The Lifecycle Management of ETDs Project:
Multi Stakeholders National Partnership

Shannon Stark
University of North Texas
Shannon.stark@unt.edu

Daniel G. Alemneh
University of North Texas
Daniel.alemneh@unt.edu

Stephen Eisenhauer
University of North Texas
Stephen.eisenhauer@unt.edu

Abstract

The transition from traditional paper and microfilm formats to electronic theses and dissertations presents a number of significant challenges for academic libraries. To address these challenges, the UNT Libraries, together with their partners, are working on a collaborative project sponsored by an Institute of Museum and Library Services National Leadership grant. This poster describes the project objectives, possible outcomes, and proposed deliverables including a toolkit of guidelines, educational materials, and software tools.

Keywords: digital curation, ETD, lifecycle management, digital preservation

Introduction

In 2011, the Institute of Museum and Library Services awarded about a quarter million dollars to nine institutions and libraries to work collaboratively towards lifecycle management of Electronic Theses and Dissertations (ETDs). With virtually all theses and dissertations now being created as digital products, lifecycle management of these valuable intellectual documents is an important new responsibility of academic libraries. But it is a responsibility that unfortunately not all libraries have the infrastructure or resources to support.

The long-term management of ETDs requires not just the curation of electronic documents, but demands that consideration be given to policy creation, copyright, embargoes, formatting, metadata, and more. The big question is: How will institutions address the entire lifecycle of ETDs, ensuring that the papers acquired from students today will be available to future researchers?

This two year project is the answer these nine institutions have devised. They will promote best curatorial practices to increase the capacity of academic libraries to reliably preserve ETDs by using a three-pronged approach that includes writing guidance documents, developing software tools, and creating educational materials—all of which will be available under an open access license at the conclusion of the project.

Participating ETD Curators

The following partners joined together to accomplish the stated objectives.

- University of North Texas (lead institution):
  - Dr. Martin Halbert, *Principal Investigator*
  - Shannon Stark, *Assistant Coordinator*
  - Dr. Daniel Gelaw Alemneh, *Project Steering Committee Chair*

_____

- Educopia
  - Dr. Katherine Skinner, *Co-Principal Investigator*
- MetaArchive
  - Matt Schultz, *Project Manager*
- Virginia Tech and The Networked Digital Library of Theses and Dissertations (NDLTD)
  - Gail McMillan, *Project Steering Committee Chair*
- Boston College:
  - Dr. Bill Donovan, *Project Steering Committee Chair*
- Rice University
  - Geneva Henry, *Project Steering Committee Chair*
- Indiana State University
  - Xiaocan (Lucy) Wang, *Project Steering Committee Chair*
- Penn State University
  - Dr. Patricia Hswe, *Project Steering Committee Chair*
- University of Arizona
  - Yan Han, *Project Steering Committee Chair*

Timeline for Deliverables and Workshop

The project began with the development phase in December 2011, with all members of the steering committee working virtually from their respective locations. During this first year, the focus has been on creating the eight guidance documents and micro-services, which will be completed by the end of 2012. In the second year, feedback on the deliverables will be gathered, and the workshop will take place. A timeline of these events can be seen in the following figure (Figure 1).
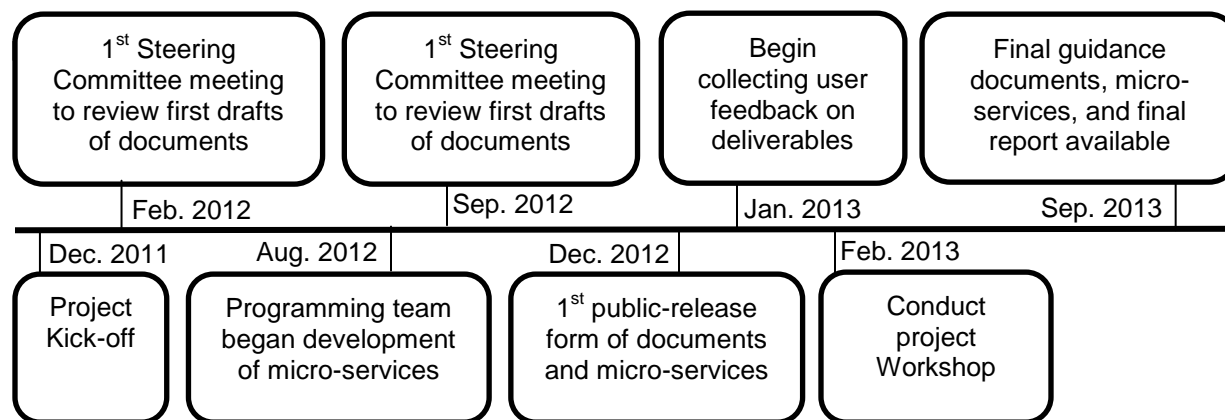

Figure 1. Timeline of the project

Project's Outcomes and Deliverables

A needs assessment process undertaken by the planning group for this project identified a cluster of informational resources required to advance the acquisition, organization, use, management, presentation, and preservation of ETDs. To address these needs, guidance documents are currently being written by experts in each specific area, based on their research and experiences.

Descriptions of Guidance Documents

Briefing on Access Levels and Embargoes of ETDs

One of the most contested topics in ETD program planning is the question of ETD embargoes and levels of access restriction. This document will directly address this difficult issue, which there is

currently no consensus on regarding "good practices" in the field. This briefing will concisely lay out the issues and the different sides of this argument for those considering an ETD program, clearly articulating the pros and cons claimed by proponents for and against embargoes.

Briefing on ETD Copyright Issues and Fair Use

There is so much uncertainty and fear over the issue of copyright and fair use that it has in some cases prevented campuses from implementing ETD programs at all. This briefing will address publisher concerns and issues squarely. Again, the point of this briefing is to provide a concise summary of the relevant information for stakeholders, with documented citations.

Guidelines for Implementing ETD Programs - Roles & Responsibilities

This document will provide guidance for understanding roles and responsibilities, together with frank acknowledgements of the concerns that different stakeholders may have, and how ETD programs have addressed these issues as well as the ramifications of *not* specifying roles and responsibilities.

Guidelines for Collecting Usage Metrics & Demonstrations of Value for ETD Programs

This document will provide guidance for institutions concerning the critical issue of assessment of ETD usage, and how communicating such assessment metrics can demonstrate the benefits of the program to stakeholders. Practical examples of how to document and convey usage metrics will be provided.

Overview of Formats, Complex Content Objects, and Format Migration Scenarios for ETDs

Additional issues of great concern to many current and prospective ETD programs are how to determine what formats they should accept, and how to manage on an ongoing basis the increasingly complex ETDs that are now being produced by students. This document will seek to set forth the relevant information and issues, and to provide scenarios for managing and migrating complex ETDs over their extended lifecycle.

Overview of PREMIS Metadata & Lifecycle Event Record-Keeping for ETDs

Another issue revealed in the needs assessment process was that most institutions do not have workflows and systems in place to capture the appropriate levels of metadata needed to manage ETDs over their lifecycle, often because of a lack of awareness of the ramifications of not maintaining such information. An overview of these critical issues will be provided, not just at the point of ingestion, but subsequently, as ETDs often have transitional events in their lifecycle (embargo releases, redactions, etc.).

Guide to ETD Program Cost Estimation and Planning

One of the most frequently asked questions identified in the CNI and MetaArchive surveys pertains to the costs and process for setting up an ETD program. This guide will provide succinct information on costs and planning, laying out the critical paths that many ETD programs have charted, together with advice about strategies.

Guide to Options for ETD Programs

Many institutions are delayed in ETD program planning simply because they do not have a clear understanding of the range of options to consider in implementing an ETD program. Like the other documents in this series, this guide will be a descriptive (rather than prescriptive) information resource

that seeks to explain the relevant decisions institutions must make and to clarify the pros and cons of different options. This document will reference and integrate all the other documents in the series.

Descriptions of Tools

The use of software tools to analyze ETD content in various ways is an essential part of the repository workflow. These tools help analyze the formats of digital files, record information about their formatting and validation, and scan them for computer viruses. One of the goals of this project is to develop a common protocol for allowing any institutional repository to use these tools effectively and with minimal effort.

Despite the usefulness of these tools, integrating them into an existing institutional workflow or ETD management system can be challenging and require large amounts of code written specifically to interact with each individual tool. We hope to simplify this challenge by creating a specification for scripts that can bridge the gap by handling the direct interactions with these tools, and making their output usable in a clean, uniform manner (Figure 2). This will greatly reduce the development burden of institutions wishing to integrate these tools.
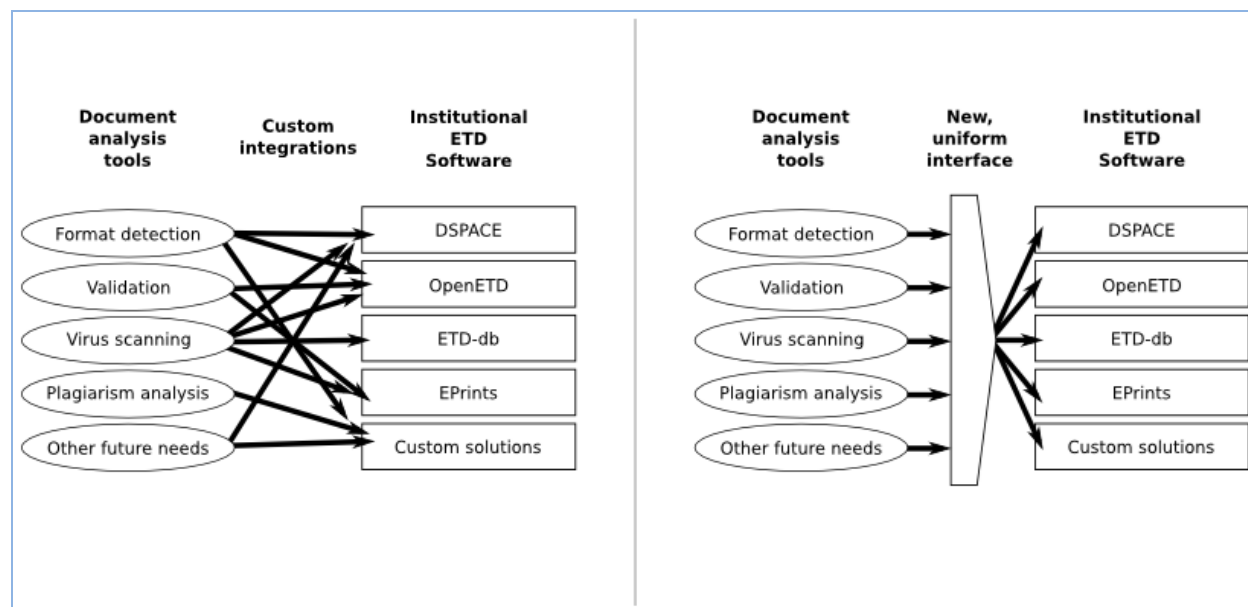

Figure 2. Tools development for ETDs lifecycle management project.

Descriptions of Educational Materials and Associated Workshop

During the second year of the project, a set of educational materials will be created on the theory and practice of lifecycle management of ETDs--which will be made freely available--and give a workshop utilizing these materials. These materials will reference the guidance documents, while the workshop will offer hands-on lifecycle tools training. The educational materials will include syllabi, training handouts and exercises, and PowerPoint presentations.

Summary

Theses and dissertations are predominantly produced in digital format today, and it is the academic libraries' responsibility to meet the challenge of managing these documents throughout their lifecycle. This project provides three key resources to help practitioners succeed. The guidance documents will provide libraries with a much needed core set of concise advisory documents to address practical implementation considerations in ETD programs. Although necessarily limited in scope, the software tools address key functions that virtually all ETD programs must provide, and will be available

via a free and open source software license for any academic library to utilize. And lastly, the educational materials and workshop will enable this project to directly reach practitioners seeking information and tools relevant to establishing or updating ETD programs at their institutions.

References

IMLS ETD Lifecycle Management Project Wiki. (n.d.). Retrieved on August 24, 2012 from:
        http://metaarchive.org/imls/index.php/Main_Page