

Sarah L. Shreeves and Thomas H. Teper

Looking backwards

Asserting control over historic dissertations

Scholarly communication programs are nearly as diverse as the institutions that support them, and the individual components of these programs tend to be highly specific to the institution in question. Components may include provision of an institutional repository, support for journal publishing, education and outreach programs, researchers on author rights and open access, and the development of open access mandates.

Most of these efforts hold one aspect in common—an attempt to support efforts that assert individual or institutional control over new, locally produced content in order to increase access. Fewer programs attempt to assert control over their institutions' legacy of research and scholarship.

This is understandable, as a large portion of legacy scholarship remains the intellectual property of the publishers that disseminated the content. Reasserting control over this content remains difficult. However, two significant parts of the research legacy remain within reach: the grey literature (technical reports, working papers, bulletins, etc.) generated by our departments and research centers, and theses and dissertations authored by graduate students.

The University of Illinois at Urbana-Champaign (UIUC) Library, with the blessing of University Counsel, actively pursues the digitization of our rich collections of grey literature and provides open access to these in IDEALS, the UIUC digital repository for research and scholarship.¹ Beginning in 2006, the library digitized and made openly available nearly 7,000 items of grey literature from

its collections.

Theses and dissertations, the subject of this column, are a more complicated target as copyright rests with the author, according to UIUC policy. However, historical theses and dissertations remain important to individual faculty members and, especially, graduate students across the country and abroad. While the advent and widespread implementation of electronic thesis and dissertation (ETD) programs largely address the access issues for new content, historical theses and dissertations represent some of the least well-disseminated and accessible scholarship generated on academic campuses.

Knowing that, UIUC embarked on an ambitious program to digitize the entirety of its historic dissertation and theses collections with the goals of expanding access, promoting the campus' past research and scholarship, and, wherever possible, connecting with authors (alumni) to provide open access to their scholarship.

From the beginning, the library viewed this effort as less of a collection management or preservation issue and more as an effort to tackle broader scholarly communication and

Sarah L. Shreeves is the coordinator for IDEALS and co-coordinator for the Scholarly Commons, e-mail: sshreeve@illinois.edu, and Thomas H. Teper is associate university librarian for collections at the University of Illinois at Urbana Champaign, e-mail: tteper@illinois.edu

Contact Claire Stewart—series editor, head of digital collections and scholarly communication service at Northwestern University—with article ideas, e-mail: claire-stewart@northwestern.edu

© 2012 Sarah L. Shreeves and Thomas H. Teper

outreach issues. Paying ProQuest to simply digitize our historical dissertation collection and provide access to these items within their commercial platform would serve to make the dissertations accessible to similar campuses, but it would not serve to make them accessible to scholars at institutions without access to this product.

Furthermore, it would neither help us to reconnect with the thousands of scholars that spent significant portions of their life developing their academic credentials on this campus, nor would it provide open access to the entire collection of theses and dissertations. For these reasons, the library (through IDEALS) wanted to serve as the ultimate aggregation and access point, even if success—digitization of the entire backfile, increased use of locally produced scholarship, and reconnection with the authors—would require a multiyear process.

One goal of this digitization effort centers on unifying access to UIUC's theses and dissertation collection. Prior to the start of the digitization program, these resources were available through a multitude of locations, platforms, and formats that complicated search and discovery efforts by users and library personnel alike:

- Dissertations prior to the mid-1940s: Available in paper and managed by the Rare Books and Manuscript Library (RBML). These paper copies are the copy of record. In some limited cases, copies may be in a subject library.

- Dissertations produced between the mid-1940s and 1996: Available in the UMI/ProQuest-provided microfilm (copy of record) or in paper in RBML or, at times, in a subject library.

- Dissertations produced between 1997 and 2009: Available in the UMI/ProQuest-provided microfilm (copy of record), in paper in RBML or, at times, in a subject library, and as a scanned document in the ProQuest Digital Dissertations platform.

- Dissertations produced from 2010 to date: Available electronically either through IDEALS (copy of record) or in the ProQuest Digital Dissertations platform.

- Theses prior to 2010: Available in paper in RBML (copy of record) or, at times, in a subject library.

- Theses produced from 2010 to date: Available electronically through IDEALS (copy of record).

On the occasion that a circulating copy of a dissertation or thesis was identified as available in a subject library, the likelihood of that copy actually being on the shelf was often in doubt.²

The first step in digitizing these dissertations and theses focused on clarifying the rights issues. UIUC's institutional policy states that copyright of a thesis or dissertation rests solely with the author. For theses and dissertations submitted since the 2010 implementation of UIUC's ETD program, the university secures clear permissions from the authors to make these openly available via IDEALS.³ This permission is not present for dissertations and theses submitted in paper prior to 2010. However, the General Rules of UIUC gives the university the right, as a condition of deposit, to provide access to and disseminate these documents.⁴

When the library brought this issue to University Counsel in 2006, it was determined that the digitization of theses and dissertations constituted a low risk as long as the library limited direct access to the UIUC community and provided a means for copyright holders to request a dissertation or thesis be further restricted or withdrawn or, conversely, to be made available openly.

The decision was prompted in large part by the belief that providing expanded access to these items is a responsibility of the university that is supported by the General Rules, the University Counsel, the Graduate College, and the library.

For the majority of the dissertations there were two options: digitize microfilm holdings or paper copies. Paper-to-digital conversion remains the only option for Master's theses. While University Counsel expressed comfort with the library digitizing the paper copies, this comfort level dropped precipitously when we discussed digitizing the ProQuest-produced microfilm directly. Given this concern, the

library opted to work with ProQuest to digitize what we could from microfilm before moving on to paper-to-digital conversion processes. Previously, ProQuest approached the library with a report detailing the entire body of microfilmed dissertations from the campus and a proposal that UIUC pay to digitize the content for exclusive delivery through their platform.

Given the restrictions imposed on access, the library was not entirely receptive. The pricing that accompanied this 2005 offer appeared excessive, and the proposed model failed to address a growing awareness of our obligation to assert control over locally generated intellectual content. By locking the content down in a commercial platform, goals of asserting control and enhancing access could not be met. Yet, the publisher remained reluctant to permit local loading of the digitized content due to assignment of rights that occurred when the authors agreed to have their titles microfilmed.

This impasse persisted until 2009, when, after extensive negotiations, ProQuest agreed to allow the digitized dissertations to appear in IDEALS under restricted access and would allow UIUC to open access to those dissertations whose authors authorized such an action. Approximately 5,000 dissertations authored between 1989 and 1997 were digitized and made available in PDF form both within IDEALS (to the UIUC community) and in the ProQuest platform. While the PDFs are not indexed by search engines because of the access restrictions, the metadata—including extended abstracts—is, and IDEALS sees many requests for this content, particularly from overseas. Such requests are referred to the ProQuest platform and to interlibrary loan at their home institution.

While this pilot taught us much, it was not an unparalleled success. The process for providing open access to a dissertation proved challenging on both ends. Authors who wished to open access to their dissertation within IDEALS were required to provide a letter stating their authorization of this action to ProQuest who would then notify the library. This was an extremely burdensome process for authors, and stymied any progress on making

the titles accessible to individuals outside the bounds of our campus.

Success eventually came about through a bit of serendipity. During discussions about an unrelated acquisition, our ProQuest sales representative asked whether UIUC remained interested in digitizing additional dissertations. We remained reluctant, concerned about the burdensome nature of the permissions process and the pricing model. But, the ProQuest sales representative suggested that something might be possible. Within a couple weeks, the vendor agreed that UIUC could locally load the digitized dissertations, open access to dissertations upon receipt of correspondence from the copyright holder, and develop a process to notify the vendor of receipt of such permissions.

With this agreement in hand, negotiations could begin in earnest. Once a contract was signed, the digitization of some 18,000 microfilmed dissertations began. Within a couple months, ProQuest completed the work, made the dissertations available through the commercial platform, and delivered PDF copies, as well as the MARC records, to UIUC for ingest into IDEALS. IDEALS staff are currently processing these dissertations for the repository; this involves metadata wrangling, including extracting departmental and disciplinary information from the PDFs to add to the metadata, and ensuring that the dissertations appear in the appropriate collections within IDEALS. We will also add the IDEALS location to records in UIUC's online catalog. The dissertations will be ready for ingest in fall 2012.

Once we have readied the dissertations for ingest, we will work with departments and colleges on campus, as well as the Alumni Association, to contact authors in order to secure permissions to make the dissertations openly available. We are developing an online form that allows authors to indicate their ownership of a specific dissertation, and to indicate their willingness to make it openly available. The library will use this information to make the dissertation openly available and to notify ProQuest of its action. We foresee this to be an excellent opportunity to develop and

strengthen connections with departments and graduate program alumni.

While completing the digitization of dissertations that ProQuest had microfilmed is an enormous step forward, we still have a substantive collection of Master's theses as well as early dissertations that exist only in paper. This year, working with the Digital Collections and Content group, we began a pilot digitization project to understand the resources needed for a larger project. The challenges are considerable. In the case of Master's theses, there is a matter of scale. Historically, the number of Master's theses produced at UIUC has nearly equaled or slightly exceeded the number of dissertations produced, though this is no longer true.⁵ We also have no definitive record of all of the theses and dissertations produced on the campus and so must rely on reports from the online catalog and what appears on the shelf.

As for the early dissertations, many of these are typewritten on higher-quality paper, but a significant number of them are handwritten documents that resemble modern-day composition books, complete with lightweight board covers and embrittled whip-stitched paper. In addition to concern about the physical integrity of these items, the library faces the challenge of delivering meaningful access to handwritten manuscripts for which OCR does not appear to be a viable option. We are exploring the potential of crowdsourcing or using new research to automate such a process, but recognize that these may be difficult for a collection of items numbering in the hundreds. Nonetheless, UIUC continues pursuing a long-term objective of making digital versions of all of these dissertations available.

Our continued efforts to digitize UIUC's collection of theses and dissertations as a component of our scholarly communication program represents a significant departure from similar digitization efforts that focused solely on access or preservation. Of course those are benefits, but UIUC pursued this program with the long-term intent of developing coherent access to its locally produced scholarship, exerting better control over these assets, and strengthening relationships with authors, departments, research

centers, and colleges. Significantly, opening access to past scholarship is part of a wider scholarly communication conversation; making this investment opens the path to initiating and continuing conversations about current scholarship with those same stakeholders.

Notes

1. IDEALS: <http://www.ideals.illinois.edu>. For an example of a grey literature collection, see the Bulletin of the Illinois Agricultural Experiment Station at <https://www.ideals.illinois.edu/handle/2142/3536>.

2. Cherié L. Weible, "Where Have All the Dissertations Gone? Assessing the State of a Unique Collection's Shelf (Un)availability" *Collection Management* 30(1): 55-62.

3. UIUC Graduate College FAQ on ETDs <http://www.grad.illinois.edu/thesis-faqs>.

4. See Article III, Section 4: <http://www.illinois.edu/trustees/rules.cfm#sec34>.

5. UIUC currently produces approximately 1,300 dissertations and theses per year, of which 60 percent are dissertations and 40 percent are theses. *zz*



ARL

2013 Workshops

brought to you by the

Association of Research Libraries
Statistics & Assessment Program

Service Quality Evaluation Academy

March 11–15
New Orleans, LA

Advanced XML: Further Adventures with XSLT

March 25–27
Washington, DC

To learn more and register, visit

www.arl.org/stats/statsevents