ETDs in Canada: Trends and Best Practices

Nancy Stuart
Technical Services Librarian, University of Victoria Libraries, Victoria, BC
Presented at ETD 2012, Lima, Peru

ABSTRACT

The focus of this paper is to present the results of a cross Canada survey of Institutional Repositories and electronic theses and dissertation programs conducted in December 2011. The purpose of the survey was two-fold. The first was to show the growth of Institutional Repositories (IRs) across Canada. The second was to illustrate the state of the electronic theses and dissertation (ETD) submission programs at Canadian institutions. The survey was a follow up to a 2009 survey and illustrates there has been steady growth in both the number of IRs and ETD programs in Canada.

Results of the survey include statistics on the number of IRs, types of materials in the IRs, the type of software platform, the number of ETD programs, whether electronic submission is mandatory, if embargoed or restricted theses and dissertations are allowed, the number of institutions having their ETDs harvested by Library and Archives Canada and the number of institutions sending their theses to ProQuest and retaining a microfiche copy. The survey was sent to Canadian institutions where a thesis or dissertation is required for graduation and to Canadian institutions who are members of the Canadian Association for Research Libraries. Responses were received from 33 institutions.

Best practices for ETD programs and workflows will also be discussed. Focus areas include the PDF format for multiple and multi-media files, OAI-PMH harvesting, mandatory submission and embargoes. By analyzing the results of the survey, it is clear that Canada is moving forward quickly in the implementation of ETD submission programs. The fact that, from the survey results, over 70% of Canadian institutions have their ETDs harvested by Library and Archives Canada puts Canada, and it’s National Library, in a unique position to share its knowledge and expertise in ETD programs and ETD workflows.

HISTORY OF THESES AND DISSERTATIONS IN CANADA

Since 1965 the Theses Canada program at Library and Archives Canada has been acquiring theses and dissertations from universities across Canada. (Reeves, 2010) For space reasons, these were acquired as microforms, in most cases, microfiche. In 2003 Theses Canada embarked on a pilot project to acquire electronic theses and dissertation. During 2004 and 2005 a few early adopters of ETD programs in Canada participated in a pilot project to harvest ETDs using the OAI-PMH protocol. This protocol is intended to only harvest metadata, but an extension was added to also harvest the ETD file. Microfiche was then considered the preservation copy and the ETD was the accessible digital copy. Open access ETDs allowed many more people to access the research in theses and dissertations in Canada. The goal of Theses Canada is to acquire all Canadian theses and dissertations in electronic format and ultimately have the digital copy as the preservation copy. Another goal of Theses Canada is to help universities
transition from paper based to electronic theses and dissertations. To that end a number of surveys have been conducted to determine the progress of ETD submission programs in Canada. Library and Archives Canada has benefited from acquiring ETDs by making them much more accessible to researchers and reducing physical space. Preservation of the ETDs in Canada is a work in progress, as it is in many countries.

SURVEY METHODOLOGY

In the Fall of 2011, a survey instrument was developed at the University of Victoria to determine the trends and some of the best practices of ETD programs across Canada. The purpose of the survey was two-fold. The first was to determine the growth of Institutional Repositories (IRs) across Canada. The second was to determine the state of the electronic theses and dissertation (ETD) submission programs at Canadian institutions granting graduate degrees, where a thesis or dissertation is a requirement for graduation. The survey was a follow-up to one conducted in April/May 2009. Responses were received from over 50% of the institutions that received the survey.

TRENDS

Institutional repositories

An institutional repository is a digital collection of an organization’s intellectual output. Institutional repositories centralize, preserve, and make accessible the knowledge generated by academic institutions. (Canadian Association of Research Libraries, 2012) “Since the term institutional repository was coined by SPARC in 2002”, (Mondoux & Shiri, 2009) Canada has seen a steady growth in the development of IRs at academic institutions. The Canadian Association of Research Libraries reports that in 2003 there were just 2 CARL members with IRs. By 2008, there were 22 CARL members that had institutional repositories. The survey conducted in the Fall of 2011 confirmed the steady growth of IRs across Canada and 25 of the 33 institutions surveyed (76%) had an IR. Half of those, that did not have an IR, have plans to implement one. It is evident from the registry of open access repositories (ROAR, http://roar.eprints.org/) that continued growth is a global phenomenon and Canada is typical of this activity.

There are many software platforms available for implementing an institutional repository. From the Fall 2011 survey, the most popular software used for the Institutional Repository is DSpace (63%). DSpace is an open-source software released by Massachusetts Institute for Technology for archiving academic content. Others mentioned were Digital Commons (Bepress), Digitool (Ex Libris), E-Prints, Fedora, Drupal, and ETD-db. A few institutions had home grown products. (Fig. 1). These findings mirror the listings in the Registry of Open Access Repositories, which also lists DSpace as the most commonly used software platform.

---

In this paper, the word “institution” refers to an institution granting a graduate degree, usually a university or college.
ETD programs

Since 2009 there has been steady growth in the number of ETD programs in Canada. The results also showed that there is planned future growth in the number of ETD programs across Canada. Twenty-two institutions (67%) have an ETD program, and in 63% of Canadian institutional repositories, ETDs are the largest collections. The two earliest ETD programs launched in Canada were in 1996 and 2001. From 2004 on there has been a steady rise in the number of institutions launching ETD programs (Fig. 2). In Fig. 2 the blue bar represents the number of new ETD programs launched each year, while the red bar represents the total for that year.

For those institutions without an ETD program, 64% have plans to implement one in the near future.

When ETD programs were first implemented, most institutions made electronic submission voluntary. It became clear that to have a complete electronic collection of theses and dissertations, institutions would have to require mandatory submission of ETDs. The findings of our Fall 2011 survey show that of the institutions that have an ETD program, almost 60% have mandatory submission as a requirement. In
the previous 2009 survey only 30% required mandatory submission. The trend for new ETD programs is to require mandatory electronic submission.

**Embargoed or restricted ETDs**

In the ETD environment, the word embargo or restricted is used interchangeably. “In academic publishing, an embargo is a period during which access is not allowed to certain types of users”. (Embargo (academic publishing), 2012) In various ETD collections, the term restricted usually refers to a collection of theses or dissertations that are restricted to a group of users, often faculty, students and staff of a particular institution. In DSpace the term embargo is used to withhold the document from all users, allowing only the metadata to be viewable. To restrict or embargo the document to a group of users, a variety of processes are available for use. The process used would depend on the software, hardware or IT expertise that is available at a particular institution.

With more and more theses and dissertations being Open Access, the embargo or restricted theses has become an issue. A number of years ago when most theses or dissertations were in print, a student could withhold their thesis for two main reasons: patent pending or pending publication. The library would not receive the copy of the print thesis until the withhold date had expired. Currently with mandatory electronic submission, embargoeing or restricting theses has become more complicated. In our survey we asked if institutions had restricted or embargoed theses and for what period. 72% said they allowed for restricted or embargoed theses and the period of time varied from 4 months to forever. Common intervals were 6 months / 1 year / 2 years / 5 years. Most institutions allowed for a renewal of the restriction. For those institutions with ETD programs, 22% did not allow restricted or embargoed ETDs as an option. The IR software platform used will determine how embargoed or restricted theses are managed.

**Library and Archives Canada harvesting**

Since 2003 Library and Archives Canada has been acquiring ETDs from university participants (Reeves, 2010). The results of the Fall 2011 survey showed that of the 22 institutions with ETD programs, 71% of those ETD collections are now being harvested by Library and Archives Canada. (Fig. 3).
The harvesting program at Library and Archives Canada (LAC) has been experimental and unique. It was unique in the sense that along with the metadata, the PDF was also harvested. In 2009, four universities were being actively harvested by Library and Archives Canada. When universities upgraded their IR software, the LAC harvester would break and require much programming time to fix it. During 2010, this issue was rectified and currently sixteen Canadian universities have their ETDs harvested regularly by the LAC harvester.

**Preservation**

“Microfiche or microfilm has been the standard preservation medium for theses and dissertations. In the United States, colleges and universities have been steadily transitioning from traditional paper/microfilm to digital ETD submission, dissemination, and preservation processes”. *(Lifecycle management of electronic theses and dissertations: IMLS NLG proposal. 2011)* Since 2011 Library and Archives Canada has discontinued keeping the microform copy as a preservation copy. The trend is to use the digital copy for preservation. Library and Archives Canada has a goal of collecting Canadian digital theses and dissertations for long term preservation. From the survey results only 36% of the institutions sending their ETDs to ProQuest\(^2\) are still receiving a microform copy, a sharp decline from 2009. This result illustrates the trend toward digital copy preservation.

**BEST PRACTICES FOR ETDs IN CANADA**

There are already a number of best practices documents for ETDs, both on the NDLTD website (Networked Digital Library of Theses and Dissertations, 2011) and the Theses Canada Portal website (Theses Canada, 2010). In Canada we base our best practices on the document from NDLTD, but our Canadian best practices reflect the unique nature of Canadian theses and the way that Library and Archives Canada collects the digital copy. Also as technology advances, ETD best practices must evolve and develop along with the technology.

**Mandatory submission**

Best practice is to start the ETD program as mandatory. If it is voluntary, it is difficult to transition to mandatory in an expedient way. Mandatory submission ensures all ETDs are archived. Policies will need to be put in place to allow for embargoes or restricted ETDs. Normally mandatory submission must be approved by the Faculty of Graduate Studies and/or Senate of the institution.

---

\(^2\) ProQuest is a US based company that provides electronic and microform information products and services to academic, school, public, corporate and government libraries around the world. It has been publishing theses and dissertations since 1938. They distribute citations and abstracts of theses and dissertations through ProQuest Dissertations & Theses (PQDT) database.
Embargoes

Recently there has been much discussion on the ETD listserv3 about embargoed and restricted theses. One U.S. University reports that 19% of their ETDs are viewable to campus only and 17% are not viewable for the entire embargo period. This is a total of 36% being restricted in some manner. Others reported similar numbers in the range of 25% – 34% being embargoed or restricted.

Thus far, there is no consensus on “good practices” in the field concerning embargoes/restrictions. (Lifecycle management of electronic theses and dissertations: IMLS NLG proposal, 2011) As stated in the “Lifecycle management ...” document, there are issues and two different sides of this argument for those considering an ETD program, clearly articulating the pros and cons claimed by proponents for and against embargoes. The point of this document is not to settle the issue, but provide prospective ETD programs with clear and concise information about the ramifications of campus policy decisions for or against different kinds of access restrictions. The document outlines the two year project (2011-2013) to develop and disseminate guidelines, educational materials and an associated workshop, and a set of software tools for life-cycle data management and preservation of ETDs. The results of the project will be of great benefit to the ETD community.

There has been discussion about indefinite embargoes or restricting a thesis or dissertation forever. This long term or forever embargo was most often from a creative writing discipline, where the student had submitted poetry or a novel as their thesis. In most cases the student wanted to publish the work and was obliged by the future publisher to have not published the work in Open Access. Academic institutions have dealt with this in a variety of ways. Some will allow an indefinite or forever embargo in exceptional circumstances. Many institutions now include information to their graduate students on the benefits of Open Access and explain the reasons a student may or may not want to restrict their thesis or dissertation. Many faculties of Graduate Studies have set policies with specific reasons for withholding a thesis4. A normal period for a restriction is 6 months, 1 year or 2 years.

Best practice is to have a written policy for restricted or embargoed theses or dissertations and to communicate that carefully to the graduate students and their supervisors. Best practice is to have the student submit a form, signed by the appropriate people, usually the student, supervisor and Dean of the Faculty of Graduate Studies. In the case of the University of Victoria, the Creative Writing Department decided to have graduate students do a project rather than a thesis, if they were going to submit an original piece of creative work such as poetry or a novel. At our institution a graduate student working on his Masters degree can choose to do a graduate project, which is not as formal as a thesis and is not required to follow the policies laid out in the Faculty of Graduate Studies for submitting a master’s thesis. The embargo issue is avoided because they are not required to submit a graduate project.

---

3 ETD-L@LISTSERV.VT.EDU

Metadata

Metadata allows for discovery of ETDs. The best practice for an ETD program is to use the metadata schema ETD-MS: an Interoperability Metadata Standard for Electronic Theses and Dissertations found on the NDLTD website. [http://www.ndltd.org/standards/metadata/etd-ms-v1.00-rev2.html](http://www.ndltd.org/standards/metadata/etd-ms-v1.00-rev2.html) If that schema is implemented from the start, it will be possible for your ETD collection to be discovered and harvested. This schema is based on Dublin Core with a new element called thesis.degree. It has four qualifiers: name, level, discipline and grantor. Depending on the software, it is possible to add elements to the schema. Best practice is to follow the qualifiers suggested by DCMI Metadata Terms [http://dublincore.org/documents/dcmi-terms/](http://dublincore.org/documents/dcmi-terms/)

The following is the metadata for an ETD from the Virginia Tech ETD collection presented in the “Vanilla” XML encoding of the ETD-MS standard taken from the NDLTD website.

```xml
<thesis xmlns="http://www.ndltd.org/standards/metadata/etdms/1.0/"
xsi:schemaLocation="http://www.ndltd.org/standards/metadata/etdms/1.0/
http://www.ndltd.org/standards/metadata/etdms/1.0/etdms.xsd">
<creator>Walker, Gordon James</creator>
<subject>outdoor recreation</subject>
<subject>recreation experience preference scales</subject>
<subject>recreation experience matrix</subject>
<description>This dissertation examines four issues ....</description>
<publisher>Virginia Polytechnic Institute and State University</publisher>
<contributor role="chair">Joseph W. Roggenbuck</contributor>
<date>1997-03-31</date>
?type>Electronic Thesis or Dissertation</type>
<format>application/pdf</format>
<identifier>
http://scholar.lib.vt.edu/theses/available/etd-3345131939761081/</identifier>
<language>en</language>
<rights>unrestricted</rights>
<degree>
<name>PHD</name>
<level>doctoral</level>
<discipline>Forestry</discipline>
<grantor>Virginia Polytechnic Institute and State University</grantor>
</degree>
</thesis>
```

To acquire the metadata for an ETD, it is best practice to have the graduate student submit their ETD to the Institutional Repository. Most Canadian ETD programs allow the student to submit keywords for subject access and do not impose a controlled vocabulary.
The NDLTD website also has many helpful resources about metadata, a MARC-21 crosswalk and includes a link to Bailey’s bibliography on IRs and ETDs (Bailey, 2011). In that bibliography there is a section on Metadata, which lists many citations to up to date articles covering ETDs and metadata.

Multi-media / Multiple files

When ETDs first came on the scene, they were digital copies of text and images, archived as a single PDF file. With changes in technology, ETDs have given graduate students wide latitude in creating a digital thesis or dissertation. NDLTD has been giving awards for innovative ETDs since 2004. The Innovative ETD Award recognizes the student efforts to transform the genre of the print dissertation through the use of innovative software to create cutting edge ETDs. Over the years, students have won the award with their ETD being a website, or one that contains video or other multi-media objects to illuminate the content of the dissertation.

In Canada, our graduate students have been embracing the ETD format and creating theses and dissertations beyond a single text file in digital form. In Canada, the Library and Archives Canada harvests our ETDs, but require a single PDF. If a submission has multiple files it will only harvest the first PDF. Many students have their data and supplementary files in formats other than PDF, such as an excel spreadsheet, a .jpg file, or a .mov file. In order for these ETDs to be harvested by Library and Archives Canada, we looked to the Adobe Acrobat software where it is possible to create a PDF e-portfolio which contains a number of files in various formats. We have mandated that our graduate students may only submit a single PDF file and if they have multiple files or formats, they are required to create a single e-portfolio. We have Adobe Acrobat Pro 10 available in our Music and Media Commons where students can create their e-portfolio and access help if needed. The students have adapted to this very well and just follow the guidelines in the FAQs http://libanswers.uvic.ca/a.php?qid=77139 and http://libanswers.uvic.ca/a.php?qid=77141 explaining how to submit multiple files using an e-portfolio. An example is Doroschuk http://hdl.handle.net/1828/3427 This thesis was submitted in print in 2005 with an accompanying VHS. To preserve the VHS we had it converted to an .mp4 and digitized the thesis and created an e-portfolio which contains the PDF of the thesis text, 1 .mp4 file and 11 .cda audio files.

Preservation

Preservation is a concern for all librarians working with digital files. Much has been written about the need for preservation of digital assets. MetaArchive is a U.S. community-owned, community-led initiative that is a secure and cost-effective repository that provides for the long-term care of digital materials. It is best practice to include in any ETD program, a strategic plan for long-term preservation. Most countries and regions of the world are moving in this direction in order to preserve these digital assets for future generations.

---

5 NDLTD ETD awards http://www.ndltd.org/events_and_awards/awards/etd-awards/

6 MetaArchive Cooperative is a digital preservation network created by six libraries in the southeastern United States http://www.metaarchive.org/
Europe's DigitalPreservationEurope(DEP)\(^7\) is one such example. 

Previously, many universities across Canada sent their theses and dissertations to ProQuest for microfilming, digitization and inclusion in the Dissertations and Theses PQDT database. The microform was considered the preservation copy and the digital copy was for easy access. With more institutions going to digital only, ProQuest could be an option for preservation of the digital ETD. If Canadian universities continue to send their theses to ProQuest, the ETD would be preserved as part of the ProQuest ETD collection. “ProQuest is committed to following best practices related to preservation, and has taken an active role in the area of digital preservation”.(ProQuest, 2012)

Nine Canadian University libraries are collaborating to preserve collections important to the provinces of British Columbia, Alberta, Saskatchewan and Manitoba through the LOCKSS\(^8\) network called the Council of Prairie and Pacific University Libraries (COPPUL) Consortium. This group has a particular focus on freely available born digital web content including government documents, e-journals, ETD collections and small presses.

FUTURE OF ETDs IN CANADA

ETD programs in Canada continue to develop and expand. To develop and expand best practices, Canadian institutions collaborate with Library and Archives Canada, the Canadian Association of Research Libraries and the Networked Digital Library of Theses and Dissertations. These organizations offer support to new and developing ETD programs, by sharing best practices and holding conferences and workshops. As stated on the Theses Canada website, “the mission of Theses Canada is to acquire and preserve a comprehensive collection of Canadian theses at Library and Archives Canada (LAC), to provide access to this valuable research within Canada and throughout the world. Its mission to preserve this portion of Canada's bibliographic heritage is achieved through collaboration with the many Canadian universities who participate in the program”.

Many ETD programs across Canada share their knowledge, expertise and workflows with each other, through networking, and the ETD list-serv\(^9\). It is an active community and we are actively working towards a Canadian ETD collection which is accessible to the world and preserved for future researchers.

---

\(^7\) DigitalPreservationEurope(DEP) fosters collaboration between many existing national and international initiatives across the European Research Area. It addresses the need to improve coordination, cooperation and consistency in current activities to secure effective preservation of digital materials. [http://www.digitalpreservationeurope.eu/](http://www.digitalpreservationeurope.eu/)

\(^8\) The LOCKSS Program, based at Stanford University Libraries, is an open-source, library-led digital preservation system built on the principle that “lots of copies keep stuff safe.”

\(^9\) [ETD-L@LISTSERV.VT.EDU](mailto:ETD-L@LISTSERV.VT.EDU)
References


