

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



TIMELINE

Phase 1
Oct 2017 - Apr 2018

Conduct needs assessment (metadata and system requirements)

Phase 2
May 2018 - Apr 2019

Develop and document B2H Toolkit and website

Phase 3
May 2019-Sept 2019

Assess, Improve, and Promote B2H Toolkit.



BRIDGE2HYKU AT...





UH

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

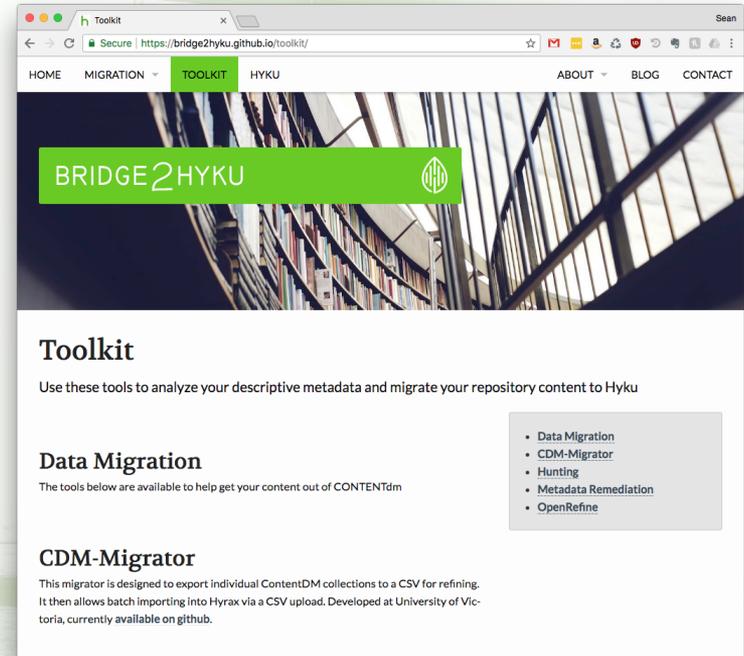


UH + PARTNERS

- 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





CDM BRIDGE- <http://bit.ly/cdm-bridge>

- Desktop Electron-app
- Works with local and hosted CONTENTdm (CDM)
- Metadata crosswalking
- Export csv and files



CDM Bridge


Export Collection ▾

Collections	Export Field	Item Level	CDM Field
1850s and 1860s Hotel and Restaurant ...	Title <i>Required</i>	<input checked="" type="checkbox"/>	Title ▾ +
Andy Warhol - Celebutants, Groupies, ...	Creator <i>Required</i>	<input type="checkbox"/>	Creator (LCNAF) ▾ +
Annual Reports from the Hospitality Ind...			Creator (HOT) ▾ -
Architectural Retail Catalog Collection			Creator (ULAN) ▾ -
The Astrodome, Eighth Wonder of the ...			Creator (Local) ▾ -
Bebe Gow Scrapbooks			
Blanche Espy Chenoweth, Her Life, Her...			
Board of Regents Records	Keyword <i>Required</i>	<input type="checkbox"/>	Subject.Topical (Local) ▾ +
Book of Hours, Use of Reims	Rights statement <i>Required</i>	<input type="checkbox"/>	Use and Reproduction ▾ +
Building Houston	Contributor	<input type="checkbox"/>	Contributor (LCNAF) ▾ +
Burdette Keeland Architectural Drawin...			Contributor (HOT) ▾ -
Captain O.C. McDavid World War II Let...			Contributor (ULAN) ▾ -
Carlos (DJ Styles) Garza Drawings and ...			Contributor (Local) ▾ -
City of Houston, Its Development and E...			
Civil War Letters	Abstract or Summary	<input type="checkbox"/>	Description ▾ +
Daily Cougar	License	<input type="checkbox"/>	-- Select a field -- ▾ +
DJ Screw Photographs and Memorabilia	Publisher	<input type="checkbox"/>	-- Select a field -- ▾ +
Domestic Crude			





HYBRIDGE- <http://bit.ly/hykubridge>

- 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

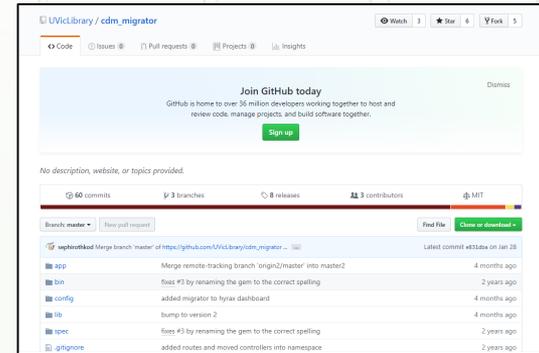
Title	Choose one Applies to boths compound objects and children: <input type="text" value="title"/> Applies only to compound objects: <input type="text"/>
Library has	Choose one Applies to boths compound objects and children: <input type="text"/> Applies only to compound objects: <input type="text"/>
Description	Choose one Applies to boths compound objects and children: <input type="text"/> Applies only to compound objects: <input type="text"/>
Genre	Choose one Applies to boths compound objects and children: <input type="text"/> Applies only to compound objects: <input type="text"/>
Subject	Choose one Applies to boths compound objects and children: <input type="text"/> Applies only to compound objects: <input type="text"/>

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

The screenshot shows the 'Vault' dashboard. The left sidebar contains a 'TASKS' section with 'Cdm Migrator' expanded, showing 'Csv Batch Uploader' and 'Cdm Mapping Tool'. A red circle highlights 'Csv Batch Uploader' with an arrow pointing to the expanded view below.

Dashboard Data:

Registered Users		Total
Current	11	Past

Administrative Sets:

- Default Admin Set 2
- Default Admin Set

User Activity: New user sign ups



The 'CDM Mapping' tool interface shows a list of fields for mapping:

- Title:** Choose one. Applies to both compound objects and children: title. Applies only to compound objects: [dropdown]
- Alternative title:** Choose one. Applies to both compound objects and children: alternative_title. Applies only to compound objects: [dropdown]
- Description:** Choose one. Applies to both compound objects and children: description. Applies only to compound objects: [dropdown]
- Genre:** Choose one. Applies to both compound objects and children: genre. Applies only to compound objects: [dropdown]
- Subject:** Choose one. Applies to both compound objects and children: subject. Applies only to compound objects: [dropdown]



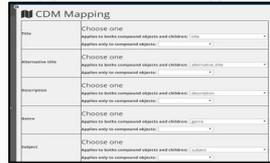
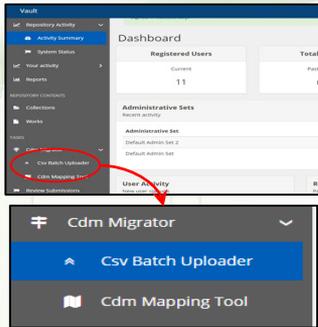
The 'CSV Batch Uploader' tool interface includes the following fields:

- Multi-value separator: [input field]
- Choose File: No file chosen
- Admin set: [dropdown menu]
- Collection: [dropdown menu]
- Save button
- template link

The expanded 'Cdm Migrator' menu shows three options:

- Csv Batch Uploader (highlighted in blue)
- Cdm Mapping Tool

UVic - Migration



Remediation

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





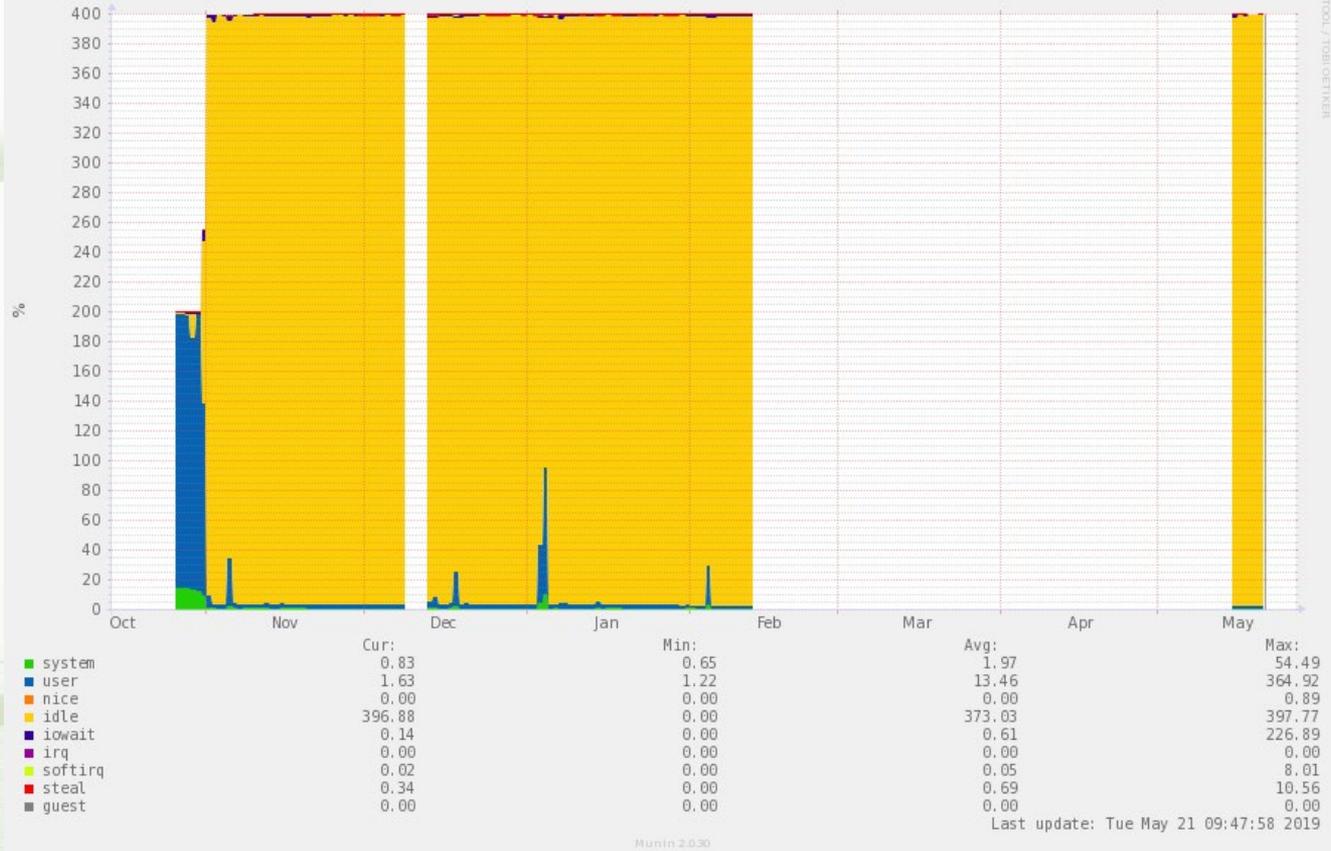
Texas Digital Library (TDL)

- 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 w/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

CPU usage - from Sat Oct 13 21:48:38 2018 to Sun May 26 21:48:38 2019





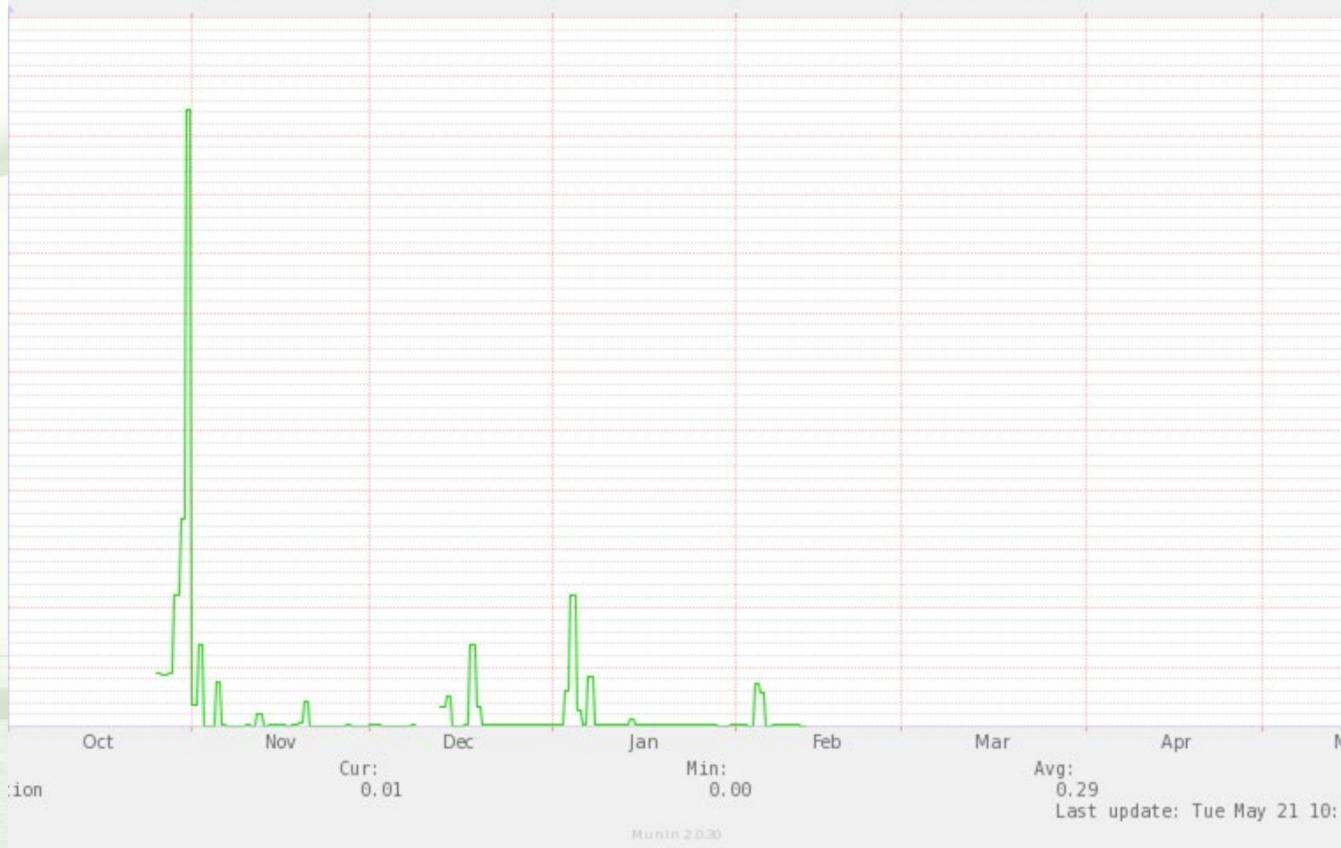
IOs for /dev/nvme1n1 - from Sat Sep 22 22:08:42 2018 to Mon May 27 10:08:42 2019



	Oct	Nov	Dec	Jan	Feb	Mar	Apr	Ma
		Cur (-/+)		Min (-/+)		Avg (-/+)		Ma
ze (KB)	0.00 /	2.37	0.00 /	2.23	308.93m/	2.72	62.60 /	
	0.00 /	7.27	0.00 /	4.66	616.92m/	8.26	142.94 /	
						Last update: Tue May 21 10:11		

MunIn 2.0.30

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!

Background from unsplash.com

