Treasure Chest on the Move: Creating and Leveraging METS Records to Transfer Complex Digital Objects

Michael J. Dobbs
Catholic University of America
Washington, DC

Background:
• Treasure Chest ran from 1946 – 1972
• Comprised 27 volumes and nearly 500 issues
  • Average of 35 pages per issue.
• Digitization began in the early 2000s at the Washington Research Library Consortium (WRLC)
• Digital Projects Office at WRLC reorganized in the middle of publishing online – stranding the project
• Frozen servers restricted access
• Need to migrate became apparent
• Metadata for migrating the collection in order was needed due its size and complexity

Simple:
- A single page
- Easily migrated

Complex:
- Entire Collection
- Tricky to migrate

Simple, flat object:
- Entire Collection

Complex object:
- Entire Collection

Work Completed:
• Complete Excel spreadsheet hierarchies
• Use Perl script to translate into structural metadata in METS
• “Wrap” descriptive MODS in METS
• Quality analysis of XML after automated processes
• Use extant indexing to curate alternative viewing structures

Lessons Learned:
• Complexity of collaborative digital projects in terms of coordination and future planning
• Variety of skills needed, including technical and management skills
• Metadata, cataloging, and naming standards need to be a collaborative decision
• Dream big, but plan contingencies in basic terms

Issues:
• Indexing inconsistencies
• Perl problems
• Duration of delays
• Unclear understanding (of dispersed responsibilities)

Next Steps:
• Evaluating final comics for re-scanning and re-indexing
• Migrating the full-run, after the repository is chosen
• Applying lessons to future digitization projects

Want more?
Check out the finding aid and digitized comics

Contact:
• http://archives.lib.cua.edu
• Email: dobbs@cua.edu
• Twitter: @CUAarchives