

Bridge2Hyku:

Tools and Strategies for Content Migration to Open Source Repository



Annie Wu, Santi Thompson, Andrew Weidner,
Todd Crocken, Sean Watkins, Anne Washington,
and Leroy Vallejo

2019 Open Repositories
Hamburg, Germany

Agenda

- Introduction
- Bridge2Hyku (B2H) project background
- B2H project needs, impact and goals
- B2H Toolkit Alpha release and pilot testing
- Communication and sustainability
- What's next

UH Libraries Move to Open Source Digital Solutions

- Evaluated DAMS, digital preservation and archive landscape
- Recommended to implement open source digital solutions
- Develop an eco-system for systems integration



The diagram illustrates the UH Libraries Digital Preservation Architecture, showing the flow of data and metadata between various systems. The architecture is organized into several components and their interactions:

- Input/Output:**
 - PM/MM/AC Files** (Preservation Metadata/Metadata/Access Control Files) feed into the **Carpenters** workflow.
 - BCDAMS-MAP Metadata Application Profile** provides **Preservation Metadata Spec** and **Access Metadata Spec** to the **Carpenters** and **Brays** workflows.
- Core Workflows:**
 - Carpenters Digitization Workflow** (UH Libraries Dev) is the central hub for digitization.
 - Brays Metadata Workflow** (UH Libraries Dev) handles metadata management.
- Intermediate Systems:**
 - Greens ID Minter** (UH Libraries Dev) and **Cedar SKOS Vocab Manager** (UH Libraries Dev) interact with the **Carpenters** workflow via **Mint Preservation Object ARKs** and **Mint Authority ARKs**.
 - Authority Control** feeds into the **Cedar** system.
 - Shared Data File** is a common resource for **Carpenters** and **Brays**.
- Output/Storage Systems:**
 - Archivematica** (Community Dev) receives **Preservation Object SIPs** from **Carpenters** and sends **POST Archival Items (If Needed)** back.
 - ArchivesSpace** (Community Dev) receives **POST Preservation Object ARKs** from **Archivematica** and **POST Access Object ARKs** from **White Oak**.
 - Halls Public User Interface** (Community Dev) is part of the **ArchivesSpace** stack.
 - White Oak ID Router** (Community Dev) sends **POST Access Object ARKs** to **ArchivesSpace**.
 - Hyku|Avalon** (Community Dev) receives **Access Object DIPs** from **Brays**.
- Key:**
 - Red boxes:** UH Libraries Dev
 - Blue boxes:** Community Dev
 - Arrows:**
 - Solid arrow: Data Input
 - Dashed arrow: Data Exchange
 - Thick solid arrow: Data Output

Samvera Community



[about samvera](#) [what is samvera?](#) [why use samvera?](#)

[Home](#) » [Who Uses Samvera?](#)

SAMVERA PARTNERS

The Samvera Community is built around a core group of Partners committed to the success of the project and more importantly to the success of each other. There are also many institutions and individuals that have formally committed to the sustainability of the Community through code contributions, participation in working or interest groups, helping with documentation, design, or planning. No contribution is too small. If you are interested in contributing to the Samvera Community or just interested in Samvera in general read more about the Community [here](#).

The number of Samvera Partners has grown steadily worldwide.



35 partner institutions
primarily in North America
and Europe

Solution Bundles:

- Avalon
- Hyku
- Hyrax

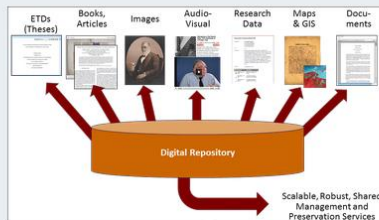
Samvera Features

SAMVERA IS FLEXIBLE AND EXTENSIBLE

The Samvera framework uses the best of current repository, search, and user-interface technologies. The separation of data and display, within the Samvera framework, promotes flexibility and extensibility.

Flexible

Within Samvera framework, one body, the digital repository, supports multiple heads or tailored, content type-specific applications. Additional heads can be implemented with the same repository to handle new forms of content as they emerge. All content stored in the repository can benefit from common curatorial functions and end-users benefit from tailored discovery and access interfaces.



Extensible

The Samvera framework is architected to extend features based on specific concerns and use-cases. Established features with long-term community commitment are available, such as:

- **Browse Everything** allows your rails application to access user files from cloud storage, such as Dropbox, Skydrive, Google Drive, Box, and a server-side directory share
- **LDP** provides a Linked Data Platform and allows publication of RDF data

The agile and collaborative nature of the Samvera Community also enables experimentation with cutting-edge extensions, which can be added to address unique needs, explore usefulness and possibilities, and eventually gain community commitment. Experiments include:

- GeoMash for linked data and geo-vocabularies
- International Image Interoperability Framework (IIIF) manifest
- ORCID integration
- SHARE notify service
- Shibboleth authorization integration

- Flexible, scalable and extensible
- Support multiple heads and special format applications
- Cutting-edge features extension
- Linked Data Platform (LDP)

[about samvera](#)[what is samvera?](#)[why use samvera?](#)[who uses samvera?](#)[getting started](#)[news & events](#)

[Home](#) » [News & Events](#) » University of Houston Libraries Partner with Samvera

UNIVERSITY OF HOUSTON LIBRARIES PARTNER WITH SAMVERA

Posted on August 18, 2017 - News

The University of Houston Libraries (UHL) have become Samvera's latest formal Partner.

UH Libraries has been actively involved in the Samvera Community through the Metadata Interest Group, CONTENTdm Migration Working Group, and Hyku efforts.

In addition to direct Samvera Community involvement, UH Libraries supports Samvera through activities in the Fedora repository community. UH Libraries helped initiate the South Central States Fedora Users Group (SCSFUG) with Texas Digital Library (TDL) to form a growing multi-state community of library faculty and staff who are actively working on or interested in Fedora-based repository projects. They have hosted events to continue to conversation and support for our base repository system.

NEWS & EVENTS

[Samvera Connect 2017 and 2018](#)

[Samvera Connect 2017 Sponsor Prospectus](#)

[Samvera Connect 2017](#)

[Samvera and the world's top universities](#)

[University of Houston Libraries Partner with Samvera](#)

RECENT TWEETS

Tweets by [@samverarepo](#)



Samvera Community Retweeted



IMLS National Leadership Grant

Bridge2Hyku Toolkit:
Developing Migration
Strategies for Hyku

LG-70-17-0217-17

Oct. 1st, 2017 – Sept.
31st, 2019



Bridge2Hyku – International Collaboration

Implementation Team



<http://www.uh.edu>

Annie Wu *Principal Investigator*
Santi Thompson *Principal Investigator*
Andy Weidner *Project Manager*
Todd Crocken *Content Strategist*
Sean Watkins *Software Developer*
Anne Washington *Migration Strategist*



<https://welcome.miami.edu>

Sarah Shreeves *Advisory Board Member*
Paul Clough *Software Developer*
Elliot Williams *Migration Strategist*



**University
of Victoria**

<https://www.uvic.ca>

Lisa Goddard *Advisory Board Member*
Braydon Justice *Software Developer*
Dean Seeman *Migration Strategist*



<https://www.iupui.edu>

Chip Dye *Advisory Board Member*
Andy Smith *Software Developer*
Kristi Palmer *Migration Strategist*
Richard Wing *Migration Strategist*

Community Stakeholders

- Stanford University
- DuraSpace
- Digital Public Library of America (DPLA)
- Indiana University – Bloomington

Recently Added Partner:

- Texas Digital Library (TDL)

Bridge2Hyku - Needs

- DAMS assessment
- Trend towards open source solutions
- Lack of migration tools and guidelines
- Lack of staff resources and expertise



Bridge2Hyku - Impact



- Fill current gaps in content migration practices
- Quicker and wider adoption of open source digital solutions
- Broaden access and expand use of digital content and collections

Bridge2Hyku – Goals and Outcomes

Bridge2Hyku Toolkit
website and Github site

- Migration strategies and best practices
- Migration software/tools
- Hyku features and requirements

Community support and sustainability



Project Timeline

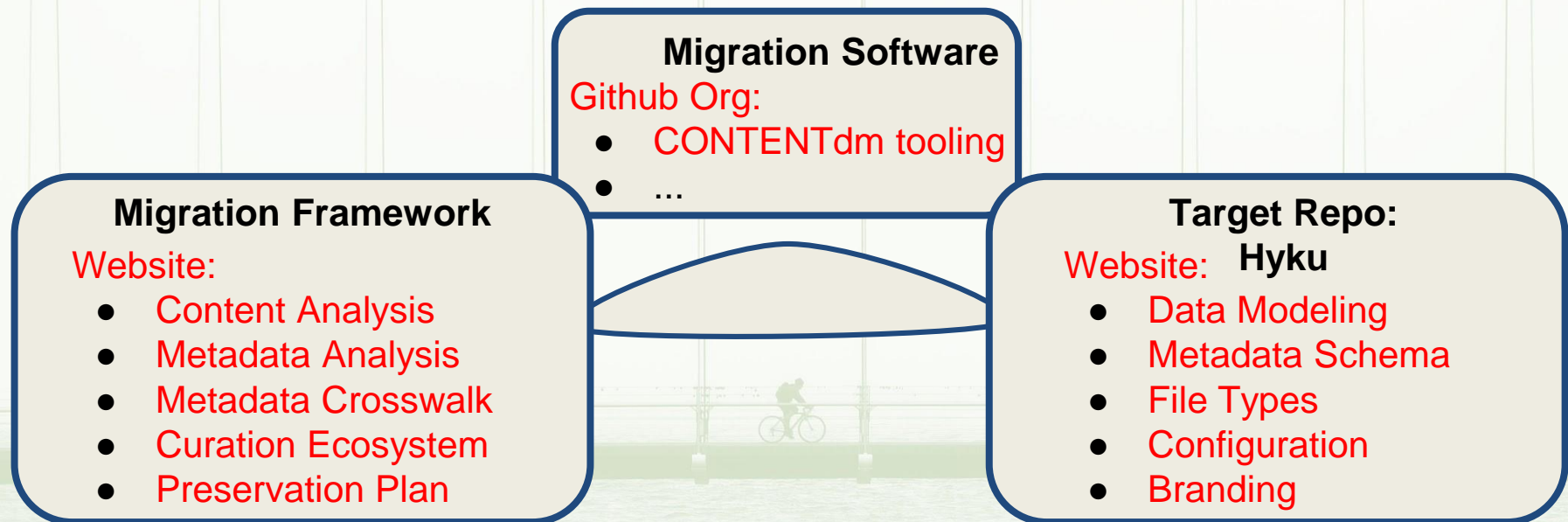
Phase 1
Oct 2017 - Apr 2018

Phase 2
May 2018 - Apr 2019

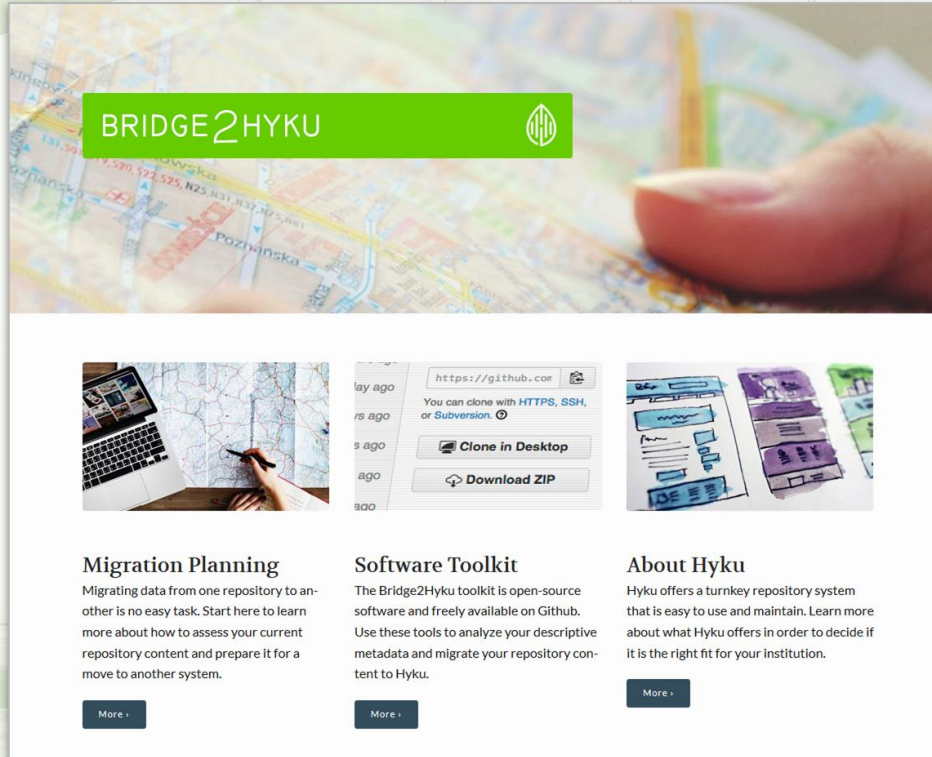
Phase 3
May 2019-Sept 2019

- Conduct needs assessment (metadata and system requirements) ✓
- Develop and document B2H Toolkit (website, tools and guidelines) ✓
- Assess, Improve, and Promote B2H Toolkit.

Project Design



Bridge2Hyku Toolkit



- Migration Strategies
- Migration Tools
- Hyku Features

<https://bridge2hyku.github.io/>

Content Migration Strategies

BRIDGE²HYKU



Migration

Learn how to assess your current repository content and prepare it for a move to another system

Migration Overview

Data migration for digital collections is a complex process. As with most difficult tasks, it can be helpful to break down what needs to be done into smaller blocks of work. In general, data migrations can be divided into four sub-tasks, each with their own sets of outcomes, tools, and techniques: Migration Planning, Metadata Normalization, Content Migration, and Content Verification.

The following sections describe the importance of each of these steps in more detail and provide resources for effectively planning and executing your next content migration. This page does not assume that you have chosen a target system for your migration or that you are migrating from a specific source platform. This information is intended as a general migration guide for both new and seasoned digital collections professionals.

Table of Contents

- [Migration Overview](#)
- [Migration Planning](#)
 - [Contextual Considerations](#)
 - [Digital Library Analysis](#)
 - [Content Analysis](#)
 - [Metadata Analysis](#)
- [Normalize Metadata](#)
- [Migrate Content](#)
 - [Model Work Types](#)
 - [Map Metadata](#)
 - [Source Repository Export](#)
 - [Target Repository Import](#)
- [Content Verification](#)
 - [Approaches](#)

- Migration Planning
- Normalize Metadata
- Migrate Content
- Content Verification
- Migration Resources

Content Migration Tools



Toolkit

Use these tools to analyze your descriptive metadata and migrate your repository content to Hyku.

Data Migration

The tools below are available to help move your content out of CONTENTdm and into Hyku.

CDM Bridge (Wiki)

<https://github.com/Bridge2Hyku/cdm-bridge/releases>

An Electron desktop application for crosswalking and exporting CONTENTdm metadata. It also provides the option of exporting files from CONTENTdm. CDM Bridge connects to the CONTENTdm API and presents a custom crosswalking interface for each collection. Metadata specialists may use the application to generate lists of field values for cleanup and normalization prior to migration. When a collection is ready for migration CDM Bridge generates a CSV, with or without the files from CONTENTdm, that may be used to ingest the collection into Hyku with HyBridge.

Table of Contents

- [Data Migration](#)
 - [CDM Bridge \(Wiki\)](#)
 - [HyBridge](#)
 - [CDM Migrator](#)
 - [Hunting](#)
- [Metadata Normalization](#)
 - [OpenRefine](#)
 - [LibreCat/Catmandu](#)
 - [DLF Metadata Assessment](#)
 - [Interest Group Toolkit](#)
 - [AutoHotkey](#)

- CDM Bridge
- HyBridge
- CDM Migrator

CDM Migrator



Toolkit

Use these tools to analyze your descriptive metadata and migrate your repository content to Hyku.

Data Migration

The tools below are available to help move your content out of CONTENTdm and into Hyku.

CDM Bridge (Wiki)

<https://github.com/Bridge2Hyku/cdm-bridge/releases>

An Electron desktop application for crosswalking and exporting CONTENTdm metadata. It also provides the option of exporting files from CONTENTdm. CDM Bridge connects to the CONTENTdm API and presents a custom crosswalking interface for each collection. Metadata specialists may use the application to generate lists of field values for cleanup and normalization prior to migration. When a collection is ready for migration CDM Bridge generates a CSV, with or without the files from CONTENTdm, that may be used to ingest the collection into Hyku with HyBridge.

Table of Contents

- [Data Migration](#)
 - [CDM Bridge \(Wiki\)](#)
 - [HyBridge](#)
 - [CDM Migrator](#)
 - [Hunting](#)
- [Metadata Normalization](#)
 - [OpenRefine](#)
 - [LibreCat/Catmandu](#)
 - [DLF Metadata Assessment](#)
 - [Interest Group Toolkit](#)
 - [AutoHotkey](#)

CDM Bridge

CDM BRIDGE

Export Collection

Collections

1850s and 1860s Hotel and Restaurant Menus

Andy Warhol - Celebutants, Groupies, and Friends

Annual Reports from the Hospitality Industry Archi...

Architectural Retail Catalog Collection

The Astrodome, Eighth Wonder of the World

Bebe Gow Scrapbooks

Blanche Espy Chenoweth, Her Life, Her Times

Board of Regents Records

Book of Hours, Use of Reims

Building Houston

Burdette Keeland Architectural Drawings & Photog...

Captain O.C. McDavid World War II Letters

Carlos (DJ Styles) Garza Drawings and Memorabilia

City of Houston, Its Development and Evolution

Civil War Letters

Daily Cougar

DJ Screw Photographs and Memorabilia

Domestic Crude

Export Field	CDM Field	
Title <i>Required</i>	Title	+
Creator <i>Required</i>	Creator (LCNAF)	+
	Creator (HOT)	-
	Creator (ULAN)	-
	Creator (Local)	-
Keyword <i>Required</i>	Subject.Topical (LCSH)	+
Rights statement <i>Required</i>	Use and Reproduction	+
Contributor	Contributor (Local)	+
	Contributor (LCNAF)	-
	Contributor (HOT)	-
	Contributor (ULAN)	-
Abstract / Summary	Description	+
License	-- Select a field --	+
Publisher	Publisher	+

example.localhost:3000/dashboard/my/works?locale=en

Apps Gmail Google Calendar BoardGameGeek Google Analytics Scrum Board Spidy GitLab Other Bookmarks

HykuEnglishSean Watkins

Sean Watkins

ACTIVITY

Repository activity

Your activity

REPOSITORY CONTENTS

Admin Sets

Collections

Works

HyBridge Ingest

TASKS

Review Submissions

Reports

CONFIGURATION

Settings

Workflow Roles

Users and groups

Home / Dashboard / Works

Works

Create batch of worksAdd new work

All WorksYour Works











21 works you own in the repository

Filter works:

VisibilityStatusAdmin Set

Show 10 per page

Enter search termsGo

	Title	Date Added	Highlighted	Visibility	Actions
<input type="checkbox"/>	 Parking lot one quarter mile from explosion		<input type="checkbox"/>	Private	Select
<input type="checkbox"/>	 Oil tanks on fire three days later		<input type="checkbox"/>	Private	Select
<input type="checkbox"/>	 Parking area one half mile from explosion		<input type="checkbox"/>	Private	Select
<input type="checkbox"/>	 Road to explosion area		<input type="checkbox"/>	Private	Select
<input type="checkbox"/>	 Row of damaged houses		<input type="checkbox"/>	Private	Select
<input type="checkbox"/>	 Corner view of damaged houses		<input type="checkbox"/>	Private	Select
<input type="checkbox"/>	 Damaged houses		<input type="checkbox"/>	Private	Select
<input type="checkbox"/>	 Warehouse destroyed		<input type="checkbox"/>	Private	Select
<input type="checkbox"/>	 Freight cars one quarter mile from explosion		<input type="checkbox"/>	Private	Select
<input type="checkbox"/>	 Motor convoy for Texas City		<input type="checkbox"/>	Private	Select

« Previous

Next »

123

HyBridge

Hyku Features

[HOME](#)[MIGRATION](#)[TOOLKIT](#)[HYKU](#)[ABOUT ▾](#)[BLOG](#)[SEARCH](#)

BRIDGE²HYKU



Hyku

Learn about what Hyku offers for your institution

National Digital Platform

Hyku, the repository software produced by the Hydra-In-a-Box project, was built as a part of the Institute for Museum and Library Services' [National Digital Platform](#) initiative. It strives to be "[a feature-rich, robust, flexible digital repository that is easy to install, configure, and maintain.](#)"

- [National Digital Platform](#)
- [Hyku vs Hyrax vs Omeka](#)
- [Hyku Features](#)
- [Hyku Metadata Schema](#)
- [Institutional Repository](#)
- [Cultural Heritage Repository](#)

Bridge2Hyku Github



Why GitHub? ▾ Enterprise Explore ▾ Marketplace Pricing ▾

Search



Sign in

Sign up

BRIDGE2HYKU



Bridge2Hyku

Repositories 7

People 1

Projects 0

Grow your team on GitHub

GitHub is home to over 31 million developers working together. Join them to grow your own development teams, manage permissions, and collaborate on projects.

[Sign up](#)

Dismiss

Find a repository...

Type: All ▾

Language: All ▾

hybridge

A ruby gem to batch ingest digital objects into the Samvera Hyku application

Ruby ★ 1 3 Apache-2.0 Updated a day ago

Top languages

Ruby Shell JavaScript
 TypeScript

People

1 >

Bridge2Hyku Blog

[HOME](#)[MIGRATION](#)[TOOLKIT](#)[HYKU](#)[ABOUT ▾](#)[BLOG](#)[SEARCH](#)[CONTACT](#)

BRIDGE2HYKU



UPDATES

Hibernation's End - Conference Season has Arrived



We haven't been sleeping, but we also haven't been posting. Find out where you can learn more about the Bridge2Hyku Project during this conference season! [Read More >](#)

Take the Bridge2Hyku [Digital Collections Survey](#) to help us understand your content migration needs.



Bridge2Hyku Twitter



Tweets **61** Following **143** Followers **65** Likes **55**

Follow

Bridge2Hyku

@Bridge2Hyku

We're building a migration toolkit for digital archives/repositories to move into Hyku. This project was made possible in part by @US_IMLS

📍 Houston, TX

📅 Joined January 2018

📷 Photos and videos



Tweets Tweets & replies Media

📌 Pinned Tweet



Bridge2Hyku @Bridge2Hyku · Mar 15

we've heard back from all our submissions, and gosh, we have a busy spring coming up. Check out our conference schedule and come learn more about our project face-to-face soon.



Hibernation's End - Conference Season has Arrived

We haven't been sleeping, but we also haven't been posting. Find out where you can learn more about the Bridge2Hyku Project during this conference season!

bridge2hyku.github.io



3



Bridge2Hyku @Bridge2Hyku · Mar 27

If you took a Workshop entitled "Strategies and Tools for Digital Repository Selection and Migration", what would be the most important information to walk away from it with?



1



2



Bridge2Hyku @Bridge2Hyku · Mar 1

this is still true!

New to Twitter?

Sign up now to get your own personalized timeline!

Sign up

Worldwide trends

#FridayFeeling

99.7K Tweets

#VietnamWarVeteransDay

16.2K Tweets

#FridayMotivation

32.4K Tweets

#FridayThoughts

21.9K Tweets

Agnès Varda

French New Wave director Agnès Varda dies at age 90

#FlashbackFriday

10.3K Tweets

LOVE MAZE

58K Tweets

Other Toolkit Promotion Venues

- Conference presentations
- Workshops and online training
- Open access publications



Pilot Testing and Enhancement

- November 2018 – February 2019
- Testing rubric
- Pilot recommendations



University
of Victoria

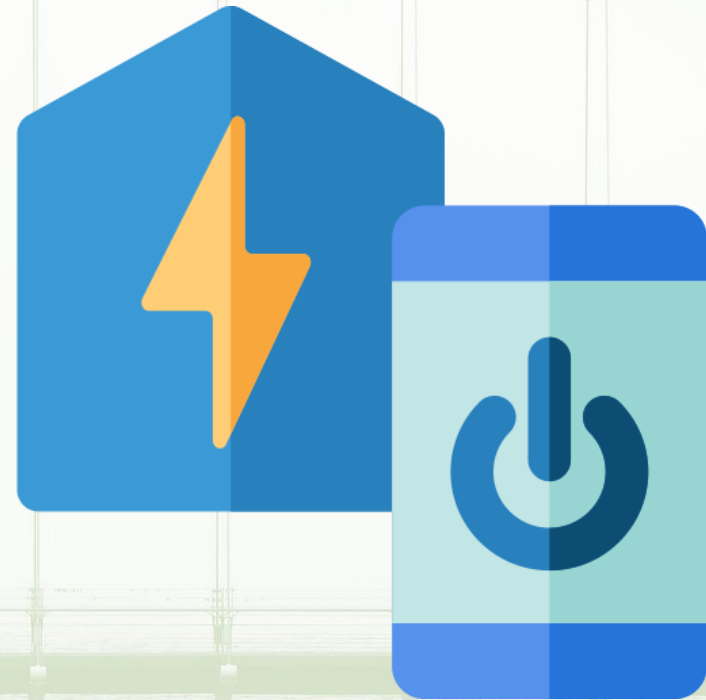
B2H Project Partnership

- Digital Collections Report
- Tools Development
- Toolkit Promotion



Next Steps and Sustainability

- B2H Github Organization (UH)
- Samvera Community
- Third party hosted support





Thank You

Contact Info:

Annie Wu
Santi Thompson

awu@uh.edu
sathompson3@uh.edu

Image Credits:

- Samvera: <https://samvera.org>
- African proverb: <https://makeameme.org/meme/if-you-want-p1qwui>
- Impact: <https://www.interstaterestoration.com/blog/why-business-impact-analysis-bia-important>
- Outcome: <https://www.qcs.co.uk/outcome-based-care/>
- Promotion: “Digital campaign free icon” Icon made by [Freeipik](https://www.flaticon.com/) from <https://www.flaticon.com/>
- Partnership: “Teamwork free icon” Icon made by [Freeipik](https://www.flaticon.com/) from <https://www.flaticon.com/>
- Sustainability: “House free icon” Icon made by [Freeipik](https://www.flaticon.com/) from <https://www.flaticon.com/>
- Thank you: <https://www.thebalance.com/sample-thank-you-notes-and-email-messages-2064015>
- Background: unsplash.com