University of Houston Libraries

Bridge2Hyku Toolkit: Developing Migration Strategies for Hydra-in-a-Box

Introduction

The University of Houston (UH) Libraries, in partnership with Indiana University (IU) and Indiana University-Purdue University Indianapolis (IUPUI), and in consultation with primary community stakeholders including Stanford University, DuraSpace, and the Digital Public Library of America (DPLA), seeks $249,103 in funding from an IMLS National Leadership/Project Grant to support the creation of the Bridge2Hyku (B2H) Toolkit. Focusing on general information and guides for migration as well as on specific content for migrating to the Hydra-in-a-Box platform, the toolkit will help institutions to better understand their digital library ecosystems and how they can prepare for migration. The toolkit will include (1) documentation on the theoretical approaches to migration, (2) instructions on how to conduct a needs assessment based on analyzing metadata structures and understanding system requirements, (3) best practices for preparing repository data for migration, and (4) specialized tools to assist users with migration to Hydra-in-a-Box.

Statement of National Need

Since the early 1990s, libraries and other cultural heritage organizations have integrated digital asset management systems (DAMS) into the suite of tools and services offered to their users. Defined by Richard Pearce-Moses as "software that supports the ingest, description, tracking, discovery, retrieval, searching, and distribution of collections of digital objects," DAMS have evolved over time as the technologies that support them have been refined and user needs and expectations have shifted. Not surprisingly, libraries have come to reassess, select, and migrate to a new DAMS based on the changes in technology and user needs. As Stein and Thompson[i] illustrate, this migration trends towards open source solutions, including DSpace, Fedora, Islandora, and Hydra. Furthering this trend, IMLS has invested funding in Stanford, the DPLA, and DuraSpace's collaborative effort to develop Hydra-in-a-Box.

While Hydra-in-a-Box will offer a robust, extensible platform for making digital objects accessible over the web, institutions cannot take advantage of the benefits the platform offers until they confront the difficulties of migrating legacy data from other systems. As the results of the Hydra-in-a-Box User Survey[ii] demonstrate, many institutions want to migrate from their current DAMS, but they lack the staff expertise and resources to plan and execute a successful data migration. These problems are compounded by a lack of general guides, tips or hands-on tools to assist libraries in completing the DAMS migration lifecycle. The UH Libraries, in partnership with IU and IUPUI, and in consultation with community stakeholders Stanford, DuraSpace, and DPLA, will leverage a National Leadership Project Grant to build the B2H Toolkit. This toolkit will fill current gaps in the migration process by helping libraries to understand their own digital ecosystem and offer them strategies and tools for migrating their digital collections from CONTENTdm to Hydra-in-a-Box and promoting them to DPLA. There is a potential for the toolkit to offer migration capabilities from other digital asset management systems in the future.

National Impact

Our project plan and outcomes aim to establish a framework for sustainable data migration. This includes understanding institutions' unique requirements, data models, and digital curation ecosystem as well as the software tools, existing and custom made, needed for data transfer. The B2H Toolkit will leverage IMLS' investment to build capacity for libraries and cultural heritage institutions in adoption of Hydra-in-a-Box nationally and worldwide. It will engage and strengthen the open source Hydra community around Hydra-in-a-Box by leveraging our collective expertise through strategic collaboration between UH, IU, IUPUI, Stanford, DuraSpace, and DPLA.

This project will deliver national impact by addressing the following IMLS performance goals:

(1) Broaden access and expand use of the Nation's content and collections, by helping libraries and cultural heritage institutions migrate their digital collections to the Hydra-in-a-Box platform and promote them to DPLA.
(2) Improve management of the Nation's content and collections, by assisting institutions in migrating to a community supported, nationally recognized digital platform, which promises better management and shared expertise of institutions' digital assets.

Project Plan

The UH Libraries will collaborate with partner institutions in the design and development of the toolkit. The toolkit will be co-maintained and managed by the UH Libraries, participating institutions and the Hydra community after release. The project will be divided into three phases. Phase one will focus on conducting a needs assessment. Activities in this phase include: analyzing partner institutions' metadata structures; understanding and documenting file types and content types; studying and determining ingest requirements and compatibilities for Hydra-in-a-Box; and establishing guidelines and recommendations for designing, developing and implementing a digital curation ecosystem. Phase two will be devoted to B2H Toolkit design and development, including: analyzing and documenting data modeling, migration lifecycle and data transfer path(s); defining and refining specific functionality and features for B2H Toolkit; mapping data elements from existing repositories to Hydra-in-a-Box; developing the B2H Toolkit software for migrating from CONTENTdm to Hyku; and testing and implementing B2H. The final phase is for the product release and outreach, including: communicating this toolkit to libraries and the cultural heritage digitization community via conference presentations, workshops, online and in-person training sessions with instructional materials; confirming sustainability of the toolkit; and improving functionality as needed.

Estimated Project Budget

The anticipated total amount requested from IMLS is $249,103. This amount reflects: (1) $120,000 for salaries and wages to develop the B2H Toolkit; (2) $35,155 for fringe benefits; (3) $25,000 for communication and conference travel to conduct presentations and workshops; and (4) $10,000 for professional development and training. Additionally, we seek $58,948 for indirect costs related to program administration, which are calculated based on UH's federally negotiated rate (31%).

Project Team

The UH Libraries will lead the project and conduct development of the toolkit as well as coordinate pilot testing with partners, IU and IUPUI. We'll solicit feedback on software framework and development from primary community stakeholders: Stanford University, DuraSpace, and DPLA. The key project team members from the UH Libraries include: Annie Wu (Primary Investigator), Santi Thompson, (Co-Primary Investigator), Andrew Weidner, Sean Watkins, Anne Washington, and the future toolkit developer hired with grant funding.

[i] Stein, Ayla, and Santi Thompson. "Taking control: Identifying motivations for migrating library digital asset management systems." (2015): https://uh-ir.tdl.org/uh-ir/handle/10657/1367
[ii] Hydra-in-a-Box User Survey Final Report (2015): https://purl.stanford.edu/jk292fy8802