Case Study of Humboldt Digital Scholar

Prepared for the CSU STIM Digital Repository Working Group
By Jeremy C. Shellhase
October 21, 2010

Objectives:

- **Institutional goals –** The project was conceived in 2004 as a response to the perceived need for preservation and organization of digital scholarship at Humboldt State. The increase in visibility and access for faculty and student scholarship that participation in the international network of Dspace and OAI institutions could bring was an equally important motivation. The project was presented as a pilot using existing library resources and staff while soliciting voluntary participation from other university stakeholders.

- **Expectations for the project when it was started –** It was expected that online electronic masters theses would be an early accomplishment and that faculty participation would develop more slowly. Units with active web presences, such as the Schatz Energy Research Center, were given priority assistance in getting started. The expectation was that offering the service would provide a form of success in establishing the value of institutional repository services for smaller schools and making Humboldt State's scholarship more widely available.

Processes:

- **Does the organization have written policies?** - The policies are written and posted on the Dspace site itself.

    What material is being collected? - The collection policies simply require that the submissions are scholarly or creative products of the University and that all collection be sponsored by either a faculty member or an official institution of the University. The submitter must also agree to a license that states that they have the rights to make the materials publicly available. The definition of what is scholarly or creative is up to the submitting entity, so no material is specifically excluded from the archive.

    Is any material specifically not collected (class materials) - Administrative materials, such as WASC reports and Academic Senate documents are not included in the archive, because of they are not scholarly or creative output. With limited staffing and support the Library chose to focus all support on the academic side of the University and not undertake an archive or document management function.

What is the process used to upload/manage objects

Are you collecting objects yourself, or allowing uploads by departments or individual faculty?—Objects are ingested by all of these methods. Students are submitting their master's theses themselves, following instructions from the Library. The Library is notified by the graduate office when theses are ready for submission. The students are allowed to submit an electronic copy instead of supplying a bound copy for circulation from the library, so there is a minor economic motivation for them. Some faculty and institutes will perform their own submissions, but library staff will manage submissions for many submissions.

How is metadata managed? What standards are used? - Metadata is Dublin Core and for users submissions no data checking is performed regularly. The students submitting master's theses do their own data entry and the data is reviewed by the library's cataloging department before the items are made public. Revisions are made to program and faculty advisor entries to ensure a consistent browse index for these entries, but user input keywords are not modified. Minor typos are sometimes corrected.

How do you manage copyright and permissions? – User self submissions are required to agree to a copyright permission statement that they have the rights to the content and will make it available from the repository. This is a minority of submissions. For recruited submissions of published works library staff checks on copyright policies for the faculty member and identify those items that can be archived. The required processing is performed by the library staff and manual records are kept of copyright compliance.

Resources:

What personnel are required? What are their roles/duties? –

In a 2009 D-Lib Magazine article we stated the following.

"Currently, a volunteer Steering Committee of four librarians and a Systems staff member, assisted by a copy cataloger, maintain and manage HDS in addition to their other assignments. Several levels of expertise are needed to run the repository successfully. The Systems Librarian and Systems staff member manage hardware, software, customizations, and file conversion; two catalogers handle metadata and thesis submission; the four librarians share the work of developing policy, creating documentation, submitting scholarship for faculty and staff who request assistance, and communicating with faculty, staff, and publishers. Time commitments

vary with the tasks to be accomplished, such as Systems involvement when software or hardware upgrades become necessary, cataloger participation at semester's end when theses are processed, and committee meetings when policy issues arise or outreach efforts are needed. While work on HDS ebbs and flows, we estimate that the ongoing time commitment of the Committee and support staff currently is equivalent to a .5 position."

In the intervening time, the majority of HSU's Librarians have been put on a reduced, ten month time base and several librarians have entered the FERP program. Current level of staffing available could be more accurately estimated at about a 0.25 position.

What is your project budget? – There is not a budget allocated for the project.

*Do you have other support from your campus library, other departments, your institution or other organizations.—*In addition to library staff mentioned, significant support is derived from student and faculty working on their submissions. Technical support is now provided through the Chancellor's Office, with Humboldt Digital Scholar running as a part of the CSU Scholarworks project.

Access:

Is access to objects restricted at all, or is it open to the public? – We have offered the option of embargo on content, but have not encouraged it. To date all content is open for the public.

Barriers:

Please describe any barriers your project has encountered, whether technical or institutional. — Having limited time available for repository work has been the major limiting factor in the growth of the repository. The complexity of some copyright requirements can be daunting, particularly requirements for pre-publication copies of articles. Technically working with Dspace is demanding, but not really a barrier. Working with shared system resources on Scholarworks has simplified some of the technical work, but offers a reduced level of hands on modifications.

Are you happy with the support you have been given? Is it what you expected? --
Support for Humboldt Digital Scholar has been what the project participants expected. The support from the CSU system in managing and adapting the open source Dspace software has been invaluable.

Promotion:

How are you marketing your project? –

 Over the five years of the project marketing and promotion have used a wide variety of methods. Formal presentations were made to administrative bodies, such as the Executive Council and the Academic Senate. A resolution favoring open access archiving of published and unpublished scholarship was passed in the Academic Senate. The repository is a regular part of new faculty orientation and beginning of the year presentations to the colleges. Library bibliographers work with individual faculty from their departments to identify material for submission to the repository. Contact with campus institutes is made by representatives of the Steering Committee to encourage participation.

Do you have incentives for faculty/student participation? --

Students submitting master's theses do not have to pay for a copy for the library's circulating collection. The potential of making their thesis more widely available works as an incentive for many students.

The incentive for participating faculty and institutes are primarily those of preservation and access. The creation of a personal page has not been much in demand by faculty, but the simple linking to permanent copies of their works has been a motivator.

The San Diego State University Library & Information Access began efforts to digitize collections and establish a digital repository in 2005 with the hiring of a full-time Digital Collections Librarian. This individual's concentration was largely focused on digitizing special collections materials. In 2008, however, with the addition of a metadata librarian, the concentration shifted to collecting digital or digitized documents created by the faculty, staff and students of San Diego State University. During this process, librarians began using a DSpace repository (hosted and maintained by the California State University system) as the main repository for institutional documents.

Currently the library has two repository projects underway at varying stages of completion. While there is no documented repository plan available at this time, strategic planning initiatives in 2009 recommended a formalization of the repository process, and final recommendations for a repository strategy are currently being developed by the library executive council.

The two current projects are 1) the digitization of print geological theses and 2) the capture of new graduate theses in digital format.

1) Geology Theses Digitization Project

The library has a collection of over 500 print geological theses, each of which includes large-scale standalone maps and photographs in addition to the thesis text itself. Since the maps are fragile and can become torn with even moderate use, digitizing them will ensure their availability and longevity. The opportunity to make the print geological theses available via a digital repository arose when the library obtained the volunteer services of a retired geologist. Currently, 57 theses have been digitized and made available online at http://sdsu-dspace.calstate.edu/.

I. Goals and Expectations

The ultimate goal is to make all geology theses freely available to the public in digital format via the DSpace repository.

II. Process
   a. There is no written documentation of processes.
   b. Work flow
        i. The digital collections librarian contacts the author for permission to make the document available via the repository.
        ii. The volunteer geologist scans the document and accompanying maps.
        iii. The volunteer geologist adds expert metadata including keywords.
        iv. The librarian verifies the scans and creates a PDF of the document using OCR.
        v. The librarian edits the map scans for readability.
        vi. The librarian adds a low resolution map image to the PDF.
        vii. The librarian uploads the PDF and a high resolution map image to the DSpace repository.

III. Resources
    a. Personnel
        i. Digital Collections Librarian
            a. Verification
            b. Editing of documents
            c. Uploading
        ii. Metadata Librarian
            a. Consultant
        iii. Volunteer
            a. Scanning
            b. Supplying expert metadata
    b. Budget
        i. The project has no budget aside from personnel.
    c. Support
        i. The project utilizes the DSpace repository supplied and maintained by the CSU Chancellor's Office.

IV. Access
    a. All documents included in the digitized collection have no restrictions as to access.

V. Barriers
    a. Some technical barriers (e.g., customization and feature addition) result from not having the repository housed and maintained locally.
    b. The project is functioning well at its current level of support.

VI. Promotion
    a. Currently the project has not been marketed or promoted. Eventually a link to the digital document will be added to the library catalog record for the individual thesis.

2) SDSU Theses Repository Project

All SDSU theses will be added to the digital repository. This project was necessitated by a campus decision to stop supplying a print version of theses to the library.

I. Goals and Expectations

The ultimate goal is to have each new SDSU thesis available online through the library catalog and digital repository.

II. Process
    a. There is no written documentation of processes.
    b. Work flow
        i. The publisher, Montezuma Press, obtains copyright clearance from student.
        ii. The publisher uploads a PDF of the thesis to a web server.
        iii. Library catalogers download the document.
        iv. Library catalogers process the document for inclusion in the library catalog.

        v. Planned: A batch upload process will move document metadata to DSpace repository.

        vi. Planned: The digital document will be uploaded to the DSpace repository.

III. Resources

   a. Personnel

      i. Montezuma Press

        1. Clear copyright

        2. Create digital document

        3. Place on server accessible to the library

      ii. Library cataloging department

        1. Download document

        2. Catalog document

      iii. Metadata librarian

        1. Creation of Crosswalk from MARC to DSpace

      iv. Library IT staff

        1. Planned: Batch upload of document metadata to DSpace

      v. Digital Collections Librarian

        1. Planned: Verification of metadata

        2. Planned: Uploading of digital document to DSpace

   b. Budget

      i. There is no budget for this project aside from personnel.

   c. Support

      i. The project utilizes the DSpace repository supplied and maintained by the CSU Chancellor's Office.

      ii. The publisher provides the document and clears copyright.

IV. Access

   a. All documents included in the digitized collection will have no restrictions as to access.

V. Barriers

   a. Some technical barriers (e.g., customization and feature additions) result from not having the repository housed and maintained locally.

   b. There have been some issues with enabling a batch upload process to the CSU-hosted DSpace.

VI. Promotion

   a. Planned: Documents will be discoverable via the library catalog.

The DigitalCommons@CalPoly:

A Case Study on the Development and Implementation of

the Institutional Repository

for

California Polytechnic State University, San Luis Obispo

Case Study prepared for the
COLD Digital Repository Working Group

By:

Marisa Ramirez, Digital Repository Librarian
California Polytechnic State University

October, 2010

Background

Institutional repositories (IRs) first gained widespread attention in 2002, when universities began offering in-house services to manage and disseminate digital materials created by students and faculty. These repositories developed in response to profound changes, both in the needs of faculty and students and in scholarly communication itself. As the Scholarly Publishing and Academic Resources Coalition (SPARC) notes, "…This increased visibility [and] demonstration of value can translate into tangible benefits, including the funding – from both public and private sources – that derives in part from an institution's status and reputation."[1]

The Cal Poly Provost and the Library Director viewed an IR as a mechanism to increase the visibility of - and access to - scholarship by Cal Poly faculty, students and campus constituents. At the same time, Cal Poly would join its peers in higher education by contributing to the emerging model of distributed, interoperable repositories that provide the foundation for the new model of scholarly publishing. The Provost pledged an initial 3-year funding commitment to the Library, with expenses from subsequent years to be covered by other means.

At the request of the Provost, a Task Group consisting of Library and Campus Information Technology Services convened in 2006 to investigate the feasibility of launching an IR at Cal Poly. After investigating the resources required to launch a campus IR using open-source and hosted repository options, the committee arrived at the recommendation of using a hosted repository system (Berkeley Electronic Press' DigitalCommons platform), and hiring two positions (Digital Repository Librarian and Library Assistant).

The Kennedy Library deliberately chose a hosted software solution from Berkeley Electronic Press to keep new staffing at minimum and to focus on contributor recruitment and content management. The Digital Repository Librarian would reach out to contributors and to manage digital assets and their metadata. A support staff position would secure intellectual property permissions and communicate with publishers. Using student assistants, the repository staff members would manage the internal process of bibliographic searching and digitizing articles for ingest into the repository. Organizationally, DigitalCommons staff was located in Special Collections & University Archives to benefit from existing collections and knowledge of the campus. A later reorganization has relocated staff to report to the Director, Information, Resources and Archives within the Library.

---

[1] Crow, Raym (2002). The Case for Institutional Repositories: A SPARC Position Paper. *The Scholarly Publishing and Academic Resources Coalition (SPARC)*. Retrieved from http://www.arl.org/sparc/bm~doc/ir_final_release_102.pdf

Staffing

The Digital Repository Librarian position was filled in October 2007 and the Library Assistant position was filled January 2008.

The Digital Repository Librarian is a tenure-track position which has the primary responsibility for the growth, development and maintenance of the library's institutional repository and digitized archival collections initiatives, including the full range of work related to program and policy research, evaluation, and analysis. The essential components of this position are to maintain and promote digital repository collections and services provided by the Kennedy Library, including developing and securing content, communicating with and providing support to contributors and users, drafting digital repository policies and procedures, working with Library Information Technology and vendors to ensure system functionality, and establishing workflow and other procedures according to best practices and established standards.

The Digital Repository Assistant (LAI/LAII) position is responsible for contributions to the institutional repository managed by the library, including bibliographic searching, requesting copyright permissions from publishers and authors, tracking and processing submissions from contributors, creating and uploading digital files and related metadata in approved formats, supervision of student assistants working on repository tasks, and other standard digitization workflow tasks.

About 20 hours/week of student assistant time is devoted to activities such as scanning, processing and ingest of content.

Implementation Timeline

While a phased implementation plan was originally envisioned, delays in paperwork resulted in an abbreviated pilot period with launch in September 2008.

Institutional Repository Task-Force (2006 – May 07)
A task force consisting of Library and ITS staff reviewed institutional repository software and implementation models. A consultant, Sayeed Choudhury from the Johns Hopkins University, delivered an on-site talk to invited campus guests including faculty, administration and staff groups regarding the value delivered by an institutional repository. The Task Force work culminated in an IR recommendation to the Library Dean and Provost.

Recruitment for Digital Repository Librarian Position (May 2007 – September 2007)
Marisa Ramirez was hired for the position starting October 2007.

Recruitment for Digital Repository Assistant Position (LAI/II) (October 2007- December 2007)
Vanessa Woods was hired for the position starting January 2008.

Pilot Phase (Nov 2007 – June 2008)

This first phase was devoted to IR customization, purchase of supporting hardware and software, policy and marketing development, hiring of support staff and students, cultivation of pilot partners and development of the service model. The Digital Repository Librarian approached faculty groups that were "friends of the library" to provide feedback on initial implementation plans and services. The service model, marketing materials and workflows were refined based on this feedback.

Launch Phase (September 2008)
This phase was characterized by a burst of marketing and outreach efforts by the Digital Repository Librarian, including presentations, meetings, software demos and development/distribution of marketing materials to campus administrators and faculty leaders.

Subsequent Piloted and Launched Services
- Electronic Submission of Master's Theses
  Piloted May 2008 – August 2008; Launched September 2008

- Electronic Submission of Undergraduate Senior Projects
  Piloted May 2009 – August 2009; Launched September 2009

Service Model

"The phrase 'if you build it, they will come' does not yet apply to IRs. While their benefits seem to be very persuasive to institutions, IRs fail to appear compelling and useful to the authors and owners of the content."[2]

While self-submission of content by contributors was originally envisioned, once the IR was launched, it was clear that a mediated-deposit model - instead of author self-submission - would best expedite population of the IR. This decision was supported by research about barriers for author self-submission, including copyright concerns and additional time and effort required for self-archiving.[3]

The DigitalCommons@CalPoly is open to the Cal Poly academic community to contribute completed scholarship and relevant university materials created by administrative offices, departments and programs at Cal Poly for long-term preservation and worldwide electronic accessibility through the IR.

---

[2] Fried Foster, Nancy and Gibbons, Susan.(2005). "Understanding Faculty to Improve Content Recruitment for Institutional Repositories." *D-Lib Magazine 11* (1): 1. Available from http://www.dlib.org/dlib/january05/foster/01foster.html

[3] Kim, Jihyun. "Faculty Self-Archiving: Motivations and Barriers." *Journal of the American Society for Information Science and Technology 61* (9): 1917.

To contribute content to the DigitalCommons@CalPoly, faculty contributors are asked to download the DigitalCommons non-exclusive license, and send the signed copy of the license along with a list of citations to their work to the Digital Repository Librarian.

The Digital Repository Librarian, in coordination with the Repository Assistant and Student Assistants, locates copies of the work from the citations, clears copyright permissions with publishers, and posts the work. An online personalized area of DigitalCommons@CalPoly is created for authors, and the posted work appears in this personalized profile area. The author receives system generated email notifications as each work is posted. Authors also receive monthly usage statistics of that work, which also serves as an effective means for reminding authors of the value the IR delivers.

The majority of student work follows a self-submission model, because their output typically does not have the same copyright constraints as faculty work and the motivation for self-archiving is more compelling. For example, as a requirement for graduation, master's students are required to deposit a copy of their work with the library using the DigitalCommons system. While it is optional for undergraduates to provide a copy of their senior project to the library, students often wish to deposit their work because it lends a Cal Poly imprimatur upon their work. Additional information about the master's theses process is available at: http://www.calpoly.edu/~rgp/gradthesis.html. Information about senior projects is available at: http://lib.calpoly.edu/seniorprojects/.

Marketing

Well before the DigitalCommons@CalPoly (http://digitalcommons.calpoly.edu) was launched in 2008, the library recognized the importance of marketing the new service. Prior to funding the project, the Library Dean and Provost discussed the advantages of an IR to the campus. Presentations were made to Academic Personnel (the faculty human resources office), and the Academic Deans Council among others. High visibility faculty and campus leaders were invited to the "kickoff" event, featuring a presentation by Sayeed Choudhury, a nationally recognized leader in repository development. Shortly after a Digital Repository Librarian was hired and the software was in place, small-scale marketing efforts took place with pilot groups of faculty and campus leadership. After the first month of launch, the library hosted a day-long colloquium to broaden the awareness of the repository among campus leaders, faculty and library colleagues from fellow California State University campuses.

With several other technology initiatives taking place on campus, it was important to clearly communicate what the IR was, and was not. The repository was marketed as a set of services to enhance the visibility and availability of creative, scholarly and intellectual work by faculty and students. It was not an e-portfolio system, nor was it a productivity tool to monitor faculty activity for promotion-tenure purposes. While there are many benefits to contributing to the IR, some benefits were emphasized more than others, depending on the background and needs of the audience.

There was clear campus support behind the notion of an IR, but it was difficult to clearly communicate the concept of a repository to faculty during the early stages of IR implementation. Faculty enthusiasm about the repository was boosted after implementing a concrete instantiation that was populated with exemplar content. Instead of delivering abstruse talks to faculty, the repository manager's presentations consisted of demonstrating the live site, which provided concrete examples of how the repository looked and worked. It was much easier to show (instead of tell) faculty how the service could further their teaching and research goals. The repository, in effect, became its own marketing tool.

One of the key approaches to marketing the IR was to approach it as a service rather than a technology. Early IR literature describes IRs as "a set of services that a university offers to the members of its community for the management and dissemination of digital materials created by the institution and its community members."[4]

Educating library liaisons on the basics of the IR and providing them marketing tools was useful. Library liaisons at Cal Poly are provided an "elevator speech", a brief summary of the repository purpose, services, benefits to participation and instructions on how to get started. The talking points are brief enough to be delivered in the span of an elevator ride, a quick hallway discussion or office visit. Library liaisons are also provided a box of "marketing tchotchkes" to distribute during college or department visits, information literacy instruction or at other times when librarians have contact with faculty. These small inexpensive gifts, customized with the IR logo, range from post-it notes to mini-optical mice and are used to incentivize and encourage participation in the IR. Once interest is piqued, faculty and students are referred back to the repository manager.

Personalization was found to be one of the most effective outreach strategies by the repository manager. For example, faculty respond favorably when they receive customized emails that reference their recently published scholarship, including a journal name and article title. This information is often mined from citation or journal databases, and is included in the personalized email to individual faculty. There also has been success in providing monthly use statistics to faculty, so that they are aware of how often their work is being used. An Annual Report to the Provost summarizing IR accomplishments and growth is another strategy employed to communicate the value the IR to campus leadership. Access an example online at http://digitalcommons.calpoly.edu/lib_dean/22.

Cal Poly's strategy for marketing the IR included: 1) finding opportunities to be visible on campus and discuss the value that the IR will deliver to campus; 2) being open to collaborations with non-traditional library partners, such as Public Affairs or campus institutes, which are rich sources of content; 3) maintaining the flexibility to adopt new marketing strategies.

---

[4] Lynch, Clifford.(2003). ""Institutional Repositories: Essential Infrastructure for Scholarship in the Digital Age." *ARL: A Bimonthly Report 266*: 2.

Workflows

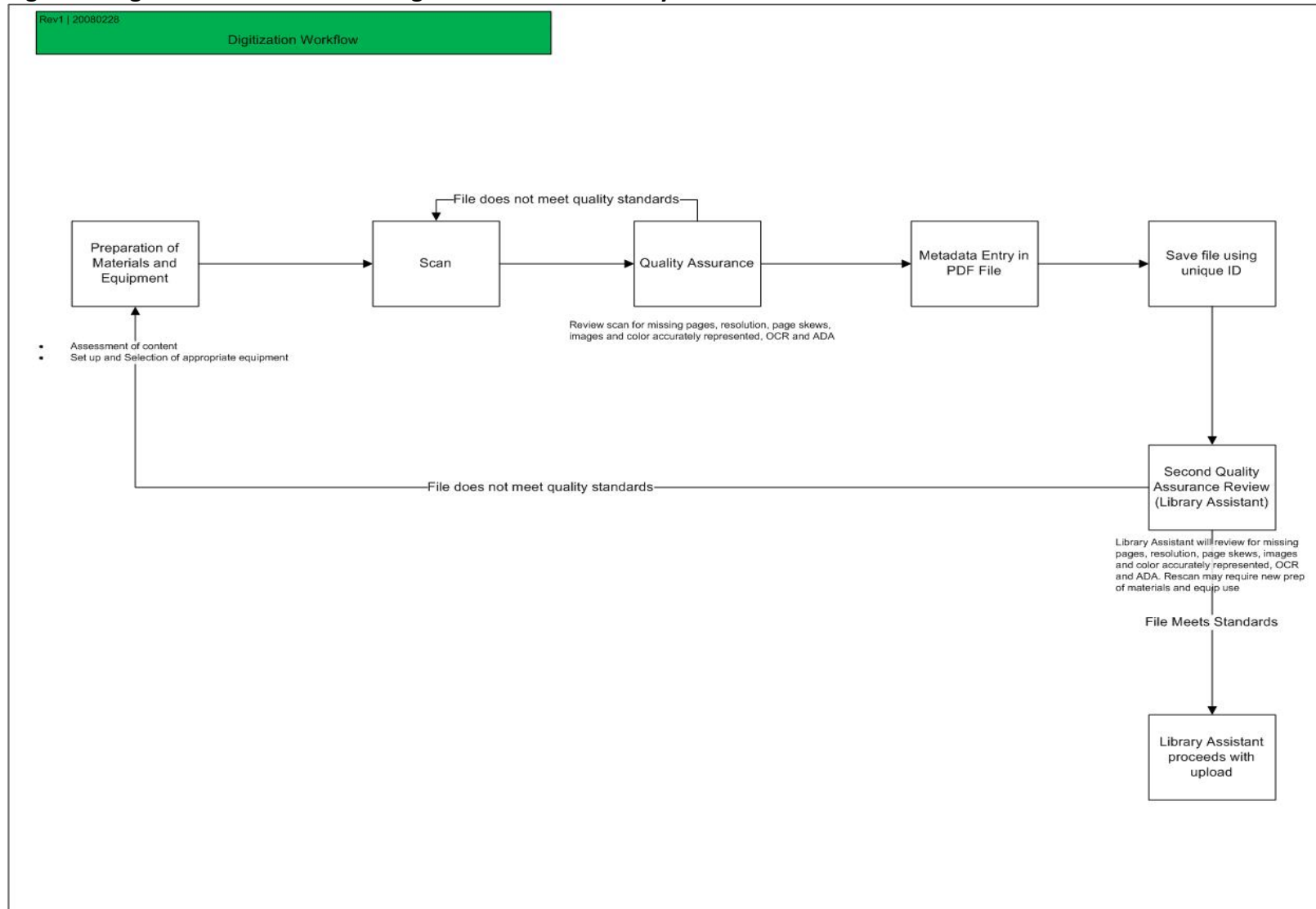### Figure 1: Digitization Workflow for DigitalCommons@CalPoly

Figure 1: Digitization Workflow for DigitalCommons@CalPoly

Figure 2: DigitalCommons@CalPoly Workflow
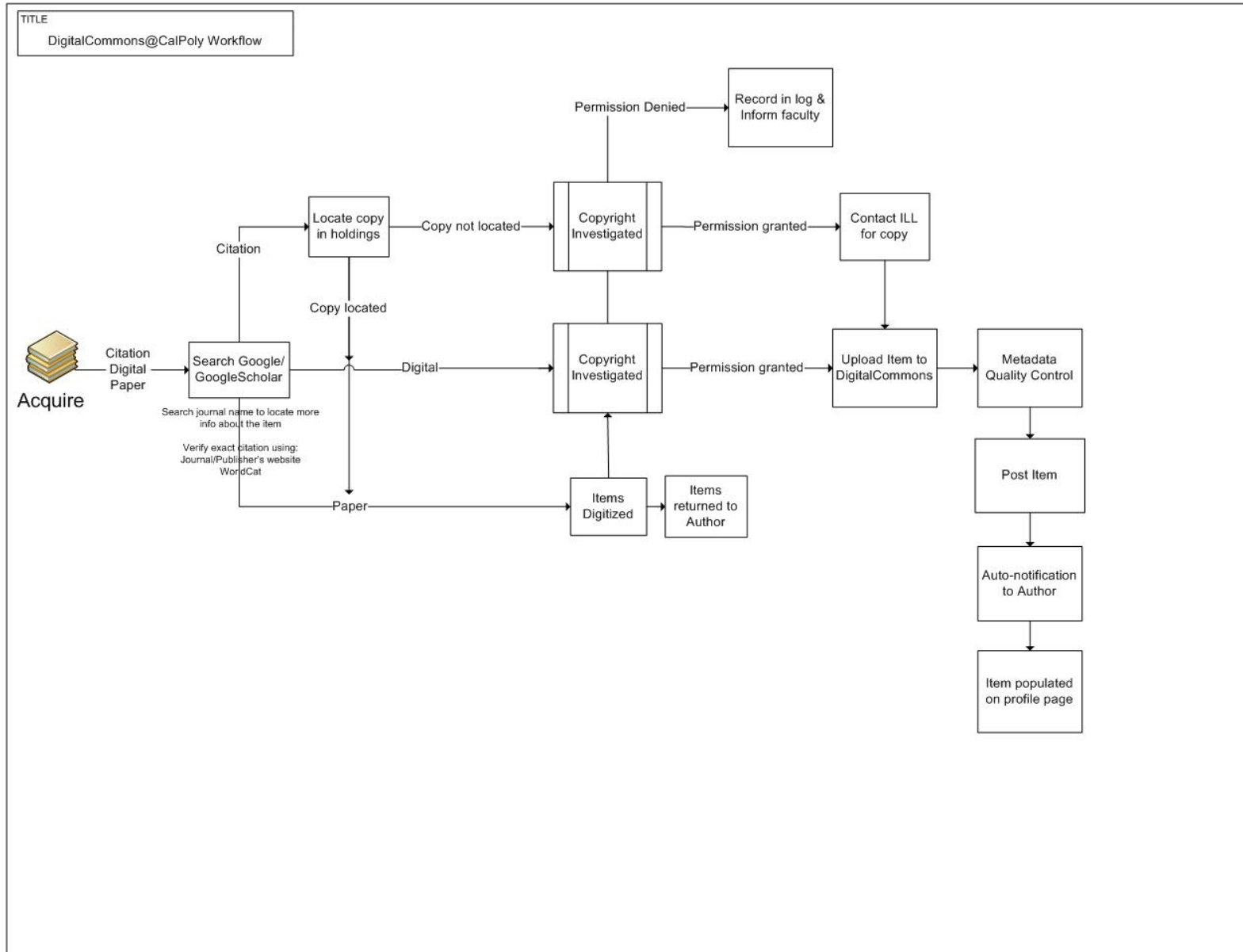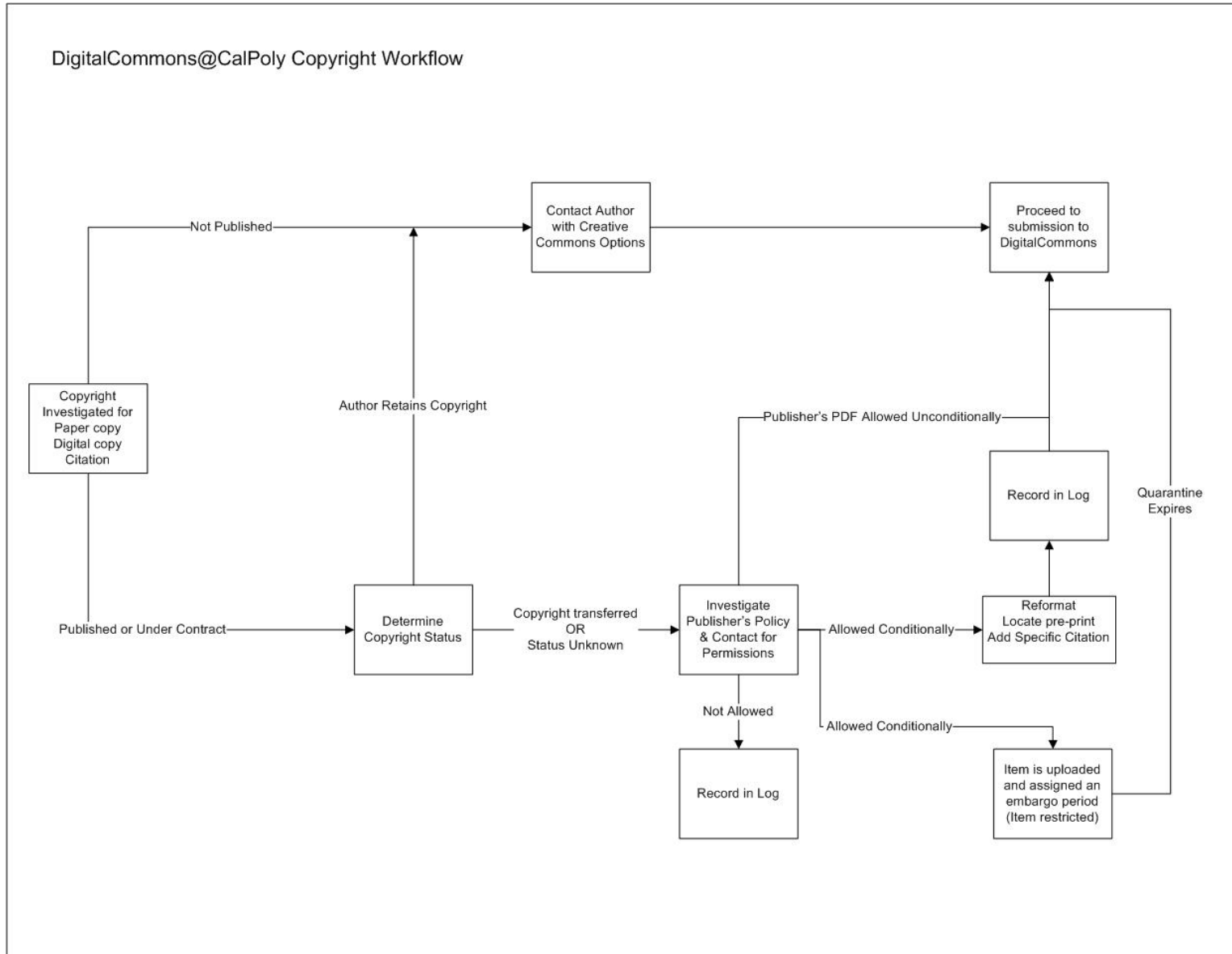
Figure 3: DigitalCommons@CalPoly Copyright Workflow

DigitalCommons@CalPoly Copyright Workflow

Content

From the outset, DigitalCommons@CalPoly was envisioned to provide access to a variety of content, which would reflect the intellectual life of the campus. As it now stands, the DigitalCommons@CalPoly promotes discovery, research, cross-disciplinary collaboration and instruction by collecting, preserving and providing access to scholarly work created at Cal Poly. The repository also provides access to relevant documents created by administrative offices, departments and programs at Cal Poly. Members of the Cal Poly academic community are invited to contribute completed scholarship for long-term preservation and worldwide electronic accessibility through the DigitalCommons.

According to Berkeley Electronic Press, the repository software vendor, the DigitalCommons@ Cal Poly features the broadest assortment of academic and institutional content among all DigitalCommons repositories. Content includes:

	Master's theses
Abstracts	Peer-reviewed journal articles
Alumni publications	Poetry
Annual reports	Posters
Architectural plans	PowerPoint presentations
Campus periodicals	Press releases
Campus photographs	Research from campus institutes
Conference proceedings	and centers
eBooks	Senior projects
Finding aids	Speeches
Grant reports	Staff publications
Images of campus	Undergraduate essays
Internship Reports	WASC reports
Master plans	

In general, contributions to the DigitalCommons@CalPoly are scholarly in nature and the author has a connection to Cal Poly. The contributor must be the creator of the work. Because deposits are intended to be permanent contributions to the repository, works that are in progress or ephemeral (such as drafts) are not recommended for contribution.

Any file type can be uploaded to DigitalCommons, but to assure long-term operability, access, and retrievability, text-based files and files using non-proprietary formats are the most common.

Levels of Access/Distribution

Two levels of electronic distribution of work are currently available through DigitalCommons@CalPoly. The level of access is determined by the copyright owner of the work: publishers as a condition of deposit of previously published work in an IR, or by a student, in concert with his/her advisor.

Open Access (Worldwide) Distribution

The first option is to make the information freely available worldwide. This means the descriptive data (author, title, abstract and basic information about the submission) will be publically visible and a link will be available to download the file.

Restricted Access (Complete restriction to file access)

The second option is to embargo (restrict) access to the work based on conditions set by the copyright owner (publisher or student). During ingest of the item, one can indicate the initial embargo time period.

Items holding the status of "Restricted Access," will only have their bibliographic information visible in the DigitalCommons@CalPoly. "Restricted Access" status is designed to secure the full work, disallowing access to everyone including the Cal Poly community. An item under "Restricted Access" status will be placed in a secure "holding pen", and the file will not be made visible or available. Only the descriptive data (author, title, abstract and basic information about the submission) will be publically visible.

After the embargo time period lapses, the document will be moved into open access (worldwide) distribution.

Policies

The policies can be accessed from the DigitalCommons@CalPoly website at: http://digitalcommons.calpoly.edu/faq.html.

Standards / Best Practices

- Dublin Core Metadata Standard
http://dublincore.org/

- ETD-MS (Metadata standard for Master's Theses)
http://www.ndltd.org/standards/metadata/etd-ms-v1.00-rev2.html

- OAI-PMH
http://www.openarchives.org/
The DigitalCommons@CalPoly is a registered data contributor to large–scale digital library projects including OAIster.org, Open Archives Initiative, and Scientific Commons.

- Copyright Clearance Guidance
http://www.sherpa.ac.uk/guidance/submission.html

- Imaging Best Practices

  Benchmark for Faithful Digital Reproductions of Monographs and Serials, Version 1, December 2002, Digital Library Federation.

  CDL Guidelines for Digital Images, Version 2.0, November 2005, CDL GDI, California Digital Library.

  Digital Imaging Workflow Procedures, 2007, Special Collections, Kennedy Library.

  Digital Projects Guidelines, Version 3.5, July 2007, Arizona State Library, Archives and Public Records.

  Minimum Standards for Faithful Digital Reproductions, 2007, Records Management Department, Arizona State Library, Archives and Public Records.

  Moving Theory into Practice: Digital Imaging for Libraries and Archives, by Anne R. Kenney and Oya Y. Rieger, Research Libraries Group (Mountain View, CA: 2000).

  The NINCH Guide to Good Practice in the Digital Representation & Management of Cultural Heritage Materials, November 2002, National Initiative for a Networked Cultural Heritage.

  Technical Guidelines for Digitizing Archival Materials for Electronic Access: Creation of Production Master Files – Raster Images, June 2004, US National Archives and Records Administration (NARA).

  Western States Digital Imaging Best Practices, Version 1.0, January 2003, Western States Digital Standards Group.

Challenges

In implementing the DigitalCommons@CalPoly, the following challenges were observed:

- Staff time to conduct bibliographic searches and interlibrary loan requests is greater than originally anticipated because contributors seldom have copies of their work
- Some academic publishers fail to respond quickly or refuse to grant permission for IR ingest of published material, delaying and/or limiting content ingestion.
- An average of five contacts by repository staff is necessary before citations are supplied by faculty, with a greater number of contacts for actual provision of articles and other content.
- Some colleges and academic departments that have launched their own faculty profiles consider DigitalCommons personalized pages as duplicative.

- Open access benefits of DigitalCommons are sometimes perceived as redundant by faculty who practice other forms of information dissemination in professional organizations
- Adding undergraduate work to DigitalCommons was challenging because there were few universal campus standards governing senior projects.
- Duplication of software functionality between repository module and profile module causes confusion for contributors, increasing staff time to manage profiles.
- Scalability of the project is a concern, given the length of time necessary to locate, secure rights, and ingest content and the number of unserved faculty and students.

Opportunities

During the implementation of the DigitalCommons@CalPoly, the following advantages were noted:
- Response to contributors is viral rather than organizational. As we continue to populate the IR and its use rates continue to climb, new contributors are contacting DigitalCommons first and asking to participate, based on "word of mouth"
- DigitalCommons is solving chronic information management issues for individual faculty members and key academic units on campus. The IR reduces the need for digital asset management at the unit level across campus.
- Centralized and convenient access to full-text research at Cal Poly is seen as a vehicle for improving the quality of graduate and undergraduate student work.
- DigitalCommons is ADA-compliant, providing campus units with a quick and convenient way of making their scholarly content universally accessible.
- Use of DigitalCommons extends beyond traditional scholarly uses. This includes Advancement, Public Affairs, Alumni and local campus journals and conference events.
- Cal Poly has joined its peers in higher education by contributing to the emerging model of distributed, interoperable repositories that provide the foundation for the new model of scholarly publishing.

References

Bailey, C., Jr. (2010). *Institutional Repository Bibliography*. Available from http://digital-scholarship.org/irb/toc.htm

Barton, M. R. and Waters, M. M. (2005). Creating an Institutional Repository: Leaders' Workbook. *Learning About Digital Institutional Repositories*, MIT Libraries. Available from http://www.dspace.org/implement/leadirs.pdf2004-2005.

Crow, R. (2002). The Case for Institutional Repositories: A SPARC Position Paper. *The Scholarly Publishing and Academic Resources Coalition (SPARC)*. Retrieved from: http://www.arl.org/sparc/bm~doc/ir_final_release_102.pdf.

Crow, R. (2002). *SPARC Institutional Repository Checklist & Resource Guide*. Washington, DC: Scholarly Publishing & Academic Resources Coalition. Available from http://www.arl.org/sparc/bm~doc/IR_Guide_&_Checklist_v1.pdf

DeRidder, J. (2004). *Choosing Software for an Institutional Repository.* Available from http://diglib.lib.utk.edu/dlc/ir_software.pdf

Foster, N. F. and Gibbons, S. (2005). Understanding Faculty to Improve Content Recruitment for Institutional Repositories. *DLib Magazine, 11*(1)*.* Available from http://www.dlib.org/dlib/january05/foster/01foster.html

Griscom, R., Hassen, M., Steiner, M.D. & Kerbel, S. (2006). *Content Recruitment and Development: A Proactive Approach to Building and Institutional Repository.* Retrieved from http://repository.upenn.edu/library_papers/10/

JISC Infonet. (2010). *Introduction to Digital Repositories.* Retrieved from http://www.jiscinfonet.ac.uk/infokits/repositories/

Johnson, R. K. (2002). Institutional Repositories: Partnering with Faculty to Enhance Scholarly Communication. *DLib Magazine, 8* (11). Available from http://dlib.org/dlib/november02/johnson/11johnson.html

Kim, Jihyun. (2010). Faculty Self-Archiving: Motivations and Barriers*. Journal of the American Society for Information, Science and Technology, 61*(9), 1909-1922.

Lessons Learned. (2010). Retrieved from the Duraspace Wiki: https://wiki.duraspace.org/display/DSPACE//LessonsLearned

Lippincott, J. K. (2006). *Institutional Strategies and Policies for Electronic Theses and Dissertations.* Available from http://net.educause.edu/ir/library/pdf/ERB0613.pdf

Mackie, M. (2004). Filling Institutional Repositories: Practical strategies from the DAEDALUS Project. *Ariadne, 39.* Available from http://www.ariadne.ac.uk/issue39/mackie/

Open Society Institute: A Guide to Institutional Repository Software. (2004). Available from http://www.soros.org/openaccess/pdf/OSI_Guide_to_IR_Software_v3.pdf.

Ramirez, M.L. and Miller, M. D. (Forthcoming). Approaches to Marketing an Institutional Repository to Campus. In Bluh, P. and Hepfer, C. (Eds.), *The Institutional Repository: Benefits and Challenges,* American Library Association.

Sonoma State University
Institutional Repository Case Study/ DRWG
October 2010
Submitted by Karen Brodsky, Librarian

Introduction

SSU ScholarWorks is Sonoma State University's digital archive for preserving and accessing the scholarly, creative, and administrative records of the University. In 2002, the Library began exploring the possibility of a repository system as a complement to the traditional university archive. Software and storage issues were explored as well as the costs of local maintenance in the absence of a system-wide centralized project. A project team was named and dSpace software was selected. Test documents were uploaded in 2004. In the fall of 2007, SSU ScholarWorks joined the CSU system-wide initiative and transferred software and storage responsibilities to the CSU server. From the onset, the Library has not been in a position to upload items in quantity or on a regular schedule because team members maintain numerous other areas of responsibilities within the library.

I. OBJECTIVES
Institutional Goals/Mission:
To create and establish an electronic archive system that:
- Captures and communicates the intellectual and administrative output of SSU
- Provides stable long-term storage and preservation of digitized and born-digital output
- Makes the University's record of scholarly information and accomplishments publicly accessible, enhancing SSU's institutional reputation

Initial Expectations:
- Early repository planning embraced the self-archiving service model. Now, as articulated throughout the literature, the full-service model is the preferred approach for many institutions, including SSU.
- There was an initial expectation that E-Theses would be the priority. Although theses would require significant staff time, high use was anticipated. Ultimately, faculty concerns about E-Theses and the lack of a campus mandate have dampened student participation. Although theses are still sought, repository efforts now focus on uploading faculty publications.

II. PROCESSES
Policies:
- Policies are evolving based on best practices throughout the IR community, including those available through SPARC.
- Policies are reviewed, as needed, and made publically available through FAQs on the SSU ScholarWorks home page.

Materials Collected include, but are not limited to:
- Pre-prints and post prints of articles
- Published journal articles [with permission confirmed]
- Teaching materials & learning objects
- Seminars
- Conference papers & presentations
- Electronic theses & dissertations

- Photographs & other visual media
- Sound & video recordings
- Web pages & born-digital items
- Datasets from research projects
- Out of print books & other endangered content
- Gray literature
- Governance & administrative records
- Award winning graduate and undergraduate research
- Serials

Materials Not Collected include, but are not limited to:
- Undergraduate student projects
- Working papers
- Drafts or otherwise incomplete works

Other Materials: As we continue to upload items, there are occasionally materials that have not been considered and that will be considered on a case by case basis.

Object Management/Upload:
- University Library is responsible for uploading and managing all materials, except the Academic Senate minutes.
- Faculty authors supply files for submission or a citation list via a CV or a link to a personal website. Administrative records are retrieved by the Library.
- Faculty authors can use an online consent form. Informal permission via email is also accepted and becoming a norm. Masters students are required to provide documentation that includes the signature of the committee chair before a thesis can be uploaded.
- A wide variety of formats are accepted and, if needed, the Library will convert items appropriately to improve accessibility.
- When possible, we maintain 508 accessibility compliance.
- Standard metadata schema is used, including LCSH, however, non-conforming terms suggested by authors are also included.
- Dublin Core Metadata standards are adhered to and additional internal standards are created when needed.
- Copyright permission is sought by the librarians and laws are fully adhered to.

III. RESOURCES
Personnel
- Two librarians are responsible for outreach, marketing, uploading, permissions, and quality control. Librarians allocate approximately four hours per week to SSU ScholarWorks.
- The Library's Director of Technology and the Web Manager make up the technical team and are responsible for communications with the Chancellor's Office, troubleshooting software issues, website design, etc. Time needed is intermittent.

Budget
- A specific allocation of campus funds does not exist, thus all expenses must come from the Library's operating budget.

Campus Support
- While there has been no financial support from the campus, there has been enthusiastic philosophical support. The Provost and Deans recognize the importance of an institutional repository and have submitted items, some are copyright pending.

Access
- Currently, all items are available through open access. The librarians have decided that, for now, items requiring password protection, often due to embargo terms, will not be uploaded until access is open.

Technology
- Further customization of appearance is desired and will be tackled in SSU ScholarWorks phase 2.0 (early 2011). Continued work with CO to address dSpace problems such as perpetual timeouts, etc.

Promotion
Expanded marketing efforts began in 2009/2010 academic year. Targets are determined each semester based on probability for population, political need, opportunity to upload new formats, etc. Some marketing approaches have included:
- Targeting faculty – those who are using the library regularly or who are close colleagues with Library faculty and faculty proponents of open access.
- Targeting Campus meetings – all meetings have generated interest in the inclusion of materials (many with copyright pending)
    - Biology
    - English
    - Academic Affairs
    - Graduate Studies meeting
    - Senate Executive Committee
    - Meetings with individual faculty members
- Email to all department chairs, faculty via Deans
- Presentation at school meetings fall 2009
- Handouts to all faculty via campus mail
- Participation in SSU Faculty Research Expo
- Prominent placement on Library homepage
- Letters from the Dean to all graduate students enrolled in thesis completion class
- Articles in campus e-newsletter
- Word of mouth

Barriers
- Without a campus mandate, utilization is not a campus priority
- Lack of understanding on part of some faculty
- Absence of self-archiving results in heavy workload for librarians
- End-user customization cannot, in most cases, be performed in-house
- Limited financial and personnel resources
- SSU SW Team members have many other responsibilities
- Publisher response time and/or lack of permission/clarity

- E-Theses paperwork thwarts student participation

Strengths
- Supportive faculty
- Library support
- Long term cost benefit of open access model
- Advancement of public knowledge
- Committed team