

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

Academe first took notice of institutional repositories (IRs) 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

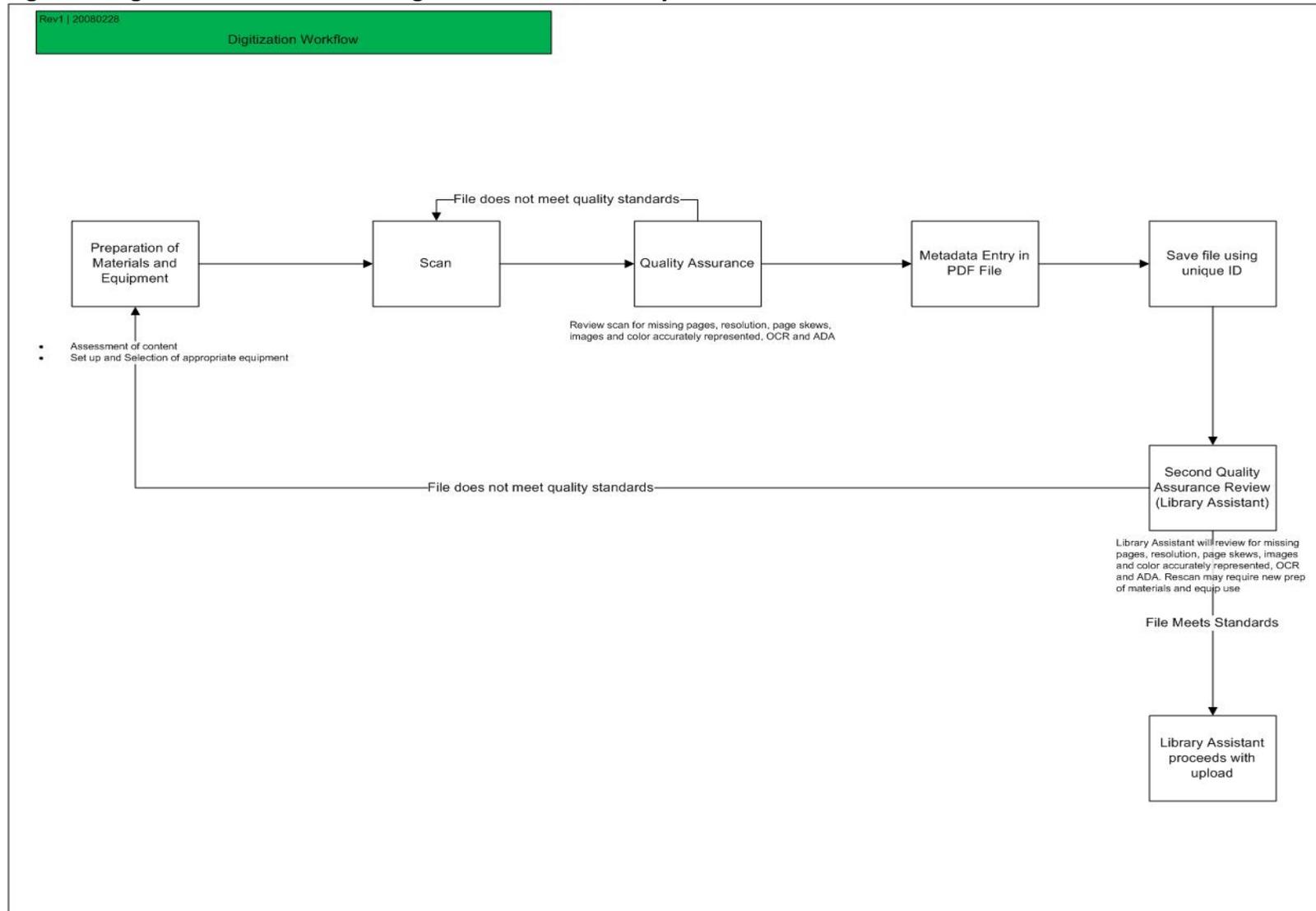


Figure 2: DigitalCommons@CalPoly Workflow

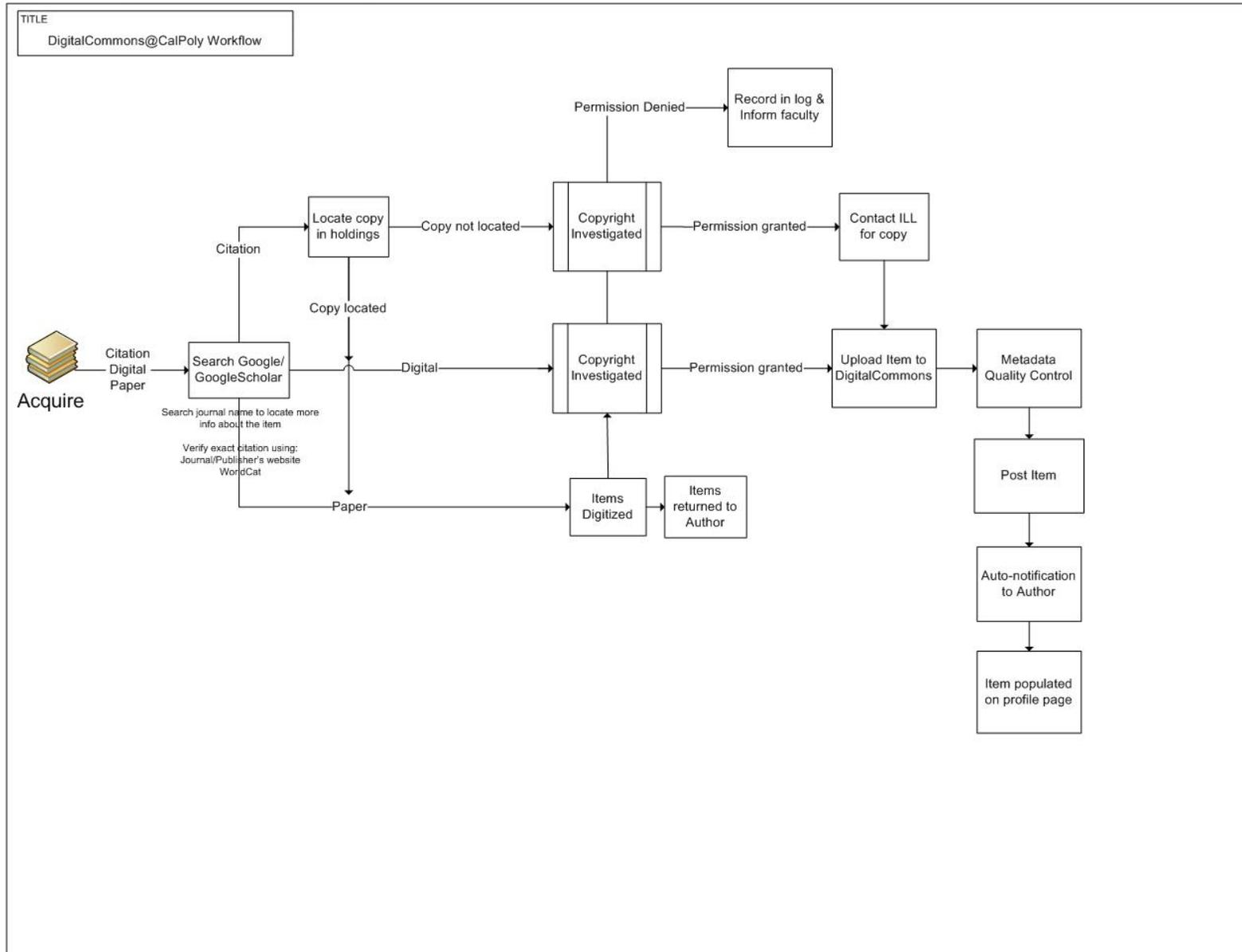
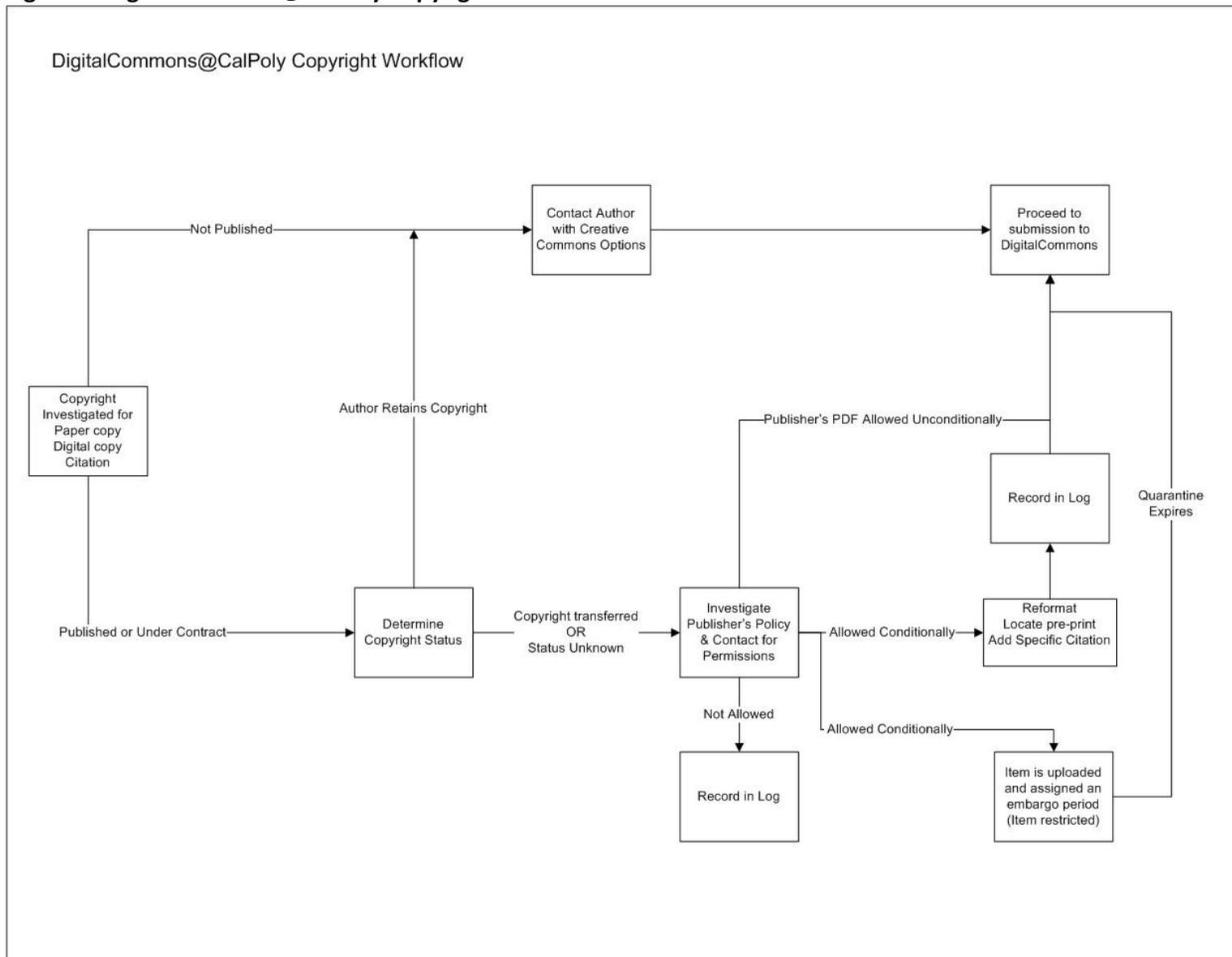


Figure 3: DigitalCommons@CalPoly Copyright Workflow



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:

Abstracts	Master's theses
Alumni publications	Peer-reviewed journal articles
Annual reports	Poetry
Architectural plans	Posters
Campus periodicals	PowerPoint presentations
Campus photographs	Press releases
Conference proceedings	Research from campus institutes 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.