Retrospective Digitization of Theses and Dissertations: Key Points

Variations in RTD practices reflect
- Institutional characteristics, values and strategic vision
- Financial resources
- Assumptions and influences in decision making
- Calculations of return on investment

Questions to Consider
- Why Digitize?
- When to Digitize?
- What to Digitize?
- Who are stakeholders? Who should be involved?
- How to capture and convert?
- How to document metadata?
- How to deliver and display?
- Where? Offsite or onsite?
- How much does it cost?
- How to manage rights issues?
- How to manage quality control?

Why Digitize?
- Philosophical benefits
- Operational benefits
- Customer Service benefits
- Not sure of the benefits

When to Digitize?
- In relation to ETD launch
- External Factors
  - Funding becomes available
  - Run out of space to keep print volumes
  - Gifted copies from college
  - As items are requested / ILL or circulated
  - Opening in internal scanning queue
  - Vendor opportunity (Proquest, Lyrasis)

What to Digitize?
- All graduate works
- Doctoral Dissertations only
- Master’s Theses only
- Subset by department/institute
- Subset by date of creation
- Subset by subject
- Subset by format
- Subset by copyright status
- Selected by request (ILL, user query)
Retrospective Digitization of Theses and Dissertations: Key Points

USTEDA Pre-Conference

May 18, 2011

From What Source to Digitize?
- Printed and bound circulating copy
- Printed and bound preservation copy
- Microfilmed copy
- How to handle: Oversize materials; Included Media; Non-standard paper (onion skin)

Who are the Stakeholders?
- Who should be involved in the project?
- Are there any surprise stakeholders?
- How do you obtain buy-in?

How to Capture and Convert?
- Imaging requirements
- Text conversion/Optical Character Recognition (OCR)
- File Format(s) for access? For preservation?

How do you Document Metadata?
- Who creates or enhances the record?
- What standard to use?
- Where to publish the records?
- Will metadata be harvestable by Internet search engines?

How to Deliver and Display?
- Institutions seems to select an approach based on their values, resources, and workflows
- Choices somewhat connected to rights management approaches
- Key decision points: File format(s); Delivery platform; Type of Access; Relationship to ETD Collection; Relationship to search and retrieval facilities

Access Options
- Open, unrestricted access on the Web
- Restricted access
  - Full-text on campus
  - Some text off campus
- Full-text on demand
- No online access
Retrospective Digitization of Theses and Dissertations: Key Points

Where to Digitize?
- In-house staff
  - Digital Library Center, Preservation, ILL, Reserves, Archives and Special Collections
  - Cataloging - metadata
  - Systems – storage, delivery, back-up, preservation
- Third-party vendors
  - Digitization service bureaus
  - Publishers (Internet Archive, ProQuest, Google, etc)
  - Binderies

How Much Does it Cost?
- Some unit costs reported by page, others by volume
- Estimates range from pennies to dollars per page
- May be dependent on services provided
  - Document handing (paper, disbinding/rebinding, microfilm)
  - Scanning
    - Bitonal, grayscale, or color?
  - Optical Character Recognition (OCR)
    - Corrected or uncorrected?
  - Metadata creation
  - Open access vs. proprietary access
- Associated costs not always figured in or reported
  - Storage costs, time compiling complete list of titles, time for clearing permissions, etc.

How to Manage Rights Issues?
- Rights Management Issues
  - Copyright
  - Privacy rights
  - Trade secrets/Proprietary information
- Copyright Issues
  - Are ALL theses and dissertations necessarily copyrighted?
  - Who owns copyright?
    - in graduate student’s work
    - in included 3rd party material
  - Do original usage provisions apply to digital reproduction?
How to Manage Quality Control?

- Quality control level can range from in-depth pre- and post-scan QC on 100% of items, to cursory check that *something* is online, and everything in between.
- Which you go with depends on:
  - Goal of the project
  - Funding available
  - Staff time available
  - Mechanisms in place that allows for corrections to be made

Going Forward

- Assess your own
  - Institutional characteristics, values and strategic vision
  - Financial resources
  - Past practice and existing relationships
  - Assumptions and influences in decision making
  - Calculate return on investment - Consider opportunity costs
- Proceed in the way that makes most sense for you!

Credits (Presenters)

Gail Clement, Texas A&M University - gclement@tamu.edu

Several years ago (before the development of its digital library program), Texas A&M arranged for digitization of dissertations from microfilm copies at ProQuest. The University is currently arranging for digitization of 16,000 Master’s theses by a service bureau. These digital volumes will be accessible via open access in the Texas A&M Digital Repository.

Lee Dotson, University of Central Florida - Lee.Dotson@ucf.edu

The University of Central Florida is performing in-house digitization of theses and dissertations. The digitized versions are made publicly available via CONTENTdm and the library catalog.

Christy Shorey, University of Florida - chrshor@uflib.ufl.edu

The University of Florida began scanning dissertations in-house, then moved to scan with Internet Archive. The catalog links to the vendor website and a copy saved to the Institutional Repository. Master’s theses are scanned on a per-request basis.
Retrospective Digitization of Theses and Dissertations: Further Info

USTEDA Pre-Conference May 18, 2011

Reading List


Additional Resources

ETD-L Discussion List Archives (discussions about RTD’s date back a decade!)
http://listserv.vt.edu/cgi-bin/wa?A0=ETD-L

Massachusetts Institute of Technology, “MIT Theses”, Dspace@MIT Website. Online, URL:

Southern New Hampshire University (2010)“ SNHU Academic Archive Project”. University Website. Online, URL:


University of British Columbia Retrospective Theses and Dissertations:
http://circle.ub.ca/handle/2429/831

University of Central Florida Retrospective Theses and Dissertations (part of Digital Collections):

University of Florida Retrospective Dissertation Scanning Project:
Project summary, publicity pieces, and author permission forms:
http://www.digital.uflib.ufl.edu/procedures/copyright/retro_diss_scan/DDDCA.htm

Xavier University Retrospective Electronic Theses & Dissertations (ETD) Permission Form:
http://www.xavier.edu/cm/electronic-resources/retrospective_etd_permission_form.cfm