Interim Performance Report 4
1 April 2009 – 30 September 2009

Submitted by:

Carole Palmer, Principal Investigator
clpalmer@illinois.edu

Katrina Fenlon, Project Coordinator
(217) 244-7809
kfenlon2@illinois.edu

October 2009

University of Illinois at Urbana-Champaign
General Project Activities

**Project Staffing**

In late May, a web_interaction designer, Sunah Suh, was hired for the duration of the summer to assist with interface development. A new project coordinator, Katrina Fenlon, was hired in August, replacing Amy Jackson. Also in August, graduate research assistant Jacob Jett joined the project to assist with the IMLS DCC Flickr Feasibility Study. Research programmer Hong Zhang went from full- to half-time; a search for a supplementary research programmer is underway.

**Infrastructure Subgroup**

A new subgroup was designated to set priorities and assist in coordination of the project’s technological infrastructure maintenance and development. The Infrastructure subgroup, composed of principal investigators and research programmers, met weekly beginning in July to discuss overarching, infrastructure-related challenges and to set project priorities. Since most solutions to infrastructure concerns are not within the scope of the currently funded activities, a thematic subset has been set up to support the advanced development that cannot be sustained in the growing IMLS DCC and Opening History databases.
**Digital Collection Evaluation (DiCE) subgroup**

The DiCE subgroup (Carole Palmer, Oksana Zavalina, Jacob Jett, Richard Urban, Amy Jackson and Katrina Fenlon) met regularly to coordinate the addition of new collections and the interactions with new content providers; to discuss development and maintenance of the DCC and Opening History portals; to evaluate evolving subject strengths in both portals; and to manage the IMLS DCC Flickr Feasibility Study.

**Collection Development**

Figure 1 represents the growth of the IMLS DCC and Opening History collection registries since “Next Generation Digital Federations” began in 2007. During this performance period, the Opening History collection registry grew by approximately 40%, a dramatic increase spurred by DiCE’s active recruitment of state library collections. State libraries from Arizona, Montana, New York, and Oklahoma contributed collections to Opening History this performance period, increasing the total number of collection records to 736 and item records in the repository to 1,028,454. The IMLS DCC collection registry and item repository also grew significantly.

![Figure 1. Growth of IMLS DCC and OH Collection Registries, 2007-2009](image)

**Collection Evaluation**

In accordance with a project objective to “Conduct a formal evaluation of the IMLS DCC content and, based on the outcomes, expand the collection for targeted scholarly communities” (from “Next Generation Digital Federations” proposal), recent collection development prioritized collections featuring subjects identified in last spring’s collection evaluation as emergent subject strengths of Opening History. Most of Opening History’s subject strengths were amplified this performance period, as charted in Figure 2. In anticipation of creating a subject-specific portal for a targeted scholarly community, DiCE developed an experimental database focused on...
transportation history, one of the recognized Opening History subject strengths. A subset of Opening History, the experimental transportation database currently contains 22 collections (120,360 items). The database will serve as a testbed for both the interface subgroup’s prototyped interface redesign (discussed under “Interface subgroup”, below), and for future evaluations of portals targeting specific scholarly communities.

![Figure 2. Performance period growth in top 10 Opening History subject strengths](image)

**Flickr Feasibility Study**

The first phase of the IMLS DCC Flickr Feasibility Study began in July. While the project awaits admittance into Flickr Commons (projects generally experience a several-month wait for a Flickr response after registration of interest), the project team decided to move forward outside of the Commons on Flickr, intending to migrate under the Commons umbrella when possible. (Other institutions, such as the U.S. National Archives, have also found it necessary to use this approach. See the National Archives’ Flickr photostream at [http://www.flickr.com/photos/usnationalarchives/](http://www.flickr.com/photos/usnationalarchives/).) In the meantime, metadata-related and collection development decisions are being made with the pending transition in mind.

DiCE, which serves as the Flickr project team, began negotiations with two collections that meet the project’s objective of targeting IMLS DCC collections that demonstrate variety (of host-institution type, collection scale, photographic content, and technological platform):

- The Charles W. Cushman Photograph Collection, a large collection of color photographs hosted at Indiana University
- The Charles Overstreet Collection, a smaller collection of black-and-white photographs at Flora Public Library (Flora, IL).

The discussions that led to these institutions’ participation in the Flickr project have been documented to contribute to the case reports covering the Flickr development process and outcomes.
DiCE met regularly to design the item- and collection-level metadata records to display with IMLS DCC resources contributed to Flickr. This discussion was opened to a larger group at the September 23 Metadata Roundtable session (slides and audio from which are available at the roundtable website: http://www.isrl.illinois.edu/~dcc/mdrt.html). In making metadata decisions, DiCE has sought to meet IMLS and host institutions’ criteria for branding and permissions, while also adequately providing informative descriptions to appeal to regular Flickr users. At the close of the performance period, an informal schema for Flickr metadata was prepared for approval by host institutions.

Jacob Jett worked with project programmers to refine procedures for uploading metadata-rich content to Flickr. At the end of the performance period, the initial upload was scheduled for October. A baseline measurement of IMLS DCC collection use has been captured for the forthcoming evaluations of use.

**Interface subgroup**

The interface subgroup (Sarah Shreeves, Michael Twidale, Tim Cole, Richard Urban, Sunah Suh, Amy Jackson, Hong Zhang, Jacob Jett, Oksana Zavalina) prototyped an experimental interface designed to improve representation of content and context for scholarly use of IMLS DCC and Opening History collections. Figure 3 shows a draft of the redesigned portal page, which will be tested on the aforementioned transportation database. Among other improvements, the experimental interface attempts to:

- Enhance the currently text-heavy interface with browseable visualizations that immediately reveal the size and scope of the aggregation to users;
- Integrate facets for search refinement;
- Achieve tighter integration between collections and items and better contextualization of search results;
- Enable seamless switching between search and browse, for a more conveniently interactive finding experience.
Continuing refinement of this experimental interface is a primary concern of the interface subgroup, which at the close of the performance period began preparing to conduct usability studies to test interface features.

The experimental interface’s visual elements are based in part on the interface subgroup’s research into prototyping a Collections Dashboard (Figure 4), aimed at providing users with immediate, low-effort understanding of collection and item contexts and relationships. Michael Twidale and Richard Urban conducted participatory design sessions for the Collections Dashboard at several venues in the spring and early summer (see Presentations). Hong Zhang successfully tested elements of the Collections Dashboard on data of the Cushman photograph collection, suggesting that elements of the dashboard may be integrated into the experimental interface.
The interface subgroup also implemented COinS (ContextObjects in Spans: a simple way of embedding citation metadata in web pages) on item-level and collection-level records, making IMLS DCC and Opening History metadata available to Zotero.

**Collection and Item Metadata Relationships subgroup (CIMR)**

In March 2009, the CIMR research group, lead by Principal Investigator Allen Renear, began to lay the foundations for a testbed to evaluate rules based on CIMR categories and metadata from the Opening History aggregation (Renear, et al. 2008). Early on in the process, we decided that taking advantage of existing Semantic Web technologies would allow us to construct a testbed using experimental, but off-the-shelf, tools. Not only has this expedited our progress towards a working testbed, but it will also serve as an adoptable model for related aggregation projects. During this reporting period we have successfully implemented the major pieces of that testbed, thanks to the help of GSLIS research programmers Larry Jackson and Amit Kumar.

**CIMR Testbed Implementation:**

- Characterized temporal properties of *Opening History* item-level metadata based on recommendations from the Digital Collections Evaluation Research Group;
- Selected 25 collections that exhibited metadata properties suitable for testing temporal relationships;
- Developed an OWL ontology to represent the isGatheredInto relationship between Collection-level metadata and item-level metadata;
- Defined additional CIMR properties to record normalized values for dc:date and dc:coverage elements. These values will be generated using the California Digital
Library's Date Normalization Utility (CDL-DNU) (http://www.cdlib.org/inside/diglib/datenorm/);

- Developed XSLT transformations to convert collection-level and item-level OAI-PMH records into RDF that conforms to our OWL ontology. These transformations were developed to work with the MIT SIMILE RDFizer (http://simile.mit.edu/wiki/RDFizers);

- Developed a workflow toolkit to ingest RDF into OpenRDF Sesame triple store (http://www.openrdf.org/).

With these components in place, CIMR doctoral researchers, Richard Urban and Karen Wickett, are moving ahead on developing SPARQL queries that test basic rules for attribute-value propagation, value propagation and value constraint.

Citations to CIMR research


Dissemination and Publicity

Presentations, April 2009 – September 2009

- Museums & the Web 2009, Indianapolis, IN - Patchwork Prototyping a Collections Dashboard (Urban, R., Twidale, M., Jackson, A.)
- HASTAC 2009 - Patchwork Prototyping a Collections Dashboard
- THATCamp 2009, Fairfax, VA - Patchwork Prototyping a Collections Dashboard
Upcoming/Submitted Presentations and Publications

- Richard Urban will represent the IMLS DCC project at the Chief Officers of State Library Agencies Fall Meeting in October, presenting to the Networking Committee and General meetings.
- Accepted: Adamczyck, P., Urban, R. & Twidale, M. Building and Developing Collections Dashboards, Museums & the Web 2010 (Workshop).

Sowing Culture Blog

Sowing Culture (http://sowingculture.wordpress.com/) continues to highlight Opening History resources. Since April 1, 2009, 51 new blog posts featured more than 60 digital items from 41 collections: photographs, textual items (newspaper clippings, books and pamphlets), posters, sheet music, video, and audio files. The popularity of the blog has blossomed this performance period, as shown in Figure 5, generating traffic for collections’ homepages and the Opening History portal. Sowing Culture also disseminates via Twitter and currently has 204 followers (http://twitter.com/sowingculture).

Blogs and other resources aimed at archivists and historians have begun pointing readers to Sowing Culture. For example:

- One commenter’s response to the question "What is your favorite history-, library-, or archives-related blog?" asked by the NARAtives (National Archives) blog on September 16 lists Sowing Culture as the first one in the list of 4 blogs where the respondent "always finds something useful or interesting" (http://blogs.archives.gov/online-public-access/?p=144).

Figure 5. Increase in monthly views of Sowing Culture blog since March 2009
• Tulane University's Louisiana Research Collection's blog lists Sowing Culture among the "blogs we follow" (http://manuscriptstulane.blogspot.com/).

• Archives Next blog by Alaska state archivist Kate Themier listed Sowing Culture as one of her favorite archives-related blogs on August 31, 2009 (http://www.archivesnext.com/?p=339).

• Indicommons mentioned Sowing Culture as “an interesting blog on history that we ran across this week when they linked to our Morse post” on May 1, 2009 (http://www.indiccommons.org/2009/05/01/carnival-of-the-commons-muse-awards-and-more/).

References and Citations to IMLS DCC publications and resources


• AfriGeneas Genealogy and History Forum lists IMLS DCC aggregation among digital collections that make African American documents available online, on October 19, 2009 http://www.afrigeneas.com/forum/index.cgi?md=read;id=81571.


• Library Usability Links 5/12/09 on A Digital Outrigger refers to IMLS DCC collection dashboard work (http://librarydigitalprojects.com/2009/05/).

• Jeannine Mjoseth refers to IMLS DCC aggregation on Wall Street Journal Community Groups site under the "What are the best online archives for historical documents, art and artifacts?" topic in her detailed post.
Related Activities

IMLS DCC continued to hold the Metadata Roundtable research study group two to four times per month. Recent roundtable topics have included Roles and Competencies of Metadata Librarians; Neuroscience Information Framework Resource Registry and Ontology; and PBCore Metadata Scheme Elements. The website, which includes a full listing of the Metadata Roundtable topics and background readings, can be found at: http://www.isrl.uiuc.edu/~dcc/mdrt.html.