

# **An Overview of the National Theses Program at Library and Archives Canada**

**Sharon Reeves**

Manager, Theses Canada, Library and Archives Canada, Ottawa

## **ABSTRACT**

A centrally coordinated program for theses accepted at Canadian universities was established in 1965 at the National Library of Canada (NLC)<sup>1</sup>. From 1965 to 2003 the program was delivered in a traditional fashion – paper theses were reproduced on microfilm or microfiche, which were then made available to researchers via interlibrary loan or purchase. In January 2004 Theses Canada launched the Theses Canada Portal, which provides open access to over 45,000 electronic theses. Staff at Library and Archives Canada (LAC) went on to develop LAC's first OAI (Open Archives Initiative) harvester and repository in order to harvest electronic theses and metadata directly from universities with electronic theses submission programs.

In this paper I will discuss the development of the electronic theses program at Library and Archives Canada and the current status of ETD initiatives at Canadian universities.

## **1. HISTORY OF THE NATIONAL THESIS PROGRAM IN CANADA**

The centrally coordinated national thesis program at Library and Archives Canada (LAC) was established in 1965 at the National Library of Canada (NLC) at the request of various deans of Canadian graduate schools. The program had two objectives:

- to facilitate access to theses by identifying what was available and by providing copies via ILL or sale; and,
- to preserve theses by microfilming paper copies and storing preservation masters according to archival standards.

For thirty seven years the program continued to fulfill its mandate in much the same way. The only significant changes to program delivery came first in 1990 when NLC outsourced the reproduction and sales aspects of the program and then in 1998 when, under the terms of a new contract, UMI Dissertations Publishing (UMI)<sup>2</sup> began to digitize the theses as well as reproducing them on microfiche. Since the beginning of the program theses authors have signed a non-exclusive license with NLC, which initially allowed it to reproduce their theses on microfilm and which was revised in the early 1990s to allow for the distribution of theses in electronic format on the Internet.

## **2. PARTNERSHIPS**

Participation in the national thesis program was and continues to be voluntary. In 1965 five universities became the Canadian Theses Service's (CTS) initial partners. Over the past forty

---

<sup>1</sup> Now part of Library and Archives Canada.

<sup>2</sup> Now part of ProQuest Information and Learning.

years participation in the program has steadily grown to the point where Theses Canada currently has partnerships with sixty universities. One of the unique aspects of this arrangement is that there is no written agreement between each university and Theses Canada. The only criteria for joining the program is that a university must belong to the Association of Universities and Colleges of Canada. This makes for a flexible situation for universities, which can stop participating in the program at any time. It also leaves Theses Canada free to develop its program for its university partners as a whole, after undertaking any necessary consultations.

### **3. ETDS IN CANADA**

During the late 1990s interest in electronic theses and dissertations started to percolate at various Canadian universities in response to the initiatives concerning ETDs that were taking place in the United States and elsewhere. Several notable examples from this time period are Simon Fraser University, which experimented with using Adobe PDF format to mount its masters theses on the Web and the TUG (TriUniversity Group) Electronic Theses Project. Partners in TUG were the University of Guelph, the University of Waterloo and Wilfrid Laurier University. While this joint initiative ultimately failed, the University of Waterloo went on in 1999 to implement the first Canadian ETD submission program. Other universities investigating the possibility of ETD submission at the time were l'Université de Montréal and l'Université Laval.

### **4. NATIONAL CONSULTATION ON ETDS**

It was natural for the Canadian academic community to look to the National Library of Canada for leadership in those early days of transition from print to electronic theses and accordingly NLC organized a national consultation on e-theses, which took place December 4, 2000. Seventy-eight participants from across Canada took part in the consultation. Represented were deans of graduate schools, university administrators and professors, library professionals and students, as well as participants from various organizations associated with academia. The main objective of the consultation was to discuss the national theses program in an electronic environment. Out of the consultation came the proposal for a Canadian portal for electronic theses.

The second important issue discussed at the consultation was that of sales of Canadian theses by NLC's contractor. Concerns had been expressed by a number of graduate students and others to the effect that the agreement between NLC and UMI, allowing UMI to sell copies of theses, violated a fundamental principle to ensure freedom of access to scholarly information. To quote William Maes, University Librarian at Dalhousie University and a speaker at the consultation: "Does the current commercial situation lead us down a path that restricts access to the very things we produced? This issue is not that different from our current situation with journal publication where authors freely give up their rights in order to be published, while the very reason for publishing is thwarted by the high cost of getting that information back. ... Part of today's discussion emanates from a belief that electronic theses will bring down costs and allow us to take back ownership of the distribution and preservation of our own works" (Maes 2000).

### **5. THE THESES CANADA PORTAL**

After the consultation a national advisory committee was established for the Canadian Theses Service. At the same time theses and systems staff at NLC started to develop a proposal for a

searchable collection of Canadian theses on AMICUS, Canada's national online catalogue. This ambitious proposal was finished by spring 2002. It included the addition of abstracts to bibliographic records, the development of a web-based search interface for the numerous theses and dissertations in NLC's collection and of several new search indexes specific to theses. It also included a proposal for the patriation of theses digitized for the national program by UMI. This proposal became the foundation for the Theses Canada Portal.

Over the course of the fall of 2002 and throughout 2003 the design and systems development for the Portal was done. On April 1, 2003, in keeping with the new direction of the national program and to match the name chosen for the Portal, the Canadian Theses Service changed its name to Theses Canada. Throughout 2003 UMI sent NLC electronic versions of all the theses it had digitized under the terms of the contract in effect between 1998 and August 2002. In all NLC received over 45,000 ETDs in PDF format from UMI.

The Theses Canada Portal was launched on January 29, 2004 at the Ontario Library Association Superconference. At the time it provided access to NLC's extensive collection of theses and dissertations on microform, a collection that currently numbers over 250,000. It also provided free online access to the subset of ETDs received from UMI as well as extensive information about the national program. The Portal was further enhanced in 2004 by indexing the theses and dissertations to enable clients to search their full text. It can be accessed at <http://www.collectionscanada.ca/thesescanada>.

## **6. OAI AT LIBRARY AND ARCHIVES CANADA**

While the launch of the Theses Portal and the provision of open access to a large number of ETDs was a significant achievement, it still left the national thesis program dependent on sources outside NLC to realize its mandate. NLC's agreement with its contractor precluded the provision of immediate open access to the theses digitized on its behalf. What was needed was a way of, in the words of Bill Maes, taking back ownership of the distribution and preservation of Canadian theses and dissertations. The solution was to develop a national electronic theses program at NLC.

At its second meeting, in October 2002, the Canadian Theses Service Advisory Committee (now the Theses Canada Advisory Committee), struck a technical sub-committee composed of representatives from CTS, l'Université Laval and the University of Waterloo. The sub-committee was charged with resolving problems with respect to the transition to electronic theses in Canada and to develop the means for universities to send their ETDs directly to the NLC. This second objective was the direct result of concerns expressed at the national consultation in 2000. At the October meeting the Advisory Committee recommended that the Canadian Theses Service conduct a pilot project to acquire e-theses from l'Université Laval and the University of Waterloo. Both universities had already established ETD submission programs.

The sub-committee began its discussions in January 2003 and its work progressed concurrently with the development of the Theses Canada Portal. Its first task was to establish how the NLC would acquire ETDs from universities. The solution was obvious – NLC would harvest metadata from universities using the Open Archives Initiative Protocol for Metadata Harvesting (OAI-PMH). L'Université Laval had already set up its program to meet OAI requirements and the University of Waterloo was taking steps to make its metadata compliant for harvesting. The advantages of setting up a national electronic thesis program at NLC based on OAI open source software were immediately apparent. Once NLC was set up as both an OAI service provider

and data provider it could harvest metadata for ETDs from universities and in turn make it available for harvesting by other organizations such as the Networked Digital Library for Theses and Dissertations (NDLTD). The harvested metadata would be converted to MARC 21 format and uploaded to AMICUS, the National Library's online catalogue.

In developing its harvesting program NLC faced a challenge that most institutions do not have. In order to continue to expand its collection of theses and dissertations, thus fulfilling the institution's mandate to acquire and preserve Canada's documentary heritage for current and future generations, NLC needed to harvest not only theses metadata but the electronic theses and dissertations themselves.

At the time that these early discussions took place the Government of Canada announced that the National Library of Canada and the National Archives of Canada would be combined to form a new institution – Library and Archives Canada (LAC). As of April 1, 2003 the two institutions would act as one, although the supporting legislation would not be passed for another year. This would have some implications for future systems development of the Theses Portal.

## **7. THE PILOT PROJECT**

In October 2003 the technical sub-committee met in Ottawa, along with representatives from LAC's Information Technology Branch and Standards Division, to make some decisions that needed to be made before the pilot project could begin. The sub-committee discussed the need for an identifier on the metadata records that would make it possible to organize the harvested theses in the directory on LAC's server. To deal with this staff at LAC developed a TC [Theses Canada] number which includes the ILL library symbol for each university so that the harvested theses could be organized by university on the LAC server.

The LAC harvester and repository were developed after the meeting and were ready by the end of 2003. An extension was written to the harvesting program to allow for the harvesting of the e-theses themselves. The target date for the start of the six-month pilot project was April 1, 2004. LAC staff intended to evaluate the pilot project in October 2004 in order to identify and resolve problems and to assess the resource implications at LAC.

In 2004 a number of factors conspired to slow the project down. As alluded to earlier a significant issue was the transformation to Library and Archives Canada, which resulted in the identification of a number of new initiatives. Many systems staff were deflected from their ongoing projects to work on major systems developments related to the new initiatives. In spite of this LAC started to harvest metadata and PDFs from Laval and Waterloo.

The delay meant that Theses Canada was able to bring two more universities into the pilot. The University of Saskatchewan started its ETD submission program in December 2002 shortly after the CTS Advisory Committee meeting. In July 2004 university personnel contacted Theses Canada to ask when it would be possible to send its metadata to LAC. Theses Canada invited the university to join in the pilot project. The university had developed its ETD submission program using the ETD\_db freeware from Virginia Polytechnic Institute and State University, which allowed a thesis to include multiple PDF files. The harvesting program at LAC, on the other hand, had been written on the basis of harvesting only one file. Participation of the University of Saskatchewan allowed LAC to identify this weakness in its program, which will be fixed when sufficient time and resources are available.

The University of Manitoba began to participate in the pilot project in fall 2004. Its Faculty of Graduate Studies gave its approval to set up an ETD pilot project in May 2004, using the DSpace institutional repository software that the University Library had already started to implement. By November the university already had eight ETD records in its repository. As a result of some communications between systems staff at the university and LAC the records were harvested. This turned out to be extremely serendipitous in that it was discovered that DSpace only supports metadata in the Dublin Core metadata standard. As a result, the programmer at the University of Manitoba developed an DC to ETD-ms crosswalk, which the university has made available to other implementers of DSpace for ETDs (available at <http://mspace.lib.umanitoba.ca/ETDMS.zip>).

The delay also gave the Manager of Theses Canada time to prepare two sets of guidelines during the summer of 2004. A draft document on harvesting requirements was prepared and sent to all stakeholders for review. At the same time guidelines on *How to Set Up an ETD Submission Program* were made developed. Since April 1, 2005 these guidelines have been available on the Theses Canada Portal at <http://www.collectionscanada.ca/thesescanada/s4-210.1-e.html>.

## **8. PILOT PROJECT EVALUATION**

The 693 harvested records from all four universities were finally converted to MARC 21 and uploaded to AMICUS on March 30, 2005, six months later than expected. In April 2005 LAC staff reviewed the records, which they found to be good quality access records although the records do not fully comply with international standards. However there were several problems with the conversion that needed to be addressed, such as HTML symbols appearing in some abstracts, as well as some issues concerning internal workflow to resolve.

Over the course of the summer and fall the conversion problems were slowly addressed and fixed insofar as possible and the internal workflow issues were resolved. This made it possible to finalize the guidelines on Theses Canada's harvesting requirements. The harvester, which was stopped in April, was restarted in December. The next step is to start setting up other universities for harvesting and to upload changes about the program to the Theses Canada Portal.

## **9. STATUS OF ETD INITIATIVES IN CANADA IN 2005**

In 2002 when the Canadian Theses Service started to develop the Theses Canada Portal, the University of Waterloo was the only Canadian university with an established ETD submission program. L'Université Laval had finished its pilot project and was just starting to accept ETDs. Since then interest in ETD initiatives has burgeoned. Theses Canada is in the process of conducting a telephone survey on the current status of ETD initiatives at Canadian universities as a more comprehensive follow-up to the quick and dirty survey it did in 2004.

At the present time ten universities have yet to be called. This will be done by March 31, 2006. As well we decided not to survey fifteen universities. This decision was based in some cases on their replies to the 2004 survey and in some cases on their size. For example ten of them are small theological colleges. Other results reveal that seven universities have implemented ETD submission programs, our four pilot universities and three that are waiting to be harvested, two have conducted pilot projects but in both cases the projects are on hold for various reasons, four are in the pilot planning process, three are working on proposals for a pilot, ten are

conducting preliminary discussions which would lead to proposal development and eight have indicated that they have no plans to look at ETD initiatives at the present time. The open source DSpace Institutional Repository software has proven to be a popular choice with a number of universities – seven universities have used or are planning to use it for ETDs.

## **10. AUTHOR ISSUES**

The majority of theses authors support the principle of open access to ETDs, however a small percentage of them do have concerns. These concerns invariably fall into one of two categories: damage to future publication potential and/or the possibility that their research will be plagiarized.

### **10.1 Future Publication Potential**

A small number of students are concerned that making their electronic theses and dissertations widely accessible on the Theses Canada Portal will have a negative impact on their ability to publish them commercially however, in an informal survey conducted during the fall of 2003 by the Graduate Students' Association of Canada (GSAC), Taylor and Francis, the University of Toronto Press, the University of Alberta Press and the University of British Columbia Press indicated that there would be no negative impact on their decision to publish manuscripts based upon a student's dissertation if the theses were first made available on the Theses Canada Portal. In fact some publishers thought that the Portal initiative was a very positive step.

Theses Canada recommends that authors with concerns contact potential publishers in advance to determine whether or not prior access to their electronic theses via the Theses Canada Portal will prevent publication of all or part of their theses later on. To determine the publishing policy of various academic journals graduate students can check the Academic Journal Policy Database at [www.etd.uc.edu/journal](http://www.etd.uc.edu/journal).

### **10.2 Plagiarism**

A small number of theses authors are concerned that exposing their research on the Internet will leave it open to plagiarism. In the *UNESCO ETD Guide*, Dr. Jean-Claude Guédon, from the Université de Montréal, points out that “by apparently giving away the results of his/her work, a scientist ensures his/her intellectual property most effectively and that the ability to compare new results to already published work makes plagiarism a very risky business”. He goes on to point out that placing theses online and maximizing their visibility offers a very efficient way to protect intellectual property and prevent plagiarism (Guédon, 2001).

## **11. BENEFITS TO AUTHORS**

The single most significant benefit to theses authors relates to citation impact – online research is more accessible and therefore cited more often. As a reflection of the degree of uptake of research results by other researchers this can contribute to rewards from universities through employment, increased salary and promotion and rewards from research-funders (Harnad, 2005). Publication of electronic theses and dissertations results in:

- free publicity for post-graduate level research – the authors of electronic theses become much more widely known and as a result their reputations and status are enhanced;
- easy worldwide access to their theses for colleagues and collaborators;
- easy worldwide access to their theses for job and grant applications;
- the ability to include multimedia in ETDs;

- a raised profile for their research institutions; and,
- savings from not having to print, copy and bind paper theses.

## **11. ETD 2006, 9<sup>th</sup> INTERNATIONAL SYMPOSIUM ON ETDs**

The 9<sup>th</sup> International Symposium on Electronic Theses and Dissertations will be hosted by Université Laval in Quebec City from June 7 to 10, 2006. This is the first time the conference will be held in Canada and provides a tremendous opportunity for staff at Canadian universities to learn about all the various aspects and issues related to ETDs.

## **12. CONCLUSION**

In the next two to three years it's expected that many of the universities currently in the planning or pilot stages of their ETD initiatives will put ETD submission programs in place. This will be a busy period for Theses Canada as it starts to harvest theses and metadata from many of its partners while at the same time maintaining the traditional program it provides through its contractor. However in three years Theses Canada will be well on its way to establishing a national electronic theses program at LAC.

## **13. REFERENCES**

Guédon, J-C. (2001). *Plagiarism*. UNESCO ETD Guide, Section 2.3.6. Retrieved October 24, 2005, from <http://www.etdguide.org/>.

Harnad, S. (2005). *Australia is Not Maximising the Return on its Research Investment*. Conference Proceedings of the 8<sup>th</sup> International Symposium on Electronic Theses and Dissertations. Retrieved October 24, 2005 from <http://adt.caul.edu.au/etd2005/program.html>.

Maes, W. (2000). *Consultation on Electronic Theses*. Retrieved June 13, 2005, from <http://www.collectionscanada.ca/icbn/040021-401-e.html>.