Managing ETDs: The Good, the Bad, and the Ugly

Dan Tam Do
University of Vermont, dan-tam.do@uvm.edu

Laura Gewissler
University of Vermont, lgewissler@gmail.com

Follow this and additional works at: <https://docs.lib.purdue.edu/charleston>

Part of the [Cataloging and Metadata Commons](#), [Collection Development and Management Commons](#), [Scholarly Communication Commons](#), and the [Scholarly Publishing Commons](#)

An indexed, print copy of the Proceedings is also available for purchase at:

<http://www.thepress.purdue.edu/series/charleston>.

You may also be interested in the new series, Charleston Insights in Library, Archival, and Information Sciences. Find out more at: <http://www.thepress.purdue.edu/series/charleston-insights-library-archival-and-information-sciences>.

Dan Tam Do and Laura Gewissler, "Managing ETDs: The Good, the Bad, and the Ugly" (2017). *Proceedings of the Charleston Library Conference*.
<http://dx.doi.org/10.5703/1288284316661>

Managing ETDs: The Good, the Bad, and the Ugly

Dan Tam Do, University of Vermont

Laura Gewissler, University of Vermont

Abstract

Mandating contribution of theses and dissertations (TDs) to university archives and their electronic equivalents (ETDs) to an institutional repository (IR) is common practice. Optimizing workflows for archival print copies while managing electronic copies in an IR can be challenging given such factors as embargoes and the skill sets required to ensure theses and dissertations are accessible, discoverable, and ultimately safely stashed where they belong. As rational processes were gradually developed at the University of Vermont, pitfalls and breakthroughs presented themselves. This article relates our experience launching an ETD mandate, including campus outreach initiatives and improvements to the various related processes (document submission, harvesting, embargo removal). Our journey encompassed a range of experiences that we designated good, bad, or ugly, depending on workflow impact. We realize these are mere labels and that beauty is in the eye of the beholder, especially regarding embargoes.

Introduction

Electronic thesis and dissertation (ETD) management is a major 21st-century challenge for academic libraries. University of Vermont (UVM) Libraries joined the ranks of libraries using ProQuest Administrator to manage their graduate college theses and dissertations in 2014. A major impetus to provide access to ETDs grew out of the desire to populate a newly acquired institutional repository (IR). In 2013 UVM Libraries began a subscription to Digital Commons from bepress. Our goal was to provide access to the intellectual and scholarly output of the UVM community, particularly student scholarship.

ETDs were the perfect content to demonstrate the caliber of academic research at UVM and also promote open access (OA) to scholarship. OA mandates in response to government grant requirements were beginning to appear at large institutions such as MIT and Harvard. There was no such mandate at UVM; in fact, there was little knowledge of OA initiatives and the idea of openly sharing scholarship was viewed with some suspicion. Faculty had concerns about copyright, ownership of data, scholarship being up to par, proprietary issues (including embargoes), and being scooped.

History of UVM (E)TDs

Another major impetus for the move to ETDs was the need to improve workflow for the UVM Graduate College. Before electronic submission to ProQuest was implemented, the Graduate College handled

hundreds of paper copies of theses and dissertations (TDs) and maintained extensive files of advisor names, versions of completed work, and lists of graduates pending document completion. This paperwork could be mostly eliminated by moving to electronic submission.

There had been prior interest in an electronic version of student TDs. Between 2006 and 2008 students were requested by the Graduate College to submit electronic versions for deposit to DSpace. Turnover in the Graduate College interrupted progress toward requiring electronic submission. Five years later, the arrival of an IR and the library's interest in providing access to campus scholarship led to a successful partnership between the library and the Graduate College. Additional benefits would include streamlined submission, less paperwork for the Graduate College, and reduced processing of TDs through the library's bindery workflow. In June 2014, ProQuest Administrator was implemented and metadata began to flow between the various players. According to a senior administrator at UVM Graduate College, "the best decision the Grad College ever made was to move theses and dissertations online with ProQuest."

Obtaining the Mandate: The "Good"

Early in 2013, UVM Libraries found a willing partner in the Graduate College dean, who set up meetings with the Graduate Faculty Council to promote the idea of migrating to ETDs. Some senior faculty were skeptical of open access to scholarship, but the dean

urged them to consider the advantages of making UVM scholarship known beyond our campus. ProQuest had been in touch with our Graduate College for several years about electronic submission, but it was not until the libraries promoted the idea of partnering to provide access via the IR that the Graduate College decided to move ahead. After several meetings in which faculty concerns were addressed, the libraries drafted a plan for ETD implementation, and collaborative work on documentation began in early 2014.

A mandate (http://www.uvm.edu/sites/default/files/Electronic%20Thesis%20and%20Dissertation%20policy_0.pdf) requiring electronic submission of all graduate theses and dissertations was drafted along with permission forms (http://www.uvm.edu/sites/default/files/Electronic%20Submission%20Permissions_1.pdf) and extensive guidelines (<https://www.uvm.edu/sites/default/files/Electronic%20Thesis%20and%20Dissertation%20Guidelines.pdf>). These forms required students to grant dual permission to add their work to both the ProQuest Dissertations and Theses (PQDT) database and the UVM IR known as ScholarWorks (<http://scholarworks.uvm.edu/>). In subsequent years, it has been pointed out that contracting with ProQuest is optional and that IRs can manage ETDs on their own. There are growing concerns about student rights (Clement, 2013) and whether a university mandate to deposit TDs is in students' best interests. However, according to an article comparing ETD management in IRs with ProQuest Administrator, "there is no single 'best' system for ETD management over all. Rather, it is up to decision makers at each institution to choose an approach that best fits their university's values, goals, and needs" (Clement & Rascoe, 2013, p. 1).

For the libraries, an immediate benefit of electronic submission was the end of binding multiple copies of each thesis for graduate departments, advisers, and students' personal copies. Instead, a standing order for a single bound archival copy of each title (printed on acid-free paper) was placed with ProQuest for \$25 per title. There were early issues related to the ProQuest bindery operation, but those have largely been resolved. The archival copy is stored in the university archives. Students bind their personal copies for a reasonable cost (\$5–\$15) at Staples or the university print shop. Print copies of embargoed (delayed release) titles are received as part of the ProQuest standing order and are sequestered in a technical services processing area until the embargo is lifted.

Changing Workflows, Adding Electronic to Print: The "Bad"

Common challenges to ETD implementation include lack of understanding of intellectual property rights, particularly regarding copyright and "prior publishing." Fear of open access to new research sparked the creation of policies allowing lengthy embargoes. By 2010, the Coalition of Networked Information reported that 87% of U.S. institutions surveyed had IR policies allowing embargoes. "These kinds of fears have led to 'knee-jerk' policies of comprehensive embargo of all ETDs deposited" (Halbert, 2012, slide 36). There was also concern that an ETD "published" by ProQuest or included in an IR would be considered prior publishing and result in rejection of derived works. When surveyed, few publishers viewed ETDs as published works and even fewer refused to publish revised versions. Consequently, McMillan urged authors, "based on the data from editors/publishers' surveys, submit works based on your ETDs. Most publishers will consider [works derived from] publicly available ETDs: 89% [in the] SoSci/Humanities; 80% [in the] sciences" (McMillan, 2016, slide 31). Some ETDs contain chapters published as articles; those situations may require coordination with publisher versions.

At UVM, faculty and student concerns about access to proprietary information and pending publications resulted in a new policy offering distinct embargo periods. At the time of submission, a TD author can select no embargo or one lasting 6 months, 1 year, 2 years, or, although it has never occurred, even longer with the authorization of a dean. Since implementation, various situations, some of which are further described below, have arisen requiring clarification of guidelines about copyright, embargo periods, and embargo extensions. Iteration of policy and procedure helped improve the overall process.

Adding electronic to print workflows presented record processing challenges as well. One early problem requiring resolution arose as part of the workflow for importing ETDs into the IR. An early step in this process is to FTP the files uploaded by ProQuest, but the lack of a clear naming convention made it difficult to tell which files had already been imported into the IR and which had not. The systems librarian collaborated with IR administration to develop ways of associating particular elements, such as authors' names, with specific files and to keep track of previous imports. Currently, the workflow for importing ETDs into the IR is similar to that described

by Averkamp and Lee (2009), although updates have been made to the bepress platform since the publication of that article.

Determining how to add TD records in Voyager, the integrated library system (ILS) used at UVM, was another challenge. The first process that was developed involved using MarcEdit to run OAI harvests of the IR, convert Dublin Core to MARC, and batch edit the records before they were imported into the ILS. This approach is not uncommon (Robinson, Edmunds, & Mattes, 2016), but import was delayed while the process was under development, and, according to documentation, harvests were scheduled to take place only twice a year. When a couple of patrons asked why their theses were not in the catalog, it was clear that turnaround needed to be faster and backlogs avoided. Unfortunately, the responsible librarian's departure for another position left no one available to continue this work.

Subsequently, the systems librarian developed a process involving a timed job that downloads new ProQuest files and extracts metadata, which a Perl script uses to create semicomplete MARC records (for other examples of the use of Perl in ETD processing, see Maurer, McCutcheon, & Schwing, 2011; McCutcheon, Kreyche, Maurer, & Nickerson, 2008). By this time, a metadata librarian had been hired, and together they developed a template for the incomplete MARC records, which has itself been updated a number of times since.

Still undecided is whether or not the incomplete records should be imported into the ILS suppressed (i.e., invisible to the public). Records for embargoed titles would be hidden, but so would records for unembargoed titles, and there are arguments for exposing the latter as soon as possible, even while they look obviously unfinished. The import process would need further refinement for some records to be loaded suppressed while others are loaded unsuppressed.

Embargo Management: The “Ugly”

In the past, TD embargoes were a nonissue: only one title was embargoed by the libraries prior to 2014. Since mandate implementation, the number of embargoes has climbed in both number and proportion: 15 (19%) in 2014, 90 (54%) in 2015, 100 (61%) in 2016, and 103 (71%) in 2017 as of October 25.

Embargo management in ScholarWorks is straightforward: the full text of embargoed TDs cannot be

viewed until the embargo end date has passed, when it automatically becomes available. For print copies, however, embargoes required new procedures. Tracking processes were developed to separate embargoed from unembargoed titles. Decisions were made concerning the location of and access to print copies, particularly after one embargoed title was nearly shipped out in response to an ILL request. A no-access ILL policy was established.

A couple problems related to embargo selection emerged after ETD implementation. Occasionally, an author indicated different embargo periods in the ProQuest and IR publishing options, which made it unclear which end date should go in the catalog record and which one should be applied to the print copy. Upon communication between the IR administrator and the Graduate College, college staff agreed to verify that the same embargoes are indicated in both places and also confirmed that that embargo period applies to the print copy as well. If an author wishes to extend an embargo, the IR administrator, upon notification, can extend it in ScholarWorks, but the author must also make the extension in ProQuest; it must be clearly communicated that one without the other is meaningless.

Postembargo processing now takes place quarterly. Up to about 30 titles come off embargo in a typical quarter. It is straightforward to generate a list of embargo end dates from the IR, so after the end of each quarter, the IR administrator is asked for a report covering that period. With that list in hand, a UVM Libraries staff member prepares the newly released print copies for archiving, removes embargo notes from catalog records, and reports any errors that she finds.

The definition of *embargo* needs to be agreed upon and widely understood, including by librarians and Graduate College administrators. A school might define *embargo* to mean *no availability on the Web* but allow on-campus access to print or electronic TDs via an intranet or catalog. UVM's Graduate College defines *embargo* to mean *no access at all* to either the print or online version until the embargo is lifted. It is also important to inform all UVM Libraries staff of embargo procedures and the importance of enforcement.

Future Considerations

Improve print check-in. Print check-in ensures that physical copies have electronic counterparts in the

IR and in PQDT; it was through this process that a missing FTP load was once discovered. It is also a check for problems with binding, printing, and illustrations and for errors in titles, author names, and dates. Currently, master lists are generated in ProQuest Administrator, but additional manipulation is required before they can be used for check-in. We are looking into whether a check-in functionality can be added painlessly to an online ETD database that is already maintained by the systems librarian.

Contribute (E)TD records to WorldCat. Although some UVM TDs dating from between 2007 and the early part of 2014 are discoverable in OCLC's WorldCat, records have not been contributed since the beginning of the current ETD implementation. Resuming contribution would provide researchers with another, "powerful" way to access current (E) TDs (Lubas, 2009, p. 259), adding to the exposure they get through ScholarWorks, PQDT, the UVM Libraries' catalog and discovery layer, the Networked Digital Library of Theses and Dissertations (NDLTD), Open Access Theses and Dissertations, and search engines. Whether records for the print or electronic version are ultimately contributed, potential workflows to explore include: exporting records from the ILS; harvest of ScholarWorks by WorldCat Digital Collection Gateway (Veve, 2016a) or another tool; or harvest and transformation of ProQuest metadata (Veve, 2016b).

Reconsider cataloging practices. Whichever method is chosen for adding records to WorldCat, this moment would be an opportunity to reexamine local cataloging practices and consider the use (or not) of RDA (whose application to ETDs is described by Ashman, 2013) and of specific MARC fields and how much authority control and subject analysis to perform.

Eliminate print copies. During preliminary discussion with the Graduate College, the dean mentioned the possibility of moving to electronic-only copies in a few years. The "version of record" discussion would need clarification as the archival print copy is still

viewed by most faculty and librarians as the preservation copy of record. The potential benefits, including saving valuable staff time from check-in, avoiding print embargo management, saving the cost of binding print copies, and gaining space in the archives, would certainly provide adequate justification.

Lessons Learned

Expect changes in the scholarly communication landscape. "The ETD movement has fundamentally changed the landscape of academic and scholarly publishing, impelling stakeholders in graduate programs to reexamine historic assumptions about thesis and dissertation management and distribution" (Clement & Rascoe, 2013, p. 1).

Practice the three Cs, **Communication, Cooperation, and Collaboration**, whenever possible with faculty, students, staff, vendors, colleagues within the libraries and the Graduate College, and other campus stakeholders.

Share positive outcomes with faculty, students, and administrators to expand awareness of campus scholarly output and the benefits of sharing research. Carefully consider the impact of embargoes: they limit access to important research but also provide time for new authors to decide next steps and resolve legitimate concerns. Fears surrounding freely available scholarship can be overcome through outreach and education about the benefits of OA.

In summary: at UVM the good is getting better, the bad is getting less bad with automated processes, and the ugly has also improved significantly. Embargoes themselves are still an open question due to the complexity surrounding their motivations and impacts.

Acknowledgments

We thank our colleagues at the University of Vermont for their assistance in documenting the history of the workflows for (E)TDs at UVM.

References

- Ashman, A. B. (2013). A brief look at how RDA is being used to catalog electronic theses and dissertations. *Kentucky Libraries* 77(3), 16–23.
- Averkamp, S., & Lee, J. (2009). Repurposing ProQuest metadata for batch ingesting ETDs into an institutional repository. *Code4Lib Journal*, 7. Retrieved from http://works.bepress.com/shawn_averkamp/1

- Clement, G. P. (2013). American ETD dissemination in the age of open access: ProQuest, NoQuest, or allowing student choice. *College & Research Libraries News*, 74(11), 562–566. <https://doi.org/10.5860/crln.74.11.9039>
- Clement, G. P., & Rascoe, F. (2013). ETD management and publishing in the ProQuest system and the university repository: A comparative analysis. *Journal of Librarianship and Scholarly Communication*, 1(4), eP1074. <https://doi.org/10.7710/2162-3309.1074>
- Halbert, M. (2012). The life cycle of electronic theses and dissertations [PDF document]. Retrieved from <https://digital.library.unt.edu/ark:/67531/metadc78271/>
- Lubas, R. (2009). Defining best practices in electronic thesis and dissertation metadata. *Journal of Library Metadata*, 9(3), 252–263. <https://doi.org/10.1080/19386380903405165>
- Maurer, M. B., McCutcheon, S., & Schwing, T. (2011). Who's doing what? Findability and author-supplied ETD metadata in the library catalog. *Cataloging & Classification Quarterly*, 49(4), 277–310. <https://doi.org/10.1080/01639374.2011.573440>
- McCutcheon, S., Kreyche, M., Maurer, M. B., & Nickerson, J. (2008). Morphing metadata: Maximizing access to electronic theses and dissertations. *Library Hi Tech*, 26(1): 41–57.
- McMillan, G. (2016). ETDs and the landscape of open access publishing [PowerPoint slides]. Retrieved from <https://vtechworks.lib.vt.edu/handle/10919/65144>
- Robinson, K., Edmunds, J., & Mattes, S. C. (2016). Leveraging author-supplied metadata, OAI-PMH, and XSLT to catalog ETDs: A case study at a large research library. *Library Resources & Technical Services*, 60(3), 191–203. <https://doi.org/10.5860/lrts.60n3.191>
- Veve, M. (2016a). From Digital Commons to OCLC: A tailored approach for harvesting and transforming ETD metadata into high-quality records. *Code4Lib Journal*, 33. Retrieved from <http://journal.code4lib.org/articles/11676>
- Veve, M. (2016b). Harvesting ETD metadata from institutional repositories to OCLC: Approaches and barriers to implementation. *Journal of Library Metadata*, 16(2), 69–79. <https://doi.org/10.1080/1051712X.2016.1215730>