

This material originally appeared in the January/February 2009 issue of D-Lib Magazine (<http://www.dlib.org>).

Institutional Repository on a Shoestring

George Wrenn, Cataloging Librarian
Humboldt State University
<glw11@humboldt.edu>

Carolyn J. Mueller, Special Projects Librarian
Humboldt State University
<cjm1@humboldt.edu>

Jeremy Shellhase, Systems Librarian
Humboldt State University
<Jeremy.Shellhase@humboldt.edu>

Introduction

Humboldt State University (HSU), with 7,800 students (fall 2008), is one of the smaller campuses in the 23-member California State University (CSU) system. Our institutional repository, Humboldt Digital Scholar¹ (HDS), originated as a pilot project during the Library's August 2004 planning meeting and became a permanent Library service in April 2006. The repository functions "on a shoestring," unfunded and reliant on contributions of time from librarians and library staff for its ongoing maintenance and development.

In this article, the authors, three members of the HDS Steering Committee, describe the process of setting up and managing a digital repository: hardware and software selection; customizations; gaining campus support; developing collections; accepting submissions; and planning for the future, including participation in a system-wide effort to create a shared repository for the CSU.

Overview

Our institutional repository, Humboldt Digital Scholar (HDS), originated as an unfunded pilot project at the Library's August 2004 planning meeting. A volunteer workgroup of seven (the Systems Librarian, a Systems staff member, the Chair of Access Services, a reference librarian, the Special Collections Librarian, the Special Projects Librarian and, in 2005, the newly-hired Cataloging Librarian) met weekly over the next 1.5 years to plan and develop the repository. The workgroup assessed the pilot in April 2006 and deemed it a success. The Library Dean agreed to

make HDS a permanent library service and offered one-time funding for a server that would provide storage for continued growth.

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.

Selecting and Customizing Hardware and Software

For the pilot, the Systems Department made available a retired Sun Microsystems Enterprise 450 server. Without a software budget, we limited our evaluation to free, open-source software and found M.I.T.'s DSpace software, released in 2002, to be an attractive, feature-rich product. Judging from their listserv² traffic, it had an active user base and support from open-source developers in the archive community. M.I.T. expressed a commitment to the continued development of the software.

Systems staff found DSpace relatively easy to implement and configure. After some experimentation with user-contributed modules, we decided to implement only those customizations available through the latest standard version of the DSpace code. Limiting customization to graphics, CSS style changes, minor changes to java server page (JSP) markup and standard configuration parameters allowed Systems staff to maintain a localized version of DSpace without the cost of more significant development efforts. The technical effort required to support an institutional repository (IR) server, with these limited parameters of customization,

was relatively easy to integrate into the workflow of the Library's four-person Systems Department. Configuring the DSpace application and Handle System® server³ proved to be short-term challenges without major implications for overall Systems workflow.

The Systems Librarian limits customizations of the standard DSpace interface to elements easily maintained through software upgrades. We have added a more colorful homepage banner, an introductory welcoming message describing the benefits of HDS, and "Submit your Scholarship" menu items linking to instructional and FAQ pages. Our instructions and FAQ draw heavily on DSpace's excellent documentation and on review of other IR sites.

Gaining Campus Support

Early in the pilot phase we looked into the possibility of having the repository hosted by campus Information Technology Services. The Systems Librarian briefed the campus Chief Information Officer (CIO) on the project and its technical architecture, which was determined to be compatible with long-term plans for campus computing. The CIO gave provisional assent to the concept of campus hosting. Details were left for later discussion; ultimately, further discussion became unnecessary when the CSU Council of Library Directors, then chaired by the Library Dean at Humboldt State, agreed in 2006 to initiate a system-wide, centrally managed repository: CSU ScholarWorks.⁴

To garner University administrative support, the Dean of the Library and the Systems Librarian met with the University Executive Committee in January 2005 to present an overview of repositories and their benefits. The response was positive, with the understanding that the pilot would be supported completely by the Library with existing resources. Later in the spring, presentations were made to graduate program coordinators, deans, and directors to introduce the repository to faculty. To further encourage faculty participation and heighten awareness of the open-access alternative, the Library's representative to HSU's Academic Senate submitted a resolution, approved in August 2005, recommending that faculty make a reasonable effort to acquire or retain the rights to place their publications in an open-access archive (Humboldt State University Academic Senate, 2005, August 30). We also touted the repository as a way to disseminate scholarship more widely, and to this end registered HDS with OAster⁵ (a union

catalog of digital repositories), OpenDOAR⁶ (a directory of open access repositories), and ROAR⁷ (a registry of open access repositories).

Developing Collections

We decided that HDS should be a repository for scholarship produced by HSU faculty, students, and staff. As with many academic repositories, a major focus is the stream of theses and projects produced by our graduate students (we require that all student submissions be faculty-sponsored). We hoped that HSU departments, and the centers and institutes sponsored by HSU, would welcome an archive in which to store work products. We considered allowing deposit of administrative records and course materials but concluded that these could become high-volume, labor-intensive commitments difficult to sustain without additional staff.

During the pilot and the year following, we focused our efforts on bringing about an institutional shift from paper to electronic thesis submission. As we developed the policies and procedures governing electronic thesis submission, we met frequently with the Graduate Council and worked closely with staff in the Office for Research & Graduate Studies. The Graduate Council established a policy permitting voluntary deposit of theses in Humboldt Digital Scholar and revised the *Application for Graduation* form to allow the student to substitute HDS submission for the Library's circulating paper copy (we continue to receive a paper copy for our non-circulating archives).

With thesis-submission routines in place and working well, we turned our attention to faculty and staff content. As many academic repository managers have discovered, without an institutional requirement, faculty participation can be a challenge. The February 2007 *Census of Institutional Repositories in the United States* reports that "operating repositories have had limited success in recruiting voluntary deposit of content" (Markey et al., 2007, p. ix). We review faculty, department, and non-academic unit web pages to identify potential participants. We sign up for alert services in aggregator databases to receive notice of new faculty publications. The Dean of the Library contacts faculty whose sabbatical reports describe published research or papers presented. We offer to verify publisher policies for faculty and staff via the SHERPA/RoMEO Publisher copyright policies & self-archiving site⁸ and publisher sites,

and contact publishers to request policies. Bibliographers are asked to invite faculty in their subject areas to submit scholarship. Advertisement in the *HSU Weekly University Notices* email has generated interest from several prolific scholars. We will continue to make a concerted effort to seek faculty and staff scholarship. We have had the most success with personal solicitations that include an offer to assist with the submission process. Outreach to staff has resulted in contributions from the Center for Indian Community Development and the University's radio station, KHSU. An enthusiastic emeritus faculty member recently volunteered to lend his name to our recruiting efforts.

For the most part, communities and collections in HDS reflect HSU's divisional and departmental hierarchy. When a faculty member joins HDS, we provide the option of joining under the faculty member's department or under an umbrella faculty scholarship community. Originally created to house cross-disciplinary scholarship, the latter has become a useful access point for all faculty publications in the repository. When a faculty member creates a departmental collection, we use the DSpace mapping feature to reproduce that collection within the faculty scholarship community. We also will establish communities housing subject-specific content drawn from other HDS collections, e.g., the Humboldt Bay Community,⁹ which brings together all scholarship on the local ecosystem. Using DSpace's alert service, which sends email "detailing new items that have become available" in user-selected collections, we are able to identify candidates for that community.

The Steering Committee saw the value in moving important campus lecture series onto HDS and in 2007 selected two high-profile series featuring visiting speakers. Our first video conversion was a 2005 lecture by Michael Pollan, which has proven to be the most popular item in the repository (Pollan, 2005). After obtaining permissions, we experimented with reproduction and compression in various formats (.mp3, .mp4, .mov), testing playback on Library computers of varying vintages. We eventually settled on .mp4 as the least proprietary of the compressed formats. An archived video file can be hundreds of megabytes, so we include minimum hardware specifications for playback on our lecture community pages.

Because the CSU's 2005 *Accessible Technology Initiative*¹⁰ requires that instructional materials¹¹ be 508-compliant by fall 2008 and that all CSU websites¹² be compliant by May 15, 2012, we were eager to attempt transcription and captioning of video. Three Steering Committee members spent more than twenty hours transcribing and captioning the Pollan video (accomplished with *Express Scribe Transcription Player* and *Camtasia Studio*), a result that suggested (rather forcefully) that we would not be able to make a habit of captioning the videos we add to HDS without additional staff support. We continue to discuss the role that HDS should play in video archiving and online video delivery, and watch with interest as the campus develops plans to deliver streaming video to the HSU community.

Accepting Submissions

The submission process for faculty¹³ and graduate students¹⁴ is documented on the HDS website. First-time users can contact the Steering Committee through the web site (graduate students are registered by catalogers after notification by the Graduate Office). When a faculty member contacts HDS, we register their email address and create for them a collection over which they have administrative control and the power to submit items, customize the collection home page with logos and html text, and delete items. We offer phone and email assistance with collection set-up and submission, which most faculty members accept. Generally, those who self-submit comment that the process is easier than expected. The last step in any submission is acceptance of the HDS license. If we submit on someone's behalf, we email them the license and retain a copy of the acceptance for our files.

With an eye to encouraging submission, we remain flexible and open to experiment where file formats are concerned. We request that theses and faculty articles be submitted in PDF format but will convert documents to PDF for submitters when necessary. Although we have done some scanning to digitize paper documents, without adequate staff time and a dedicated scanner we are able to offer this service only on an occasional basis. We have converted web pages to PDF, CDs to .mp3, and DVDs to .mp4, and we recently agreed to accept GIS data files. The Systems Librarian consults a list of the forty-six formats supported or known by DSpace¹⁵ when file format questions arise. Technical assistance from a Systems staff member conversant with file

types, close-captioning, and conversion software has been crucial to the success of the format conversion experiments we have done.

Two Cataloging staff members (one librarian and one library assistant) manage thesis submissions. They receive applications from the Graduate Studies Office, register submitter email addresses (authorizing submission to the Master's thesis collection) and send submission instructions. The student describes the thesis in metadata fields, uploads the PDF file, and agrees to the terms of the HDS license. The thesis remains publicly inaccessible until approved by Cataloging staff. Cataloging staff receive automated email notification of each submission. Thesis adviser name is authority controlled, capitalization is made to conform to AACR2R practice, a thesis note is added, and text is reviewed for problem characters and typos. Subject keywords assigned by students are not revised, eliminating the need for time-consuming cataloger subject assignment. In most cases, the review process can be completed in a matter of minutes, after which the student is notified that the thesis is online. The "handle" URL of the electronic thesis is added to the local catalog record for the hard copy of the thesis.

Most students complete the process without assistance. The percentage of graduate students electing to provide an electronic thesis has increased steadily: in fall 2007, 84 percent of the theses submitted to the Graduate Office were submitted to HDS. Electronic theses have proven popular (for example, for the period Jan 2, 2008 to Nov 24, 2008, five theses received between 401 and 500 item views, twelve received between 301 and 400 item views, and forty-nine received between 201 and 300 views). Some faculty members link to theses in HDS from their web pages, and faculty, students, and staff use the repository to locate theses produced by a given program. We have received requests to add earlier theses and will place a brief notice in the online alumni newsletter asking former graduate students with theses in digital format to submit to HDS. Although we would like to convert older theses to digital format and offer on-demand digitizing for faculty, we currently only accept scholarship already digitized.

ScholarWorks and the Future

As mentioned above, a promising development for long-term hosting is CSU ScholarWorks, the California State University system-wide DSpace repository, initiated in 2006 at the urging of the

former HSU Library Dean. Although not designed expressly for it, the DSpace software is well-suited to a multi-institutional implementation with individualized member repositories, and our experience with HDS enables us to make significant contributions to this endeavor. CSU ScholarWorks offers the prospect of long-term stability and growth. It will have meta-search capabilities to provide researchers with access to scholarship produced across the CSU and will offer robust, centralized support, freeing campus-level systems staff from maintenance responsibilities while allowing individual repositories to control their own submission routines. We plan to move all HDS operations to CSU ScholarWorks in the next six to twelve months but will continue simultaneous local hosting until we are satisfied that the migration is successful and that all needed flexibility in customization is available.

Humboldt Digital Scholar continues to be a shoestring operation, with content increasing slowly but steadily. In fall 2008, we had over 300 items archived and over a dozen faculty contributors. With the available staff time and resources, we will continue to be selective in the projects that we choose to pursue, but we have demonstrated that an institutional repository, even for a smaller institution with limited resources, can be a valuable addition to the library's services to the university and beyond.

Notes

¹ Humboldt Digital Scholar. <<http://dscholar.humboldt.edu:8080/dspace/>>.

² DSpace Mailing Lists. <<http://www.dspace.org/feedback/mailling.html>>.

³ Handle System® homepage <<http://www.handle.net/>>.

⁴ California State University ScholarWorks.

<<http://www.calstate.edu/UIAS/ScholarWorks.html>> .

⁵ OAIster. <<http://www.oaister.org/>>.

⁶ OpenDOAR. <http://www.opendoar.org/>>.

⁷ Registry of Open Access Repositories (ROAR). <<http://roar.eprints.org/>>.

⁸ SHERPA/RoMEO. <<http://www.sherpa.ac.uk/romeo.php>>.

⁹ Humboldt Bay Community home page.

<<http://dscholar.humboldt.edu:8080/dspace/handle/2148/209>>.

¹⁰ California State University Accessible Technology Initiative.

<<http://www.calstate.edu/Accessibility/index.shtml>>.

¹¹ California State University Accessible Technology Initiative, Instructional Materials.

<<http://www.calstate.edu/accessibility/instructionalmaterials/index.shtml>>.

¹² California State University Accessible Technology Initiative, Web Accessibility Deliverables and Timeline.

<http://www.calstate.edu/accessibility/webaccessibility/deliverable_timeline.shtml>.

¹³ Humboldt Digital Scholar Submission Guidelines.

<http://dscholar.humboldt.edu:8080/dspace/help/HDS_Submission_Guidelines.htm>.

¹⁴ Electronic Theses. <http://dscholar.humboldt.edu:8080/dspace/help/Electronic_theses.htm>.

¹⁵ DSpace Format Support. <<http://libraries.mit.edu/dspace-mit/build/policies/format.html>>.

References

Humboldt State University Academic Senate (2005, August 30). *Resolution on Scholarly Journal Publication and Faculty Rights For Open-Access Archiving*. Retrieved from the Humboldt State University website: <<http://www.humboldt.edu/~acadsen/Resolution02-05-06-EXFINAL.doc>>.

Markey, K., Reih, S. Y., St. Jean, B., Kim, J., & Yakel, E. (2007). *Census of Institutional Repositories in the United States: MIRACLE Project Research Findings*. Retrieved from the Council on Library and Information Resources website: <<http://www.clir.org/pubs/reports/pub140/contents.html>>.

Pollan, M. (2005). *The High Cost of Cheap Food*. Retrieved from the Humboldt Digital Scholar website: <<http://dscholar.humboldt.edu:8080/dspace/handle/2148/250>>.