

The DPLA's Content Infrastructure and the Digital Hubs Pilot Project

DPLA Midwest

October 11-12, 2012

Chicago Public Library

DPLA Director for Content

In September 2012, the DPLA hired the first employee of the new independent organization: Emily Gore, Director for Content. In this role, Gore will spearhead the development of the DPLA's content infrastructure, including implementation of state and regional digital service pilot projects as part of the Digital Hubs Pilot Project. Gore will also oversee content provider relationships with digital library partners throughout the United States.

Digital Hubs Pilot Project

Under the Digital Hubs Pilot Project, launched in late September 2012, the DPLA will undertake the first effort to establish a national network out of the over 40 state or regional digital collaboratives, numerous large content repositories, and other promising initiatives currently in operation throughout the US, bringing together myriad digitized content from across the country into a single access point for end users. The approach is to work with 5-7 states or regions (Service Hubs) and a similar number of large content providers (Content Hubs) to aggregate content on a pilot basis.

The object of the Hubs Project is to strengthen and connect existing state and/or regional infrastructure to create a system of Service Hubs from which aggregated data from libraries, museums, historical societies and archives are harvested. Each Service Hub will offer a full menu of standardized digital services to local institutions, including digitization, metadata consultation, data aggregation and storage services, as well as locally hosted community outreach programs bringing users in contact with digital content of local relevance. The two-year Hubs Pilot aims to help existing state programs offer these services to all institutions in their state or region. Service Hubs will serve as an on-ramp for every institution in a pilot state or region to participate in the DPLA network. If this model proves successful, the DPLA would like to implement it in all US states or regions.

Service Hubs initially identified for the pilot are:

- Mountain West Digital Library (Utah, Nevada and Arizona)
- Digital Commonwealth (Massachusetts)
- Digital Library of Georgia
- Kentucky Digital Library
- Minnesota Digital Library
- Oregon Digital Library
- South Carolina Digital Library

In addition to Service Hubs, Content Hubs are those existing large digital libraries that will have a one-to-one relationship with the DPLA. Large digital content producers like

the National Archives and Records Administration (NARA), Harvard, and others will work with the DPLA directly to identify and prepare their collections for aggregation by the DPLA. Content Hubs that will participate in the initial launch of the DPLA are currently being identified with anticipation of many more to follow.

The Digital Hubs Pilot Project is supported by the National Endowment for the Humanities¹ and the Institute of Museum and Library Services.²

Data Provider Agreements

In preparation for their August 2012 workshop, the Content & Scope Workstream developed draft data exchange agreements governing the provision of metadata to the DPLA by DPLA Content and Service hubs. These agreements, modeled after the various Europeana Data Exchange Agreements, include a summary,³ a legal agreement,⁴ and metadata requirements.⁵ These agreements are currently being finalized and will govern the relationship between DPLA and the Content and Service hubs. These agreements are expected to be in place by the end of 2012.

Technical considerations

A metadata schema⁶ has been drafted for the Hubs Pilot Project that is currently being reviewed by experts at each Hub site and by the larger DPLA community. Metadata from the Hubs will be harvested using OAI-PMH along with other techniques, mapped to the DPLA schema, and made available through the DPLA portal and as a dataset via the API. Front-end tools to access the content are being developed as part of hackathons and other venues, and a front-end portal to the DPLA is being developed by iFactory, a Boston, MA-based web design and development agency. As part of the initial launch in April 2013, the DPLA and its Hubs partners plan on developing virtual exhibitions via the Omeka platform which feature content from the Hubs' digital collections.

¹ See: <http://dp.la/2012/07/26/national-endowment-for-the-humanities-announces-award-to-support-development-of-dpla/>

² See: <http://dp.la/2012/09/13/imls-award/>

³ See: Data Provider Agreement (Summary): <https://docs.google.com/document/d/1d4W3HZmivDunrhSesaAebmZ-3SMxJ-eNN4rOuhNBGXo/edit>

⁴ See: Data Provider Agreement (Legal Version): <https://docs.google.com/document/d/1Gtt3E1eI-fc4xEnR2JWS9MGWZjWIL-d3hZI-yaVG-Sg/edit>

⁵ See: DPLA Metadata Requirements: <https://docs.google.com/document/d/1uZNX6S4FtxW7LTRyxAvgn5yGAjJqk7T3kWRysZ6cj3E/edit>

⁶ See: DPLA Dublin Core Application Profile: <http://dp.la/wiki/sites/dpla/images/MetadataSchema.pdf>