Cloud-based environment for original metadata creation (WP1)

Work in Progress

- Cataloging library resources using Sinopia
- Sinopia User Group meets monthly

Completed Work

Work Cycles 1 and 2 are complete and Sinopia Release 2.0 is ready for use at http://sinopia.io. Here are background docs. More information available on the Sinopia help wiki: https://github.com/ld4p/sinopia/wiki

- Work Cycle 1: Project Plan for Stakeholders
- Work Cycle 2 docs: https://drive.google.com/drive/folders/1xkZ5V_ie6LDnE7cjHDFYh4eJ9_rNwA83

Preliminary work

- Requirements for and evaluations of form-based linked data editing tools (CEDAR, Library of Congress BIBFRAME Editor, VitroLib)
- Lessons Learned: VitroLib User Feedback and Usability Review

Summary and background

LD4P2 is building a cloud-based environment for the creation of linked data according to PCC guidelines, with technical development led by Stanford. Within this environment, a cataloger can create and edit data. By making use of the Questioning Authority (QA) service developed in the first grant, the cataloger will be able to search for and use additional bibliographic or authority data in various data stores such as the LC Name Authority File (LCNAF), RDA Vocabularies, the SHARE-Virtual Discovery Environment, or Wikidata.

There will be formal training in the use of the cloud environment as a production environment provided by the PCC/LC. The training will include the use of the editor to create bibliographic data that conforms to current cataloging rules (RDA).

The emphasis is on developing the skills and infrastructure to allow for implementation. Although the cloud environment created for the PCC will be supporting a single editor for managing scope of this project as well as ease of training, the data produced will be linked open data. Any metadata created or loaded into the data pool in the cloud environment will be freely available to anyone. In addition, if there are other institutions or projects interested in accessing and working with the data in the cloud environment, they will be allowed to do so as long as they can support any local needs such as the use of a local editor or the ability to ingest data into a local triplestore.

In support of Work Package 1, Cornell developers and UX designers will integrate authority lookup in external authority sources into the linked data environment using the caching and query services developed with Iowa (see Work Package 3: **Linkage to External Authorities and Web Context Data**). The Iowa services marshal data from multiple external providers, many of which do not provide public query interfaces, and provide the environment with these resources in a unified, consistent framework. This project will focus on cataloger-friendly user experience design to make the lookup process as convenient and efficient as possible. This work is critical to making sure that LD4P produces data with links across libraries and the open Web (i.e., shared URIs), and not just pools of locally connected RDF.

Harvard technical services staff in Metadata Management and Metadata Creation, well-positioned with official roles in PCC leadership, will drive the creation of widely adoptable transitional workflows for the library community. Coordinating with LD4 partners, Harvard staff will spearhead the development of policy decisions, metadata standards, and partnerships with expert metadata communities outside of libraries. Harvard staff will draw upon these workflows and policy decision efforts as it contributes to training and documentation for both tools and metadata best practices to support implementation in the cloud-based environment and inform environment tool development, as well as immediate, practical paths for moving libraries into linked data implementation in an incremental fashion (including the integration of URIs into MARC metadata) and building the skills needed for sustainable adoption of new tools and workflows.

The critical factors for this portion of the grant to succeed will be: the successful implementation of a flexible linked data creation and editing tool for the cloud-based environment; training for early adopters in tools and linked data production workflows; implementation policies on acceptable data creation established by the PCC; and, a sufficient number of implementers to make the transition significant.

The development of the cloud environment, its maintenance and support, and the iterative enhancement to the included editing tool will be supported by the developers and UX Designer at Stanford. Integration of the external authorities lookup will be supported by the Cornell developers and UX Designers. Rigorous testing of the environment and implementation of production workflows together with coordination of iterative policy development and metadata training activities with the PCC will be supported by the PCC leadership based at Harvard.

- See also Linkage to external authorities and web context data (WP3)