can list a set of SCE metadata elements, their mapping to the elements in a standard Dublin Core scheme, specifications of the properties of the elements or fields, and the use of controlled vocabularies. These documents serve as local standards to guide metadata creation, control metadata quality, preserve and share metadata in different metadata schemas or repositories.

To ensure that faculty and staff will know what are appropriate materials to be deposited in an institutional repository, it is important to publicize on the library website the collection policy and minimal metadata requirements for digital objects. Faculty, staff, and students will appreciate the Question and Answer section, samples of the digital collection, and descriptive records on the web. These documents are helpful for users to get an initial knowledge of copyrights and digitalization policies, and to know what an open-access

resource is and how to protect author rights. The college library should be ready to provide assistance and services on a campus-wide basis, and assure that everyone receives professional, courteous, and timely help.

Collaboration & Knowledge Sharing

To have a highly effective campus-wide collaboration, some committees, such as, the Digitization Steering Committee, Scholarly Publishing, or Electronic Publishing Committee, could play an important role in promoting an institutional repository, and advise scholars how to get started or involved in digitization and scholarly publishing. Collaboration across departments might begin with students digitizing a special collection or a case study. A case study of user behavior could assist librarians to understand how users search information in an IR through observing their submission and browsing information. Suggestions from faculty and students could also contribute to the improvement of a user-friendly interface.

Sharing knowledge with the NITLE DSpace consortium greatly benefits the college community. There are 27 liberal arts colleges using NITLE DSpace. Sharing best practices with the NITLE DSpace community helps affiliated institutions get hands-on experience to initiate a digital project. Professionals in the consortium propose questions, discuss emerging issues, and learn from peer institutions through NITLE-DSpace-Community and online Multipoint Interactive Videoconference (MIV). Technology Fellows lead discussions and teach courses via MIV cross the country. Each year, the NITLE DSpace User Community Meeting is another opportunity to acquire new knowledge and share experiences with DSpace users from internal or external consortium. Intensive discussions also assist librarians to keep abreast of emerging trends and technologies. Technology professionals, librarians, and library administrators have been a valuable knowledge source for the consortium. Cross-institutional collaboration deepens the understanding of digital technologies and promotes the application of an institutional repository.

Building an institutional repository is a task for the college community, not just for the library. Without the support of members of the community, an institutional repository cannot succeed and achieve its goals. Librarians should not hesitate to reach out and ask for help from other departments, administration, or even outside institutions. Support from college administrators is also a strong incentive to the library and a campus-wide influence. As long as librarians build an institutional repository with an eye to its future and play the role of leadership and collaboration, faculty and students will benefit from the scholarly communication in an IR.

Conclusion

A liberal arts college can create an institutional repository by joining a consortium with the benefit of knowledge sharing in the consortial community. The content development in an IR requires an open channel of communication to understand user behaviors, requirements, and concerns. Metadata preservation is as important as metadata creation. Metadata documentation is an important tool to control metadata quality and share best practices. An institutional repository will achieve its goal with a thoughtful plan and the collaboration of all related parties. When people ponder the future of an institutional repository, libraries can be proud of the prosperity of scholarly publishing and communication in an IR.

References

- Barrow, D. (2004). Networked collaboration transforms curricula. *Liberal Education*, 90(2), 48-53. Retrieved May 26, 2009, from the Academic Search Complete database.
- Chapman, J. W. (2007). The roles of metadata librarian in a research library. *Library Resources & Technical Services*, 51(4), 279-284.
- CDP Metadata Working Group. (2006). Dublin Core metadata best practices. Retrieved May 4, 2009, from <http://www.bcr.org/cdp/be2009st/index.html>
- Gierveld, H. (2006). Considering a marketing and communications approach for an institutional repository. *Ariadne*, 49. Retrieved Sept. 11, 2008, from <http://www.ariadne.ac.uk/issue49/gierveld/>
- Hixson, C. (2006). First we build them, then what? : The future of institutional repositories. *BiD, textos universitaris de biblioteconomia i documentació*, 15, 1-7.

- Jantz, R. C. & Wilson, M. C. (2008). Institutional repositories: Faculty deposits, marketing and the reform of scholarly communication. *The Journal of Academic Librarianship*, 34(3), 186-195. Retrieved January 22, 2009, from the WilsonWeb database.
- Llona, E. (2007). The librarian's role in promoting digital scholarship: Development and metadata issues. In M. Remnek (Ed.), *East European and Eurasian Culture: Publishing, Acquisitions, Digitization, Metadata*, (pp. 151-163). Philadelphia: Haworth.
- Lynch, C. (2003) Institutional repositories: Essential infrastructure for scholarship in the digital age. *ARL 226*, 1-12. Retrieved July 7, 2008, from the WilsonWeb database.
- Marcondes, C. & Sayão, L. F. (2003). Brazilian Digital Library of Theses and Dissertations. *The International Information & Library Review*, 35(2-4), 265-279. Retrieved May 27, 2009, from the ScienceDirect database.
- Miller, S. J. (2008) Metadata design, management and documentation. In *Metadata for Digital Collections, an online workshop*. Unpublished manuscript, University of Wisconsin at Milwaukee.
- Morgan, P. (2007). Facilitating the deposit of experimental chemistry data institutional repositories: Project SPECTRa (Submission, Preservation, and Exposure of Chemistry Teaching and Research Data). *IATUL Proceedings 17*, 1-8. Retrieved November 18, 2008, from the WilsonWeb database.
- Prosser, D. C. (2003). The next information revolution-can institutional repositories and self-archiving transform scholarly communications? *IATUL Proceedings 13*, 1-9. Retrieved December 20, 2008, from the WilsonWeb database.
- Reilly, W., Wolfe, R., & Smith, M. (2006). MIT's CWSpace project: Packaging metadata for archiving educational content in DSpace. *International Journal on Digital Libraries 6*(2), 139-147. Retrieved February 15, 2009, from the Academic Search Complete database.