Katherine Skinner and Martin Halbert from the MetaArchive Cooperative will be attending the Dec. 15, 2010, NDLTD Board meeting (and Gail McMillan via skype) to discuss two principle objectives:

- How to continue to build on our joint affiliation in the coming year to benefit both of our communities.
- The draft document below, which is a starting point for a conversation about what type of grant-funded project we might consider on ETD use cases for lifecycle management and preservation [such as the Institute for Museum and Library Services National Leadership Grant].

Abstract (rough draft): Managing the Lifecycle for Electronic Theses and Dissertations

The Educopia Institute with the Networked Digital Library of Theses and Dissertations and the libraries of University of North Texas, Virginia Tech, Georgia Tech, Boston College, and xxxxxxx, proposes to study, document, and model a set of use cases for the life-cycle data management and preservation of Electronic Theses and Dissertations (ETDs).

Colleges and Universities have been steadily transitioning from traditional paper/microfilm to digital ETD submission, dissemination, and preservation processes. Increasingly, colleges and universities worldwide are accepting and archiving only electronic versions of their students’ theses and dissertations. While this move from print-based to digital-based theses and dissertations greatly enhances the accessibility and sharing of graduate student research, it also raises grave concerns about the potential ephemerality of these digital resources. How will institutions ensure that the electronic theses and dissertations they acquire from students today will be available to future researchers?

Since the launch of the first required ETD program at Virginia Tech in 1997, institutions have been using locally derived methods of storing and providing access to their ETD collections, including a highly diverse and ever-evolving set of repository environments, metadata schemas, formats, and file/folder structures. Surveys conducted by the NDLTD and the Educopia Institute in 2009 and 2011 show that (x, y, z: something here about % of institutions that are accepting e submissions, % of institutions that are e-only, range of content formats accepted, and % of institutions that are currently preserving their content beyond backing it up). This content genre is of great value to institutions and to the scholars and researchers that they serve, and it is in critical need of preservation attention.

As institutions cease to collect and store paper or microfilm versions of this legacy content, serious questions regarding the preservation readiness of ETDs arise. The diversity of file types, formats, metadata, and structures that constitute this content genre raises two major concerns: How can curators effectively manage this growing set of data? And how can they ensure that these collections will be accessible in the long term?

We need to better understand, document, and address the preservation challenges presented by ETDs to ensure that colleges and universities have the requisite knowledge to properly curate these new collections.

We propose to investigate these issues through the following research questions, such as

1. **How can curators effectively manage the lifecycle of their ETDs?**
   We will develop guidelines and a range of use cases for the data management and preservation of ETD collections. We will also conduct a cost-benefit analysis of these use cases to provide institutions with a range of options to consider for their local needs.
2. **What are practical guidelines and specifications that ETD curators may use when preparing SIPs and ingesting content?**

The project team will study and document the creation of Submission Ingest Packages specifically for ETDs in conjunction with a set of leading preservation repositories (MetaArchive, Chronopolis, HathiTrust, Rosetta, DAITSS). We will develop guidelines for format selection, normalization, and migration practices. We will use the NDLTD’s existing standard set of metadata elements as a base upon which we can build a METS profile geared explicitly for preservation (using PREMIS) to provide institutions with a full profile that they may use as they prepare their ETD content for ingest into these preservation systems/repositories/networks.

This research will result in lifecycle management guidelines and use cases and models for SIP creation and data preparation for the preservation of ETDs.

**Some potential next steps:**
- Establish deliverables
- Seek partners
- Name an Advisory Committee
- Conduct a survey?
- From the surveys (2009 and 2011), build a readiness evaluation (basis for narrative and suggestions)