

Evolution of a Digital Repository: One Institution's Experience

Terry M. Owen

University of Maryland

Author Note

Terry M. Owen, University of Maryland Libraries, University of Maryland

Correspondence concerning this article should be addressed to Terry M. Owen,

University of Maryland, McKeldin Library, College Park, MD 20742. E-mail: towen@umd.edu

Abstract

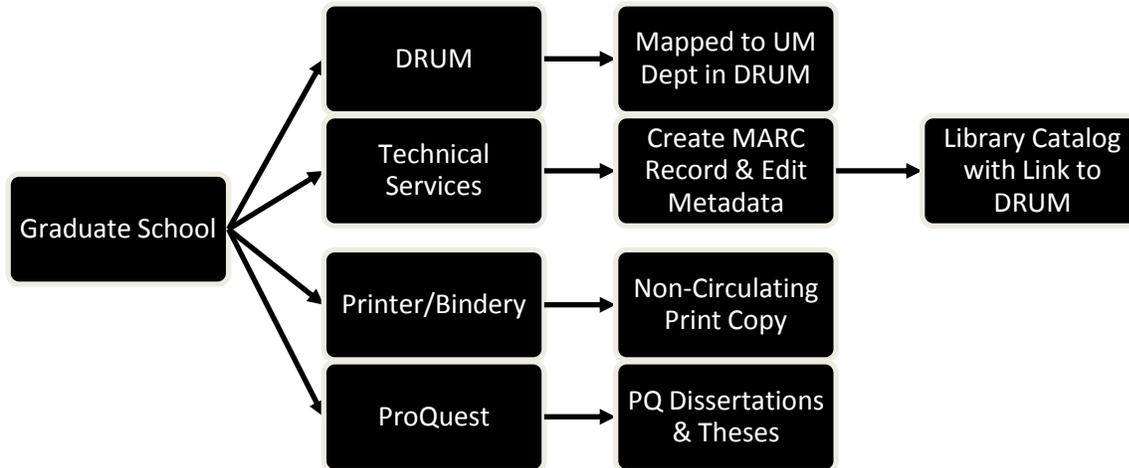
In this article, the development of a digital repository is examined, specifically how the focus on acquiring content for the repository has transitioned from faculty-published research to include the gray literature produced by the research centers on campus, including unpublished technical reports and undergraduate research from honors programs. This material has the benefit of fewer copyright restrictions, making acquisition much less problematic. Inclusion in the repository increases the creditability of the research center and provides wider distribution of this often under-recognized research.

Keywords: Digital repositories; institutional repositories; gray literature; outreach strategies; University of Maryland

The mission of the Digital Repository at the University of Maryland (DRUM), available at <http://drum.lib.umd.edu>, has always been to store, index, distribute, and preserve the research works of the University of Maryland (UMD) faculty. During the past three years, however, the mission of the digital repository has gradually expanded to include not only faculty-published research but student research, both graduate and undergraduate, and unpublished technical reports from many of the research centers and institutes located on campus. Broadening the coverage of DRUM has increased the number of documents in the repository and made previously undiscovered research available to a wider audience.

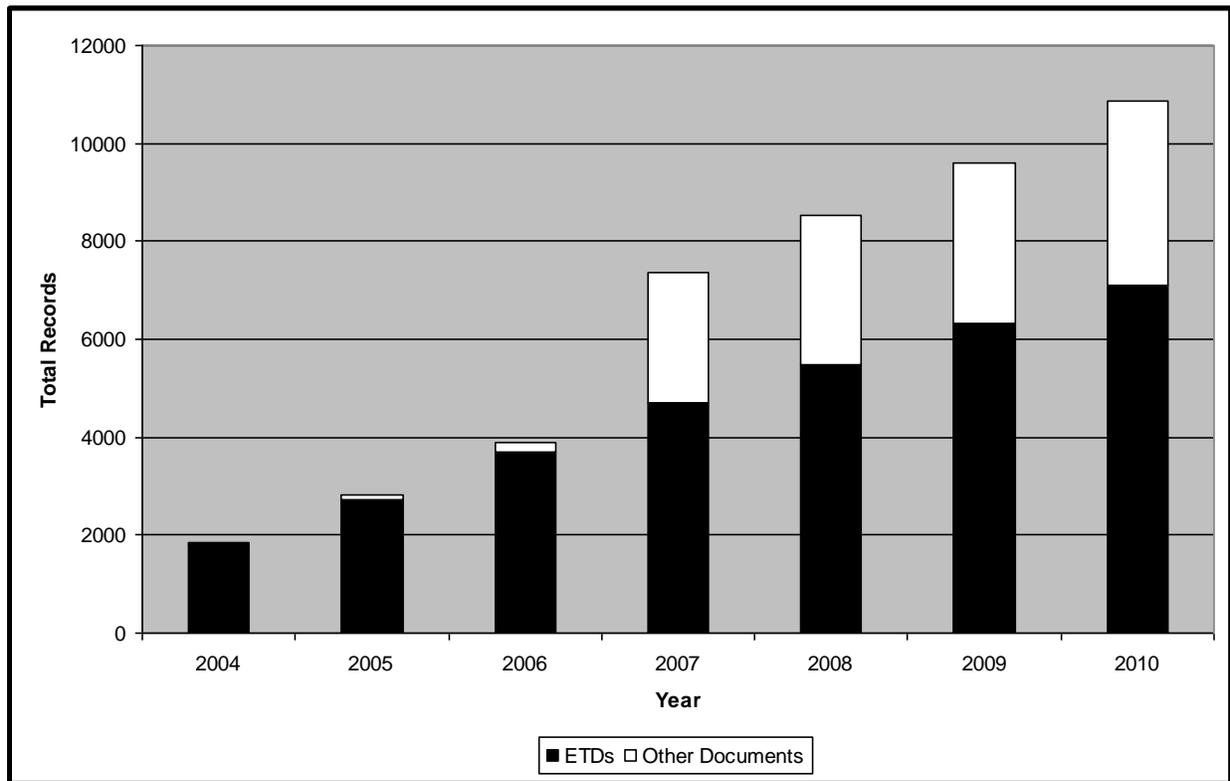
Phase I: ETDs

The University of Maryland Libraries launched DRUM in 2004, populating the repository primarily with electronic theses and dissertations (ETDs). Because electronic submission of theses and dissertations at UMD has been mandatory since 2003, inclusion of these documents seemed logical and easily accomplished. The ETD submission process is completely automated and batches are delivered at the end of each semester (spring, summer, and fall), adding 800 to 900 documents to the repository annually. As shown in the ETD workflow outlined in Figure 1, documents are simultaneously deposited in DRUM; transmitted to the library's technical services department for processing and inclusion in the library catalog; and sent to ProQuest to be made available in *ProQuest Dissertations and Theses*.

Figure 1. Workflow for Electronic Theses and Dissertations (ETDs).

ETDs account for approximately 58 percent, or more than 6,000 records, of the total documents in DRUM. But there has been a significant increase in other document types, specifically technical reports from various research centers on campus, over the past three years (see Fig. 2).

Figure 2. Population of Electronic Theses and Dissertations (ETDs) vs. Other Documents in DRUM.



The UMD Libraries hired a DRUM coordinator in 2006 to advance the development of the repository and promote services to the campus. As the ETD submission process was implemented two years previously, one of primary responsibilities of the coordinator was to encourage faculty members to self-archive their publications in order to make them more widely available via DRUM.

Initial Faculty Outreach

Initial DRUM marketing efforts to faculty at UMD included placement of articles in campus publications highlighting the benefits of increasing the visibility of research and providing workshops and seminars on the DRUM self-deposit process. Unfortunately, most of these were poorly attended. Though these first attempts were successful in promoting the

repository, the response from faculty was unenthusiastic and resulted in very few self-submitted faculty deposits in DRUM.

Faculty members at UMD often post their research on personal or departmental websites, despite the fact that this often constitutes a violation of copyright. In an attempt to capture this research for the repository, the DRUM coordinator sent e-mails to more than 400 faculty members with an offer to deposit the papers in DRUM and, if needed, verify or obtain publisher permissions on their behalf. Approximately 60 faculty members expressed an interest, but would participate only if it did not involve additional work on their part and they were assured that there were no copyright violations. These findings confirmed the results of earlier studies at the University of Rochester (Foster & Gibbons, 2005) and Cornell University (Davis & Connolly, 2007). However, the campaign proved to be highly successful, eventually adding more than 400 documents to the repository. Unfortunately, the DRUM staff greatly underestimated the time and effort required to obtain publisher permissions and deposit the research. To ensure that copyright was not violated, publisher policies were verified by checking the SHERPA/RoMEO Publisher Copyright Policies & Self-Archiving database (<http://www.sherpa.ac.uk/romeo/>). In instances where policies could not be located, publishers were contacted directly for permission to deposit in DRUM, prolonging the submission process. Even though the project was successful, this deposit model was too laborious to be sustained by a repository staff consisting of only 1.5 staff members. Because much of this formally published research was most likely available on the journal website or in another repository, such as PubMed Central, the decision was made to discontinue the project and instead concentrate on acquiring and making available the unique gray literature produced at the university. At UMD, gray literature is material that has

not been published, or if published, cannot be easily located. It takes many forms across multiple disciplines and includes technical reports, presentations, and theses and dissertations.

Phase II: Focus on Research Centers

With highly regarded and nationally ranked departments in engineering, physics, mathematics, business, and education, along with a close proximity to the Washington, DC, area, UMD is fortunate to be affiliated with almost 200 research centers, laboratories, and institutes, covering a wide variety of disciplines. Notable examples include the Institute for Systems Research, the University of Maryland Institute for Advanced Computer Studies and the Center for International and Security Studies at Maryland. While researching faculty websites for potential articles to deposit, it was noted that many of the research centers, especially those in computer science and related disciplines, had posted many of their publications on their sites. A majority of these documents were unpublished technical reports that ranged in number from fewer than 25 to more than 2,000 papers for an individual research center. A major advantage in soliciting these institutes to deposit their material into DRUM was the avoidance of copyright complications. In many cases, the institute is the copyright owner and is free to deposit in a repository or post to their website. In instances when the material had been formally published, the DRUM staff continues to obtain permission from the publisher.

Initial conversations with directors and faculty members associated with these affiliates revealed a great eagerness to participate in DRUM in order to share their research more widely. Owing to the highly competitive nature of obtaining research funds, DRUM provides an attractive mechanism to increase awareness and creditability of these programs and research data. As a result, the DRUM coordinator worked with each research center to develop workflows to suit individual needs. Even though the repository staff continues to monitor all

submissions, most of the responsibilities for depositing and approving documents have been assigned to individuals in each center.

Some of the research centers requested major changes to the DRUM search and/or display interface, which presented somewhat of a challenge. DSpace, developed by the Massachusetts Institute of Technology Libraries and Hewlett-Packard Labs in 2002 (Smith et al., 2003), is the open-source software currently used for the repository and, though significant in-house enhancements have been made to the software since the launch of the repository (e.g. embargoes for electronic theses and dissertations), access to the library's programmers is very limited, requiring the need to justify even the smallest software modification. Though most of the requests for changes have been minor, each required individual evaluation with a final determination being dependent on whether the change would benefit all DRUM users and not just a specific population. Examples of these requests included

- option to change the number of records that display per page
- improve sorting capabilities or add new sort fields (e.g. report number)
- alter display, organization, or format of items that appear on the screen
- display number of downloads for each document.

Fortunately, many of these requests were addressed through upgrades to the DSpace software. The modification of displaying the number of times a document had been downloaded or viewed proved to be very popular with researchers. In fact, this enhancement is often cited by junior faculty members at UMD to illustrate the impact of their research.

Overall, concentrating outreach efforts on research centers has proven to be very successful. As the burden of depositing and approving documents from the various centers has

been removed from the DRUM staff, it has allowed expansion of the repository into other areas, notably undergraduate research.

Phase III: Undergraduate Research

Initially, DRUM deposits were limited to faculty research only. It was thought that including undergraduate research would dilute the quality of the work contained in DRUM and weaken the importance of the repository. It has been our experience, however, that this actually is not the case. Within the past year, several professors have inquired as to the possibility of including research of their graduate or undergraduate students in DRUM. In each instance, the DRUM coordinator evaluates all requests based on scholarly merit for inclusion in the repository. Potential submissions must be substantive works of scholarship and deemed acceptable for publication. Research works submitted from undergraduate honors programs, specifically the Gemstone Program at UMD, were the first student papers to be included in DRUM. The Gemstone Program is a unique multidisciplinary 4-year research program for selected undergraduate honors students of all majors. Teams of students design, direct, and conduct significant research that is completed in their fourth year and presented as a thesis to experts in their field. As all graduate theses and dissertations from UMD are automatically included in DRUM, adding the Gemstone projects was a logical addition. Other student research soon followed, including graduate student projects from the Historic Preservation program in the School of Architecture, Planning and Preservation, undergraduate research from the Ronald E. McNair Post-Baccalaureate Achievement Program, and papers from the Archaeology in Annapolis project, affiliated with the anthropology department at the University.

Even though undergraduate and graduate research is now included in the repository, students are not authorized to self-deposit their documents. Self-deposit is restricted to faculty

members only and, in most cases, the repository staff deposits the works on behalf of students to ensure quality submissions.

Moving Forward

Today, most of the efforts in collecting research for DRUM will continue to focus on the gray literature produced by the wide variety of research centers on campus, along with evaluating relevant research from undergraduate and graduate students. This material is not readily available elsewhere, has few copyright complications, and research centers are typically motivated to deposit their works. However, much trial and error goes into building and maintaining a successful digital repository and each institution is unique. Various strategies to explore include

- working with the graduate school to ensure that all theses and dissertations are deposited in the repository;
- reporting usage statistics or number of downloads to faculty;
- addressing the specific needs of the various disciplines with tailored messages for individual departments and faculty members; Strategies that succeed in the sciences might not translate to the arts and humanities;
- developing different messages for administrators, deans or department chairs; Administrators are typically more interested in the added prestige and visibility of the institution or department whereas faculty members are usually more focused on his/her personal research; and
- searching faculty websites for potential research to deposit, keeping in mind the time required to process the documents.

Even though efforts are centered on obtaining more gray literature, faculty members are still encouraged to deposit their published research in DRUM. Coupled with the passage of the NIH Public Access Policy (<http://publicaccess.nih.gov/>) in 2008 and the defeat of an open-access resolution by the UMD Senate in 2009 (Hackman, 2009), discussions with faculty members have focused on issues related to open access, specifically authors' rights and publishing in open-access journals. Educating faculty members and the campus at large about open access has become a priority for the near future and will eventually result in more deposits in DRUM. Finally, in an effort to expand library services to the campus, the UMD Libraries plan to broaden coverage of DRUM to include presentations and proceedings of conferences held at UMD and also to explore the potential for establishing an open-access journal publishing program managed by the Libraries.

When deciding how best to increase deposits in a digital repository, each institution must first identify potential resources of appropriate material, including both published and unpublished research. Concentrating efforts on acquiring gray literature that is not readily available, however, will provide wide distribution for research that might have never been previously discovered and increase the population of the repository.

Received: January 19, 2011

Accepted: February 3, 2011

References

- Davis, P. M., and Connolly, M. J. L. (2007). Institutional repositories: Evaluating the reasons for non-use of Cornell University's installation of DSpace. *D-Lib Magazine*, 13 (3/4). doi:10.1045/march2007-davis
- Foster, N. F., & Gibbons, S. (2005). Understanding faculty to improve content recruitment for institutional repositories. *D-Lib Magazine*, 11 (1). doi:10.1045/january2005-foster
- Hackman, T. (2009) What's the opposite of a pyrrhic victory? Lessons learned from an open access defeat. *College and Research Libraries News*, 70 (9), 518-521, 538.
- Smith, M., Barton, M., Bass, M., Branschofsky, M., McCellan, G., Stuve, D., Tansley, R., and Walker, J. H. (2003). DSpace: An open source dynamic digital repository. *D-Lib Magazine*, 9 (1). doi:10.1045/january2003-smith