ETD Workflow and Metadata: Best Practices

by Nancy Stuart
University of Victoria

Canadian ETD and Open Repositories Workshop
Carleton University, May 11, 2010

Good morning. We have had some good sessions yesterday on IR platforms and some of the copyright issues that come up with managing ETD collections. This morning we are going to look at the nuts and bolts of the ETD submission process. If you are just starting out, looking at best practices that others have developed, can save a lot of time and angst.

I got involved in ETDs at UVic back in 2004 when we realized that space to store the bound volumes was getting very tight and ETDs were moving into mainstream. After attending an ETD conference in 2004, we set up a pilot in 2005, worked on implementation procedures and processes through 2006 and launched our ETD program in June of 2007. In 2009 I had the opportunity to do a 6 month study leave at Library and Archives Canada with Sharon Reeves in the Canada Theses Office. Over those six months I likely talked to you or your institution about your ETD program or plans for one. I know back in 2005 there were some resources for ETDs but I didn't always know all the right questions to ask. So we thought at a workshop such as this, a Best Practices session might be very useful.

Overview

- What is an ETD program?
- ETD workflow
- Issues to decide before implementation
- ETD website
- Tutorials
- Quality control
- OPAC records
- Harvesting issues
- Sending ETDs to ProQuest
- Restrictions
- Copyright
- Preservation

2   May 10, 2010

This morning in our hour and a half we will cover the following topics. There are likely many more things we could talk about, but at the end hopefully there will be some time for comments and questions. Also feel free to ask questions along the way, but we may refer comments to the end in order to stay on time.

So we will look at:

What is involved in an ETD program

What is involved in the ETD workflow

Issues that should be decided before implementation

Creating an ETD website

ETD tutorials are very helpful

We will look at the value of doing quality control and the pros and cons of metadata editing

Do you want records in your library catalogue

We will look at harvesting issues

Issues around sending your theses and/or dissertations to ProQuest

Do you need a restricted ETD collection

We will look at copyright licenses for ETDs and also preservation issues.

What is an ETD Program?

- "How to guide" on Theses Canada Portal
- Best practices
  - Preparation ETD Guide
  - Collaborate with FGS, Library, IT, & others
  - Design a plan and pilot project
  - Revise Theses guidelines to reflect ETDs
  - Contact LAC to harvest your ETDs

3   May 10, 2010

An ETD program is a program where graduate students save their thesis or dissertation as an electronic file and submit it to an ETD Collection in the university's digital archive. The How to Guide on the Theses Canada Portal gives step by step guidance for setting up an ETD program. The NDLTD website is a wealth of information and excellent preparation. There is an ETD Guide that outlines many of the steps listed here.

After you have done your preparation reading, or possibly attended an ETD Conference, you will want to gather together the people on your campus who should be involved in an ETD program. Normally Faculty of Graduate Studies, Library and IT personnel. It will vary with each institution. Try to find people who will champion the program.

Next you will need to design a plan and develop a pilot project.

Before implementation, your current Theses guidelines will require a complete overhaul to convert to the ETD program. This step can take a lot longer than planned for.

Once you have launched your ETD program, you should contact Theses Canada to request harvesting. We will talk more about harvesting and your technical questions can be answered in the IR Systems Best Practices for ETDs this afternoon.

ETD Workflow Diagram

- Visually see the steps
- Many variations
- Every institution will be different
- Open access / Self-publishing
- Uvic
- Waterloo

4   May 10, 2010

Having a workflow diagram can be very helpful in understanding the all the steps in the ETD process. Each institution will have different steps. Some begin training in the ETD process early in the thesis writing process. Others have a more traditional approach with converting to a PDF and uploading to the IR as the end part of the writing process. One of the benefits of ETDs is to give students an opportunity to be part of the open access movement and the self-publishing process. The grad student should be involved as much as possible in the process and have a good understanding of what they are doing.
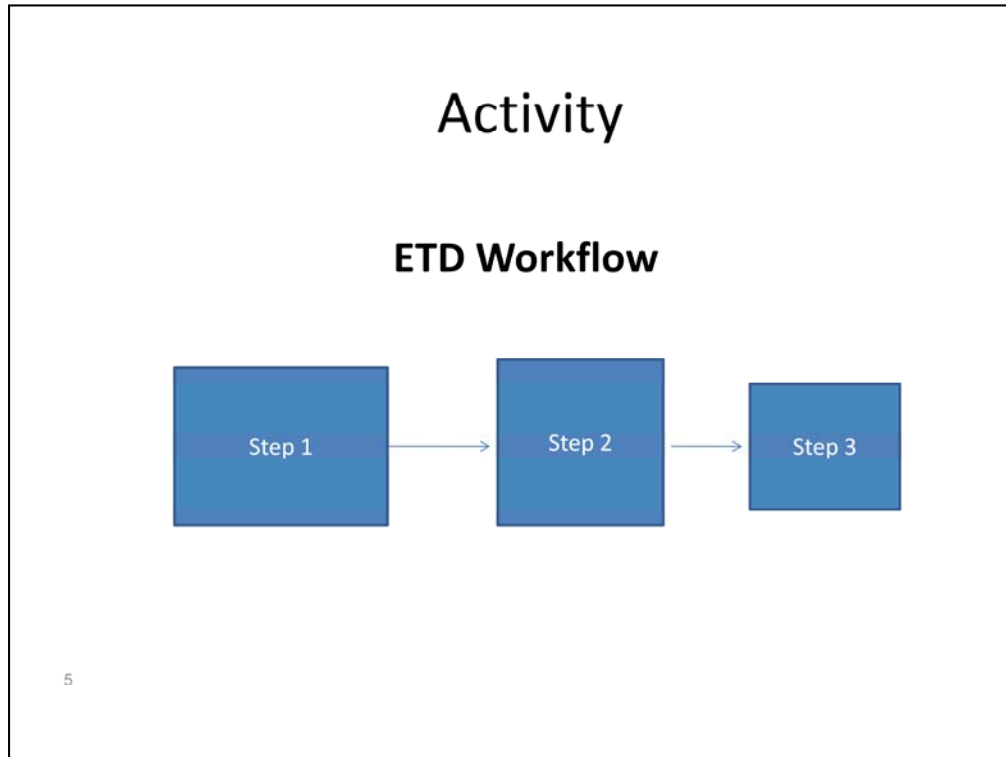
Idea – Let's work together in small groups to come up with as many steps as necessary to create an ETD Workflow. Keep it IR neutral if possible. The idea is to talk about the overall model and share ideas on which steps are crucial, important or just nice to have. Let's take 10-15 min. and then a couple of groups can share their workflow. The idea is to identify all the steps before you begin your ETD program.

The point here is to get participants to exchange ideas about best practices and to consider elements they may or may not have thought about in their implementation. They can also address issues revolving around what works well and what doesn't. A few groups can then share their models with the larger audience and we can move on to the next slide to provide further information. I think that this would be a great opening activity to get them involved early on. I can see this activity easily taking up approx. 20 mins.

ETD Workflow

- Write thesis or dissertation using a template (Word or LaTex)
- Obtain PDF training
- Consult with thesis supervisor re format and content
- Submit final version to supervisor
- Defend thesis
- Make any revisions
- Submit final version to supervisor / or GSO for approval
- Convert final version to PDF
- Hand in PDF on CD-ROM to Graduate Studies office
- **Upload and archive metadata and PDF in the IR**

6   May 10, 2010

Briefly here are the basic steps, regardless of what IR software is used.

1.Student writes the thesis. Many institutions have available thesis templates in MSWord, or LaTex. A template is an excellent way to ensure uniformity and adherence to guidelines.

2.At some point along the way, the student may need help with creating the PDF file. More about that later.

3.The thesis development is under the guidance of the supervisor. Awareness of the ETD process is key for supervisors. This ETD education should be done through the Faculty of Graduate Studies.

4.The supervisor is responsible for content of thesis. The student takes charge of format by following the guidelines.

5.After defense the student makes required corrections which the supervisor or GSO checks. Proof-reading is very important at this step, because once finalized, that is it. No changes.

6.Student converts document or documents and supplementary files to PDF. Again this may happen earlier in the process.

7.Student then hands in forms and possibly extra copies to Grad Studies Office.

8.Student then uploads ETD and it is archived in the IR.

9.Student is basically finished.

ETD Workflow cont'd

- Harvested by
  - Google,
  - CARL Harvester
  - LAC Harvester
    - Harvested by NDLTD
    - Automatic creation of Amicus catalogue record
- Automatic creation of local catalogue record
- Send ETD to ProQuest
- Microform copy to institution and LAC

7   May 10, 2010

There are still some processes after the student has submitted that are Library or Grad Studies workflows.

1.Since your IR is open access, think about where you would like your ETDs to be discovered.

2.All IRs have an OAI-PMH layer to allow for harvesting of the metadata. You will want to go beyond just a Google/Internet harvest. You will need to register with a harvester, to be harvested. One such harvester is the CARL IR harvester. Another is Library and Archives Canada. As you have heard, LAC is building a digital theses collection and is already close to 100,000 ETDs in its collection. Its aim is to collect all current ETDs produced at Canadian universities through its harvesting service. The LAC collection is then in turn harvested by the NDLTD harvester.

3.Another activity is the automatic creation of catalogue records through metadata mapping. LAC displays all theses records in its public catalogue AMICUS, by mapping the metadata from your IR into a MARC record. You can also create MARC records for your local collection from the metadata in your IR by using DC – MARC crosswalks.

4.Many institutions send a copy of their ETD to ProQuest in order to get a microform perservation copy. More about this later in the presentation.

University of Victoria
ETD Workflow Diagram
June 2009
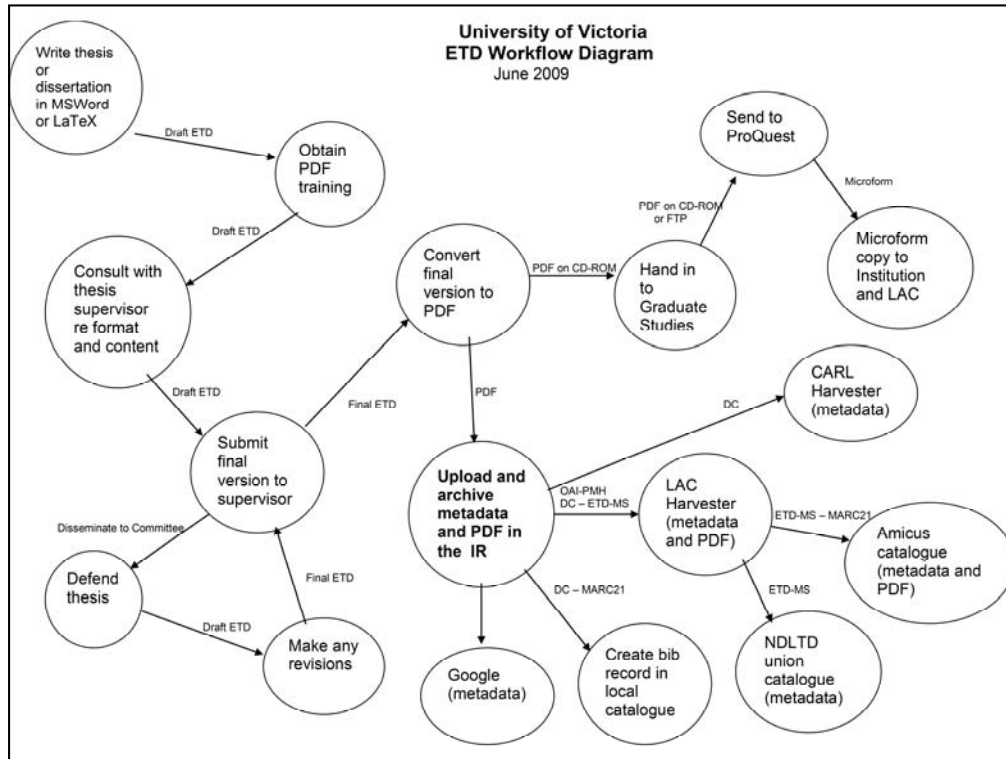
Here is the workflow we used when starting at UVic. The central circle could also be broken down into a more discrete workflow. At Uvic, Grad Admissions approve the preliminary pages of the thesis, or checks that it meets the format standard, outside the IR workflow. This summer we are moving to change this so that we use the IR software Approve / Reject workflow.
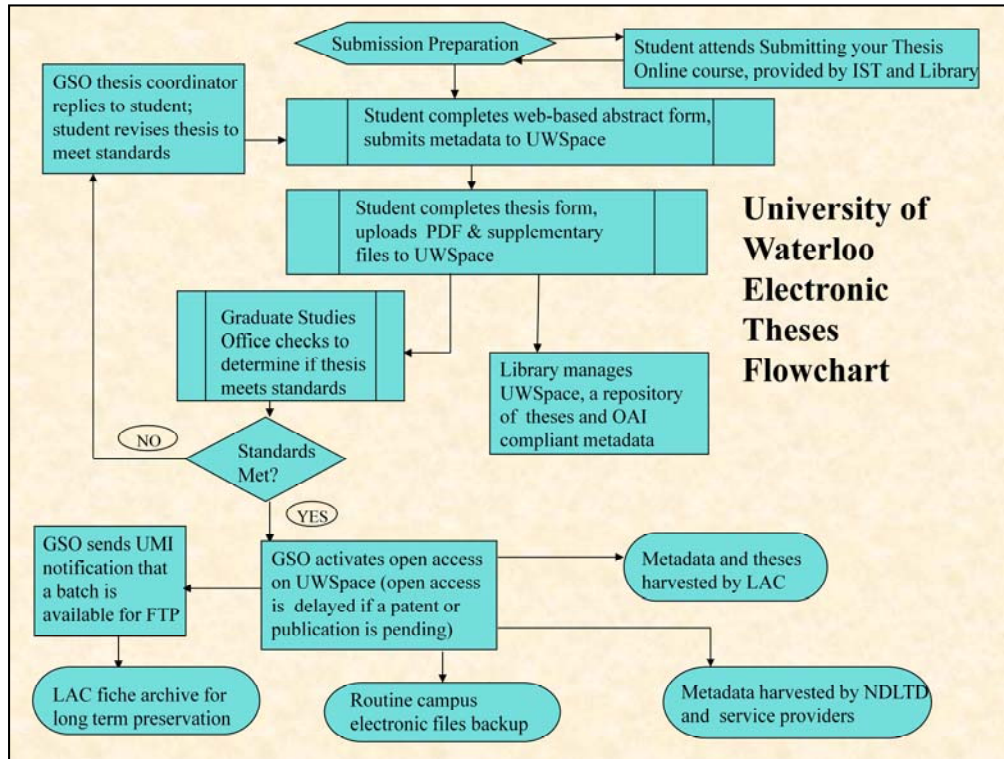
At University of Waterloo their workflow is slightly different because they use the built in workflow in their IR. The student uploads their thesis for approval within the IR software workflow. It is approved or rejected and sent back for revisions all within the IR workflow. The advantage is that the approver can see the entire thesis, not just the preliminary pages.
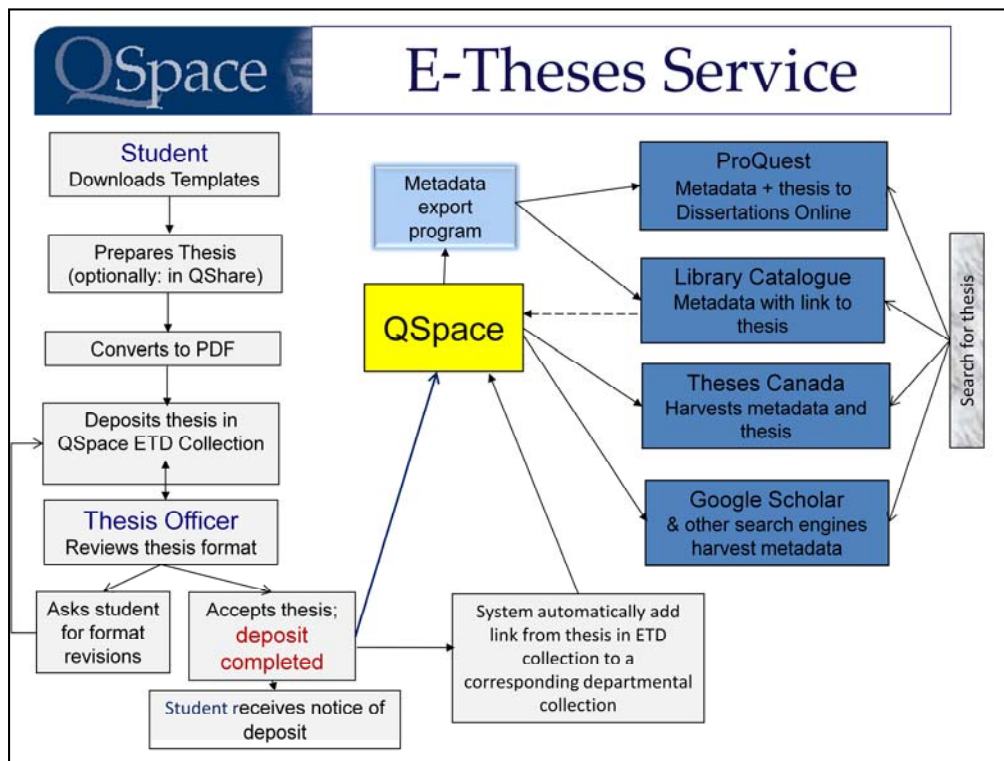
Here is Queen's ETD workflow that we saw yesterday in Sam's presentation.   Both Waterloo and Queen's FTP their ETDs to ProQuest whereas at UVic, we send ProQuest a CD-ROM with a PDF on it. At UVic, instead of handing in any paper theses, the student hands in a CD-ROM with their PDF on it. Both processes are acceptable.

Best practice for ETD workflow would be to use the IR for the Approve /Reject step. Of course this is easier said than done, as we all have to work within the framework of existing workflows and change is never easy.

Issues to decide before implementation – overview

- Institutional repository
- File formats
- File naming conventions
- Mandatory submission
- Metadata fields
- ETD collection sets

11

May 10, 2010

1. We have been talking about ETD workflow and process, but all of this is with the assumption that you have a digital archive or Institutional Repository in which to store the ETD collection. This was discussed yesterday, so the rest of the presentation will attempt to be IR neutral.

2. The other issues to think about are:

3. File formats

4. File naming conventions

5. Whether submission should be voluntary or mandatory

6. Metadata fields

7. And ETD collection sets

So let's look at these one by one.

File formats

- EDTs expand beyond print medium
- Many file formats acceptable: .pdf, .xml, .xls, .ppt, .jpg, .mpeg, .gif, .wav, .mov,
- PDF:
  - CutePDF Writer;
  - MSWord – Adobe tab or print to Adobe;
  - Adobe PDF/A; - Who requires this?
  - Adobe Acrobat 9 with PDF e-portfolio

12   May 10, 2010

1. One of the benefits of ETDs is that it expands beyond the print medium and allows students more options of expression. This of course means that their ETD may include multi-media files or other supplementary files. These file formats are normally acceptable in an IR.

2. The standard file format for ETDs is the PDF. Students can create a PDF by using a free tool, such as Cute PDF writer. If your institution subscribes to Adobe, students can create their PDF using the Adobe Tool bar on MSWord by printing to a PDF file. What students do will depend on what you require them to do. If you require it in your Guidelines, then it will happen.

3. As I said, the standard file format for ETDs is PDF. But you can go beyond PDF to PDF/A.
PDF/A is an International Standards Organization (ISO) standard for long-term preservation of electronic documents.
It offers assurance that documents archived in that format will maintain their appearance and readability regardless of which applications and systems were used to create them.
Best practice would be to store your ETD PDFs as PDF/A.

4. **Question:** How many of you require that format? How do you make this a requirement for your students and do you check that it is in that format?
The PDF/A standard does not define an archiving strategy, but instead identifies a format for electronic documents that ensures they can be reproduced consistently and predictably in the exact same way well into the future. PDF/A offers users a way to preserve electronic documents in a manner that maintains their visual appearance over time, independent of the tools and systems used for creating, storing or rendering the files.
Based on portable document format (PDF) technology from Adobe Systems Inc., PDF/A eliminates PDF features not suited to long-term archival, such as audio, video and transparency. A vital component to PDF/A reproducibility is that the documents must be 100 percent self-contained. This feature ensures that all information needed to display the document in the same manner every time is embedded in the file. Included in each PDF/A file is all content, including text, raster images and vector graphics; as well as fonts and color information.

5. **BUT** at the moment, LAC only has the ability to harvest a single PDF file. So how can we accommodate this restriction and still archive the entire ETD. Adobe has come out with an e-portfolio which is like a wrapper to hold all the files in one PDF. This may be a way to accommodate having only one PDF file per ETD. More about this later.

File naming conventions

- Good examples
  - ubc_2009_fall_ferch_elizabeth.pdf - (UBC)
  - Beauchemin_Stéphanie_2008_mémoire.pdf - (Montréal)
- Bad examples
  - Roxanne_Paul_Dec06Master's_Thesis_Final_for_UVIC_D Space_19Aug07.pdf - (UVic)
  - Thesis.pdf - (UVic)
- Part of theses guidelines
  - Lastname_Firstname_YYYY_mémoire.pdf
 - YYYY_Lastname_Firstname_Degreename.pdf

13   May 10, 2010

Before you start accepting ETDs, write into the Theses Guidelines your naming convention for the ETD PDF file. Here are some good and bad examples. If you ever want to import or export batches of files, consistent naming will help immensely. Also these urls are displayed on websites and a consistent, meaningful file name is more readable.

Having a File naming convention embedded in your ETD workflow and Guidelines gives you the ability to enforce it.

Mandatory submission

ETDs added by Year
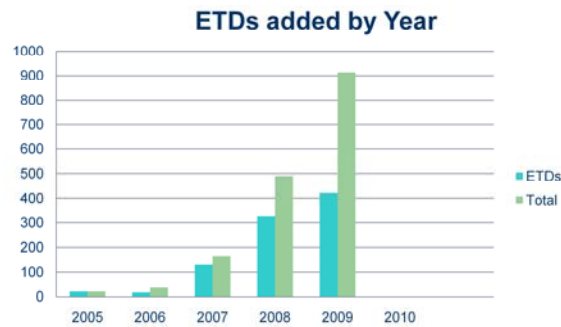
14

May 10, 2010

Most institutions begin an ETD program with voluntary submission by students. At some point they decide to mandate it, in order to collect all of their theses and dissertations in electronic format. Currently some institutions are mandating electronic submission right from the beginning. With voluntary submission, buy-in can be slow, but when it is mandated by Graduate Studies or the institution, the ETD collection grows very quickly. We have just mandated student submission for all theses this year, but have at least been receiving the PDF file, so have had staff submitting them to the ETD collection. If we only received print, we digitized it and submitted it. We have just caught up so from now all, all will be student submitted. Just start mandatory from the beginning.

Metadata fields

- Consider where the metadata will display
  - The digital archive
  - Your library catalogue
  - AMICUS
  - Carl harvester
  - Google / GoogleScholar
  - NDLTD ETD union catalogue

15

May 10, 2010

In the next 4 slides we will talk about metadata fields. First we need to consider what metadata fields to use, and you may wish to consider where the metadata will be displayed. If your metadata is going to be harvested, you may make some fields mandatory. Fields like Author, Title, Date, Type and Degree. If you are mapping your metadata to your catalogue you may require some additional fields to keep your catalogue records consistent, such as Department and Supervisor. Amicus, the LAC catalogue, will require some special fields for the degree information. LAC uses the metadata schema ETD-MS which is based on Dublin Core. It can be found on the NDLTD website and being familiar with it before starting your ETD program is very helpful.

Metadata fields, cont'd

- Mandatory
  - Author
  - Title
  - Date
  - Identifier – persistent is best
  - Degree.name – controlled list for your institution
  - Type – Thesis; Electronic Thesis or Dissertation
  - Language – ISO list

16

May 10, 2010

Here are some fields that you may want to consider as mandatory. If they are not filled in on the submission form, an error message appears to alert the user to fill in the field. The Identifier field would be system generated, but still mandatory for an ETD record.

Metadata fields, cont'd

- Optional
  - Abstract
  - Keywords or subjects
  - Department – controlled list
  - Degree.discipline – controlled list
  - Degree.level
  - Degree.grantor
  - Rights
  - Supervisor
  - Publisher
  - Others your institution may want

17   May 10, 2010

Here are the optional fields that you would want in an ETD metadata record. Best practice would be to have as many fields as possible under controlled vocabulary drop down boxes, so the student chooses from a list. This provides consistency and a user friendly form.

You can create these controlled lists applicable to your institution.

Question: Is anyone using controlled lists for Supervisors names? Possibly from your Faculty lists. Would be helpful if your records go into your catalogue where names are under Authority Control. What other controlled lists are you using? Degree; Dept. / Faculty

Some institutions record defense date or convocation date. These are all local decisions.

Metadata fields, cont'd

- Metadata schemas
    - Dublin Core qualified
    - ETD-MS: DC with degree element
      http://www.ndltd.org/standards/metadata/etd-ms-v1.00-rev2.html
    - MODS – expressed in XML
      http://www.loc.gov/standards/mods/
- How this works with Marc and RDA
    - Require crosswalks to map the metadata

18   May 10, 2010

Before you begin you should have some understanding of Dublin Core and ETD-MS metadata schemas. These are two common schemas used for ETDs. A number of IRs come with the Dublin Core qualified schema. You can add other schemas if desired. If you are being harvested by LAC, you will need to use ETD-MS. More about this in the afternoon IT sessions. ETD-MS is basically Dublin Core with 4 qualified Degree elements.

Another common schema used for ETDs is MODS.

Question: Is anyone using MODS? You can get more information on MODS on the Library of Congress website.

Another thing to consider before implementation is if you will need crosswalks to map your metadata to MARC for use in your catalogue. It is good to read up on these topics before implementation, so you are aware of issues that may arise.

ETD Collection sets

- Currently submitted ETDs in one collection set
- Digitized backfiles in a separate set
- Restricted ETDs in a separate set

19   May 10, 2010

Before implementation it is important to think about how you will organize your ETD collection or collections. Many libraries are now having their back files of theses digitized, either in house or out-sourcing them. Often these backfiles are acquired from ProQuest. Library and Archives Canada now have the ProQuest PDFs back to 1997 and up to 2004. If you have these same files you should keep them in a separate collection from the ETD Collection of current theses that you want harvested by LAC in order to avoid duplication in the Amicus catalogue. LAC also must not harvest any of your restricted theses, so these should be kept in a separate collection.

Once you have considered these issues and how you are going to deal with them, you are ready to move on to setting up your submission forms in your IR, creating an ETD website and re-writing your Thesis Guidelines.

Even though there is a lot of prep work, once you launch your pilot, or ETD program, you are almost done.

ETD Website

- A place for all pertinent ETD info
  - ETD procedures
  - How to submit your ETD
  - FAQs
  - ETD blog
  - ETD templates: MSWord, LaTeX
  - Graduate Studies guidelines for Theses
  - Forms – LAC license, ProQuest form
  - Link to your digital archive
  - Tutorials, workshops, computer support

20

May 10, 2010

So let's look at the ETD website. Your ETD website should gather together all the pertinent information needed by students and faculty for the creation of ETDs. There are a number of excellent examples of ETD websites on the next slide. You can put the overall ETD process here with submission procedures, templates, links to tutorials, FAQs, an ETD blog and FGS guidelines and forms. Ideally this is created in collaboration with all the parties involved in ETDs on your campus.

ETD IRs and ETD websites

Institution	IR Software	IR URL	Website URL
Acadia	ContentDM	http://library.acadiau.ca/ContentDM/theses.html#	http://library.acadiau.ca/guides/etheses.html
Alberta	DSpace	http://repository.library.ualberta.ca/dspace/	http://www.gradstudies.ualberta.ca/degreesuperv/thesis/submissionethesis.htm
Laval	Archimede	http://archimede.bibl.ulaval.ca/archimede/	http://www.fes.ulaval.ca/sgc/mte/
Manitoba	DSpace	http://mspace.lib.umanitoba.ca/	http://umanitoba.ca/libraries/elibrary/mspace/etd_index.html
McGill	DigiTool	http://www.mcgill.ca/library-findinfo/escholarship/	http://www.mcgill.ca/gps/students/thesis/programs/ethesis/
Queen's	DSpace	https://qspace.library.queensu.ca/	http://library.queensu.ca/webir/e-theses/sgs-index.html
Sask	DSpace	http://ecommons.usask.ca/	http://grad.usask.ca/etd/
Toronto	DSpace	https://tspace.library.utoronto.ca/	http://www.sgs.utoronto.ca/informationfor/students/finup/etds.htm
UBC	DSpace	https://circle.ubc.ca/	http://www.grad.ubc.ca/students/thesis/index.asp?menu=007,002,000,000
Victoria	DSpace	https://dspace.library.uvic.ca:8443/dspace/	http://library.uvic.ca/site/lib/dig/uvthesis/index.html
Waterloo	DSpace	http://uwspace.uwaterloo.ca/	http://www.lib.uwaterloo.ca/ETD/index.html

These institutions have created ETD websites for their students. I put this chart together last summer while I was on study leave. I am sure there are other institutions that have come on board since then. But, feel free to browse through for ideas and best practices. Don't re-invent the wheel.

Tutorials

- Training
  - students
  - supervisors
  - grad secretaries
  - Graduate Office and Library staff
- Education and promotion

22   May 10, 2010

Self-help guides for students are much appreciated and a number of institutions have done these. Queens has given us some very good information on tutorials. Unfortunately they can get out of date as the IR is upgraded or the ETD collection moves to another platform.

Training should not be overlooked and it is a place that you can collaborate with other folks on campus.

Remember that as you implement an ETD program, many new processes will have to be put in place and staff and students will need training.

Use your existing communication channels to promote and educate students, faculty and staff on the new developments as you move to ETDs.

Quality Control

- ETDs and quality control
  - Approval – final revisions
  - Metadata – spelling, caps, fields
  - Signatures
  - Files – supplementary
  - Identifiers – persistent ID
  - Servers – if changed must forward forever; can't change links in Amicus

23   May 10, 2010

There still needs to be some quality control around ETDs and the ETD process. Just as in the print world, the final corrected version needs to be approved, usually by the Faculty of Graduate Studies and/or the supervisor. This needs to be included in the workflow of the ETD. Once the metadata has been entered by the student, the metadata should be checked by library staff for consistency and typos.

Question: What are the pros and cons of metadata editing? How many are doing it?

Also there can be **NO** signatures in the ETD for privacy reasons. In an ETD there is no formal signed approval page as there was in print.   If there are supplementary files or multiple files, these should be checked. As we have seen the use of the Adobe e-portfolio will wrap up these multiple files into one files so that LAC can harvest its single PDF file. Some institutions use the handle.net system for their persistent url. If the IR moves to another server the url address to the file doesn't change. Because LAC harvests not only the metadata, but also the PDF, links to servers that move should be permanently forwarded, because the harvested link in the Amicus catalogue can't be changed or updated.

OPAC records

- Want or need?
- Manual or automatic
- URL links to the digital archive
- Authority control
- Discovery layer

24

May 10, 2010

Most institutions have traditionally catalogued their theses and dissertations, which were then available in their library catalogues. Now with IRs, is it necessary to have records in two places? How do you maintain two places for metadata? This is clearly a want rather than a need. The easiest way is to install a crosswalk to convert the metadata in your IR to MARC and load the MARC file into your catalogue. This can be done automatically with little or no human intervention. You would want links to the ETD in your catalogue, just as you load e-book collection MARC records into your catalogue. You can maintain authority control in your catalogue which is much harder, if not impossible, in your IR.

Of course a discovery layer such as Endeca, Summons or Primo could find your ETDs, most libraries are not there yet.

Harvesting issues

- LAC can only harvest 1 PDF file
- Metadata harvesting requirements
- Metadata must be in ETD-MS
- Dublin Core qualified must be cross-walked to ETD-MS
- Issues
  - Date
  - Degree elements

25

May 10, 2010

Much more about the harvesting issues will be in the session after lunch. Before you get your ETD program launched, the best practice is to have your DC metadata mapping to the correct ETD-MS field. This is one of the most problematic areas for LAC. Displays in the Amicus catalogue are based on this mapping.

Two problematic areas

Date – this should be the copyright date, not the submitted date to the IR.

Thesis note field – Thesis (M.A.) – Institution granting degree, date (year from t.p.) If metadata fields not mapped correctly, the degree name doesn't show, or the institution doesn't show or the date is the date submitted to the IR, not date from t.p.

Sending ETDs to ProQuest

- Send ETD to ProQuest
  - On CD-ROM
  - Via FTP
- Send LAC licenses and ProQuest form to ProQuest via mail.
- Send LAC licenses to LAC if not going to ProQuest.

26   May 10, 2010

First some history:

In the print world we sent our theses and dissertations to ProQuest (formerly UMI) for two reasons.

1. To have the thesis listed in Dissertations Abstracts and

2. To have a microform copy for preservation and ILL. One copy for your institution and the preservation copy and 2 ILL copies for   LAC

Currently, LAC negotiates a contract on our behalf with ProQuest every three years, and that is why LAC has a license the students sign.

ProQuest began digitizing theses in late 1997, under the terms of its first contract with Library and Archives Canada.

To date LAC has received digitized theses and dissertations from ProQuest covering the period 1997 to mid 2004.

LAC has a goal of creating a digital Canadian theses collection by the middle of this decade. This will be accomplished by harvesting our ETD collections in our Canadian institutional repositories. This is why it would be great if more institutions could begin ETD programs.

Question: Are there continuing benefits to sending our ETDs to ProQuest? Discussion.

Restrictions

- Available to World Wide Web
- Available to campus only
- Withheld
  - Reasons
  - Period
  - How to withhold

27

May 10, 2010

Of course one of the big benefits of ETDs is their availability on the web. But, there are some who don't want their thesis to be that available for valid reasons.

If an ETD needs to be restricted (withheld), there must be valid reasons and a documented workflow. Every campus will be different, but the traditional reasons for withholding are patents pending and imminent publication of the work. Some institutions have 6 month withholding blocks which can be renewed, others have 1 or 2 years. The important thing to remember here is that these theses must **NOT** be harvested. For some IRs, the easiest way to deal with this is to put the withholds in a separate restricted collection. Dspace 1.6 has an embargo feature which may be a slick way to deal with this issue.

Some institutions want to restrict access to campus only. This could be analogous to a print thesis in Special Collections or Archives, which are usually non-circulating.

Lately there have been conversations on the ETD-I and on campuses about Creative Writing students who are creating original novels, plays and poetry that they want to publish.

Question: How are institutions dealing with this?

Copyright

- Authors
  - Your institution's partial copyright license
  - Creative commons license
  - LAC partial copyright license
- Permission
  - pictures, maps, articles, etc. in the thesis
  - digitized back files of theses

28   May 10, 2010

This is one area where we get many questions. The first is author's rights. Much of this was covered yesterday in the copyright session.

For ETDs your institution's partial copyright license will need to cover distribution on the web. You may have to have your university lawyer draft you a more up-to-date version to cover ETDs. You can have your students choose the CC license to cover their digital rights. All students should be signing the LAC partial copyright license as this allows LAC to make it available on the web.

The other piece of copyright is when the student needs to obtain permission to use something in their thesis. This is traditionally where the supervisor should give guidance. Often the student has asked for permission, may put the letter of permission in their PDF, but the fine print says that they may use the item for the defense of their thesis, but it may not be included in an online version. So it must be removed from the PDF before uploading into the IR. This is where having the approval workflow within the IR software is so important. That way the approver can see all the pages of the PDF and check for this type of thing.

Many universities are digitizing older theses or having them digitized.   Some are keeping these in a restricted collection because they do not have permission from the author. Others put them in collections available to the web and invite people to contact them if they don't want them available that way.

Preservation / Archiving

- Your Digital Archive
- Your Digital Archive backup
- LAC – Trusted Digital Repository
- LOCKSS – U. of Sask

29   May 10, 2010

The final aspect of an ETD program is of course preservation. The session right after the break will deal with this topic in more depth. I just want to mention a couple of points. In the print world, we considered the microfiche to be the preservation copy. There was a copy on your campus, the original and two copies at LAC and a copy at ProQuest. That is pretty good back up.

Many faculty are skeptical of digital files being preserved over the long-haul. This is still a developing field, but let's look at the copies we have in the digital form. Your own copy in your IR. Of course your IR is backed-up both locally and off-site. There is a digital copy that LAC harvests. We will hear in the next session about LAC's plans for a Trusted Digital Repository. Some places are using LOCKSS. So the ETDs shouldn't disappear.

Preservation is different than many copies. PDF does seem to becoming a global standard with many interested parties discussing this issue. I will let Gail and Pam in the next session explore this further.

Work in progress

- How can the Canadian ETD community share best practices?
  - List-serv
  - Blog
  - Wiki
  - Theses Canada Portal

30   May 10, 2010

Best practices around ETDs is evolving and developing. We need to share best practices so we don't re-invent the wheel and we have consistently good practices. We are open to suggestions as to the best way for Canadian institutions to share knowledge and mentor institutions just coming on board with ETDs.

Question: Does anyone have ideas for sharing best practices? Any other questions?