Bridges to Hyku: A Multi-Institutional Effort Towards Migration Tooling

Anne Washington, Dean Seeman, Nick Woodward, Todd Crocken
Sean Watkins, Annie Wu, Andy Weidner, Santi Thompson, Leroy Vallejo

TCDL
MAY 22, 2019
OUTLINE

● Bridge2Hyku basics
● B2H at...
  ○ University of Houston
  ○ University of Victoria
  ○ Texas Digital Library
● What’s next?
BASICS

- 2 Year IMLS grant project (lg-70-17-0217-17)
- Based at University of Houston with partner institutions: University of Victoria, University of Miami, IUPUI, and Texas Digital Library
- Advisors/Stakeholders: Stanford, Indiana, DPLA, Duraspace.
GOALS

● Migration strategies and best practices
● B2H website
● Migration software
  ○ Export
  ○ Import
● Community support and sustainability
Conduct needs assessment (metadata and system requirements)

Develop and document B2H Toolkit and website

● System selection...
  ○ Limitations of CONTENTdm
  ○ Initially chose Hydra-in-a-Box (Hyku)
● ...And migration
  ○ Bridge2Hyku
● UH pivot to Hyrax
UH + PARTNERS

- Develop and prioritize use cases for the toolkit
- Lead development of products:
  - Website
  - CDM Bridge
  - HyBridge
WEBSITE

- Migration best practices
- Online resources
- Toolkit instructions and info
- Hyku features

- Desktop Electron-app
- Works with local and hosted CONTENTdm (CDM)
- Metadata crosswalking
- Export csv and files
<table>
<thead>
<tr>
<th>Export Field</th>
<th>Item Level</th>
<th>CDM Field</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Title</strong></td>
<td></td>
<td><em>Title</em></td>
</tr>
<tr>
<td><strong>Creator</strong></td>
<td></td>
<td><em>Creator (LCNAF)</em></td>
</tr>
<tr>
<td></td>
<td></td>
<td><em>Creator (HOT)</em></td>
</tr>
<tr>
<td></td>
<td></td>
<td><em>Creator (ULAN)</em></td>
</tr>
<tr>
<td></td>
<td></td>
<td><em>Creator (Local)</em></td>
</tr>
<tr>
<td><strong>Keyword</strong></td>
<td></td>
<td><em>Subject, Topical (Local)</em></td>
</tr>
<tr>
<td><strong>Rights statement</strong></td>
<td></td>
<td><em>Use and Reproduction</em></td>
</tr>
<tr>
<td><strong>Contributor</strong></td>
<td></td>
<td><em>Contributor (LCNAF)</em></td>
</tr>
<tr>
<td></td>
<td></td>
<td><em>Contributor (HOT)</em></td>
</tr>
<tr>
<td></td>
<td></td>
<td><em>Contributor (ULAN)</em></td>
</tr>
<tr>
<td></td>
<td></td>
<td><em>Contributor (Local)</em></td>
</tr>
<tr>
<td><strong>Abstract or Summary</strong></td>
<td></td>
<td><em>Description</em></td>
</tr>
<tr>
<td><strong>License</strong></td>
<td></td>
<td>-- Select a field --</td>
</tr>
<tr>
<td><strong>Publisher</strong></td>
<td></td>
<td>-- Select a field --</td>
</tr>
</tbody>
</table>
HYBRIDGE  


- Ruby Gem for Hyrax and Hyku
- Creates easy to use tab for bulk ingest
- Allows collection-specific ingest
- Seamlessly integrated with CDM Bridge’s output
UVic - CDM Migrator

- Created by Braydon Justice at UVic
- Allows for collection properties within CONTENTdm to be mapped to Hyku properties without developer mediation
UVic - CDM Migrator

- CDM-Migrator within hyku-vagrant stack
  - Modified for Hyrax 2.0 compliance
  - Hyku-vagrant stack created by Sean Watkins at UH
  - Technical users will be able to bring this up on a vagrant box and do a full test migration on it.

https://github.com/UVicLibrary/cdm_migrator
https://bridge2hyku.github.io/toolkit/
UVic - CDM Migrator

Dashboard
Registered Users
Current: 11

CDM Mapping

Csv Batch Uploader
UVic - Migration

Remediation

- URI-centred (OpenRefine, Vaultify)
- Faceted Environment: Exploding out Subject Headings to component parts
- RightsStatements
Texas Digital Library (TDL)

- TDL Hyku pilot - January-June, 2018
  - University of Houston, UT Austin, Trinity University
  - Needs assessment and feasibility testing
- Joined Bridge2Hyku in September, 2018
  - Migration workflow tooling
  - Performance testing
TDL in B2H

- Launched two Hyku instances on AWS with monitoring
  - Set up for performance testing
- Refactored deployment to handle migration
  - Includes cdm_migrator and hybridge gems
- Mounted shared storage with Secure FTP access
  - Separated by Hyku tenant
- Additional tooling for generating derivatives
  - Custom built ImageMagick
Disk utilization for /dev/nvme1n1 - from Sun Sep 30 22:03:41 2018 to Sun May 26 22:03:41 2019
TDL in B2H

- Optimizing configuration for server load
- Attempted to add additional metadata to filesets
- Custom work type for cdm_migrator
- Lessons learned
  - Staging point for the data (Secure FTP, S3, etc.)
  - Managing spikes in system resources
  - Deeper Samvera knowledge/experience
WHAT’S NEXT?

● Migration work continues
● B2H Project: Assess, improve, and promote toolkit
● We need your help!
● Test our tools and share your feedback!
THANKS!