

Out of the Woods

Charting Metadata with Digital Tools

This poster highlights the MARC to Dublin Core metadata transformation and the use of the automation tool kit to streamline the metadata process, a necessary step in a large-scale digitization project that promotes accessibility to scholarly materials.

A. Laura Ramirez | Marian R. Smith | Salima Bowaniya | **University of Houston Libraries**
Project Collaborator: Andrew Weidner



Misty Metadata Mountain

PHASE 1

Transform and Crosswalk Metadata

- MARC (.mrk) export of metadata records is downloaded
- Transformed metadata -MARCEdit is used to remove MARC formatting and coding
- OpenRefine and General Refine Expression Language is used to clean up metadata and format to Dublin Core
- Use Terminal commands to run a script and create .txt files of each metadata record



Copyright Woods

PHASE 2

Editing and Information Input

- Generated information (.txt) is edited and compared to digitized thesis (PDF) for errors
- Additional information (advisors, abstracts, etc.) is filled into appropriate fields in the .txt
- Digitized thesis (PDF) is evaluated for errors (pagination, private information is redacted, etc.)

Copyright Evaluation

- Digitized thesis (PDF) is checked for copyright attribution and any additional copyrighted materials
- Copyright statement is assigned



Ruby Castle

PHASE 3

Validate YAML Stage

- Script checks .txt file for configuration/formatting issues
- Script verifies that the copyright designation is filled in properly

Add Metadata Exif Stage

- Script adds metadata information to the digitized thesis with Exiftool
- Script moves the documents to ingest based on copyright status



Sustainability Compass

Communication

Establish communication channels for information sharing.

Documentation

Capture large and formal decisions in appropriate documentation.

Cross-training

Key staff should gain understanding of the processes throughout the workflow to help resolve any issues that may arise.



Cougar Bayou

To Ingest

Items have been moved to ingest and are ready to be uploaded into DSpace!

