

Campus Collaboration: The University of Waterloo Electronic Thesis Project
<http://www.lib.uwaterloo.ca/~cjewell/ETD/Show/index.htm>

By Christine Jewell
University of Waterloo
Ontario, Canada
March 15, 2000

The University of Waterloo Electronic Thesis Project Team is a collaborative effort of individuals and groups across the Waterloo campus and beyond. This collaboration has provided the foundation needed to face the many challenges of ETDs. UW's project team has progressed through a series of collaborative phases. This paper will focus on four distinct phases. The initial formation of the team followed discussions between the University Librarian and the Dean of Graduate Studies. In the information gathering phase the team contacted individuals and groups across North America and around the world. In the Pilot Project phase, we worked with other departments on campus to test the feasibility of electronic access to theses. In the TriUniversity collaboration phase, we formed an alliance with librarians and graduate officers from two neighbouring institutions. With successful collaboration, the benefits are ongoing. I'll conclude this paper with an illustration of how our cooperative connections have been relied upon through to the present phases of the project's development.

The initial collaboration began when the University Librarian and the Dean of Graduate Studies first discussed the idea of electronic theses. They discussed the advantages of electronic submission and distribution of theses. Advantages such as increased access for researchers, powerful medium for expression of ideas, reduced photocopy costs for students, and time and space savings for the library and graduate office. The University Librarian and the Dean of Graduate Studies agreed to collaborate on a project to make electronic theses a reality at the University of Waterloo.

As a librarian experienced in electronic text preparation I was charged with the mandate to head up a team. The Graduate Studies Office and the Library form the core of the UW Electronic Thesis Project. The Graduate Studies Office has focussed on thesis submissions; the Library is concerned with storage and access. Branching out from this core membership, the team also includes computer specialists. Graduate students and faculty have been involved throughout all phases of the project.

The team's first challenge involved a very steep learning curve. We needed to become familiar with the existing procedures for dealing with theses in the Graduate Studies Office and the Library. We needed to learn the terminology used in the different departments. And we needed to develop a basic understanding of some technical concepts, such as file size and formats. We also needed to learn from each about our various concerns regarding electronic theses, as well as our hopes for electronic theses, and what we each saw to be the benefits of electronic theses.

As we became more familiar with the assignment, we came to the realization that we needed to acquire a fuller understanding of the area before we could carry out our mandate. This was an assignment without a precedent on campus. There was no pre-existing plan to determine how we should go about making electronic theses a reality at the University of Waterloo. We realized very early on that our first step must be to initiate an information-gathering phase. We resolved to seek out others with experience in this area and gather information from beyond Waterloo.

The team designed a Web-based questionnaire (January 1997) to find out what other ETD projects were in existence, to learn what others saw as the issues, and to learn how others were dealing with the issues. The team met weekly for several months to develop the questions that would be addressed in the questionnaire. The exercise was helpful in itself for clarifying the complex area of electronic theses. We discerned 7 distinct issues: governance issues, submission issues, access & distribution issues, preservation & storage issues, intellectual property issues, and social & philosophical issues. Team members composed the specific questions to fall under each category: governance and submission questions were the domain of the graduate studies office, the librarians focussed on distribution, preservation, and storage issues. Our geography graduate student Martin Bunch had worked on the University's intellectual property committee, and accepted responsibility for articulating the questions for that category. Our team player from the computing department and the computer science graduate student assisted across all categories.

Governance issues included:

- Who is responsible for storage of ETDs for distribution purposes?
- Who is responsible for storage and preservation of canonical copies of ETDs

Submission issues included:

- What format or formats should be accepted?
- Should the author submit the thesis in a format suitable for distribution?
- How should multimedia files be handled?

Access and distribution issues included:

- How should electronic theses be provided and delivered?

Storage and Preservation issues included:

- In what formats should theses be stored?
- Who is responsible for preservation of the canonical copy?

Intellectual Property issues included:

- Will improved access result in copyright violations?

Social and Philosophical issues included:

- Have graduate students and faculty been canvassed for opinions regarding ETDs?
- Have the long-term implications of storing ETDs been considered?

The URL for the questionnaire was posted on electronic discussion groups with an invitation to fill in the form, and send us experiences and opinions. The message was sent to graduate studies groups, library groups, and discussion groups dealing with issues related to academic research and electronic texts.

It was through this questionnaire that we first made contact with the Virginia Tech project. We invited Ed Fox to come to Waterloo to talk on this topic. He accepted with enthusiasm. The response was very good, and furthered stimulated the interest in pursuing this project at Waterloo. Waterloo became an official member of the Networked Digital Library of Theses and Dissertations.

It was also through this questionnaire that we learned of the joint project being conducted at two other Ontario institutions: York University and the University of Toronto. Associations with York and U. of T. expanded to include meetings with David Balatti, of the National Library of Canada. We met regularly for some time, and continue to maintain contact for discussion and sharing of ideas.

The results provided valuable information (April 1997). We learned that there were only a very few projects in existence at the time. But we also learned that this was an area widely regarded as about to develop dramatically. Most respondents saw their institution as moving forward in this area in the near future. And everyone was interested in learning more about the issues. We learned that PDF was the primary formats being used for both submission and distribution of ETDs on the Web. We learned that some projects used a searchable interface of abstract records encoded in HTML. The results raised issues that we'd been unaware of, for example, the concern that ETDs may be considered prepublications. But overall, the results were encouraging. We felt ready to move into an experimental phase, where we could get some practical experience.

Our proposal for a pilot project would be acceptable, and the project itself more productive, if it had a precise and narrow focus (April 1997b). Our survey results showed us that PDF was widely regarded as an effective format for distribution of theses on the web. Our stated overall goal was to determine the feasibility of electronic distribution of theses.

Specifically, we decided to test the feasibility of procedures for providing access to theses originating in electronic form and also theses available only in paper, both typed and photocopied and laser printed. We decided to test multimedia as well. We found a fine art MA thesis that contained a substantial appendix of colour slides, and a geography thesis that included over-sized maps.

With the decision to work with paper, slides, and oversized maps, we recognized the need to expand our project team to include members from the University's graphic services. The team player from the main Graphics office was able to cut bound theses and scan pages into an image file for use by Adobe Acrobat's Capture program. The player from the graphics photography office, with access to an oversize scanner, was able to experiment, and advise re digitizing the slides and maps.

We hired a research assistant to take responsibility for producing a database of abstract records. These records contained the bibliographic information, keywords, and the text of the abstract that served as the searchable portion of the project. Each abstract record contained a link to the full text of the thesis in PDF. The Open Text search engine was applied to the database of abstracts (October 1997b).

With the successful completion of the pilot project, a report was written and recommendations put forward (October 1997c). The team recommended to the Senate Graduate Council that the University proceed towards adopting a policy of accepting storage and providing access to electronic theses. We recommended that the University establish a Web site for a database of theses in Adobe PDF format.

During this period, the University of Waterloo Library began working in a consortium with two neighbouring university libraries, the University of Guelph and Wilfrid Laurier University. The consortium is called the TriUniversity Group of Libraries, or TUG. Discussions regarding electronic theses had been occurring at both Guelph and Laurier, and graduate officers and chief librarians from those schools agreed that the Electronic Thesis Project should proceed as a TUG Project.

The core TUG team consisted of a graduate officer and a librarian from each of the institutions. The new members of the project were brought up to speed. Within a few months, the groundwork was laid for the TUG team to consider the recommendations that came out of the pilot project. The TUG team agreed with the recommendations, and they were revised to incorporate the TUG group as authors.

However, as time went by, the TUG project did not progress. A number of conditions may have been factors in the delay. The geographical distances among the institutions, the fact that the pre-existing consortium was libraries-based, not university-wide, and the fact that the learning curve was not climbed as a team, may all have contributed. But the primary factor is likely the fact that governance procedures varied. Thesis submission regulations vary from institution to institution, and the steps to be taken for revision of these regulations are on disparate time lines. Because the project could not proceed in unison, TUG administrators decided that the project should revert back to a UW initiative.

The TUG collaboration continues in an informal mode. Communications exist among the graduate officers and librarians. As the other schools approach the regulation revision stage, assistance and advice will be available. But until the University of Guelph and Wilfrid Laurier University are prepared to allow electronic submission, they will be unable to contribute to a TUG database of electronic theses. Such a database may appear in the future.

UW has been able to move ahead with the revision of thesis regulations to allow for electronic submission and Web access to theses. A business plan (Jewell) for access to the theses was developed and submission procedures defined. The business plan includes a definition and description of an ETD system, personnel requirements, and cost

estimates. Future considerations include a discussion of the need to keep an eye on changing technology and the development of innovative distribution and storage formats.

The development of the Library's database, however, encountered a number of complications. I'll describe one such complication in order to illustrate the value of ongoing collaboration and good channels of communication.

With procedures in place for submission and access, the Graduate Studies Office was beginning to receive the first several theses officially submitted solely electronically. At this point, one of the Graduate Studies Office staff expressed some concern re the convention that the Library had been using for naming abstract records and PDF files. The filenames for both the abstract records and the full text theses were based on student ID number. The student ID number was being used because it is unique, and allowed the cgi script to pull the number from the abstract submission form automatically. The Graduate Studies staff member who receives the electronic submissions expressed concerns about making a student's ID number available publicly on the Web.

Though no longer meeting on a regular basis, the team came together once again to discuss this latest development, this time however only via email. We concluded that the confidentiality issue was a concern, and an alternative unique file-naming convention was needed. A suggestion incorporating UW's computer user identification code was proposed. All students enrolled at UW are assigned a unique user ID. It is formed from the first one or two initials and the first 5 or 6 characters in the student's last name. The uwuserid is openly available information; it forms the student's email address while he or she is enrolled at UW. It was determined that combining the author's UW computer user ID (uwuserid) and the date field together would automatically derive a unique filename. This solution has been put in place. The abstract submission form (March 2000a) now includes a field for the thesis author's UW computer user identification code.

It is this type of co-operation that has facilitated the success of the project. Success resulted from observations from the perspective of one department, combined with a creative solution and knowledge of technology from another department. The sharing of ideas, the use of a variety of formats for communication, and the energy that comes with learning and working together, are all part of successful collaboration.

From an early beginning as an interesting idea worth developing, E-theses at Waterloo have become a reality (March 2000b). Various forms of collaboration have made the success possible. The team consisted of players from across campus. Individuals from around the world participated through their responses to the ETD Questionnaire. Additional members joined the original team to carry out the pilot project. The potential for ongoing consultation exists for the members of TUG, as well as the University of Toronto and York University, the National Library of Canada, and the Networked Digital Library of Theses and Dissertations. Without this kind of collaboration, the University of Waterloo E-thesis Project could not have succeeded.

References

- Jewell, Christine. (February 1999). *Electronic Thesis and Dissertation System Business plan* [Online]. Available HTTP: http://www.lib.uwaterloo.ca/ETD/business_plan.PDF (11 March, 2000).
- University of Waterloo Electronic Thesis Project Team. (April 1997a). *ETD Questionnaire Results* [Online]. Available HTTP: <http://www.lib.uwaterloo.ca/~uw-etpt/results.html> (11 March, 2000).
- University of Waterloo Electronic Thesis Project Team. (April 1997b). *Pilot project proposal* [Online]. Available HTTP: <http://www.lib.uwaterloo.ca/~uw-etpt/Report/proposal.html> (11 March, 2000).
- University of Waterloo Electronic Thesis Project Team. (October 1997b). *ETD Pilot Project* [Online]. Available HTTP: <http://www.lib.uwaterloo.ca/~uw-etpt/pilot.html> (11 March, 2000).
- University of Waterloo Electronic Thesis Project Team. (October 1997c). *ETD Pilot Project Summary Report* [Online]. Available HTTP: <http://www.lib.uwaterloo.ca/~uw-etpt/summary.html> (11 March, 2000).
- University of Waterloo Electronic Thesis Project Team. (March 2000a). *Abstract Form* [Online]. Available HTTP: <http://www.lib.uwaterloo.ca/ETD/form.html> (11 March, 2000).
- University of Waterloo Electronic Thesis Project Team. (March 2000). *University of Waterloo Electronic Thesis Project* [Online]. Available HTTP: <http://www.lib.uwaterloo.ca/ETD/form.html> (11 March, 2000).