

Review of the Digital Public Library of America (DPLA)
Peter Sutton

The Digital Public Library of America (DPLA) is the recently initiated and ongoing realization of a vision that has been gaining momentum among archivists, scholars, and others since the early 1990's. This vision was to create a single digital platform that provides access to the digital collections from all libraries, archives, museums, and other cultural heritage institutions across the U.S. There are many such institutions that have born-digital or digitized materials available to the public, but these collections are generally isolated from each other. Building on previous efforts to pull such materials together, the DPLA brings all of these collections into a single interconnected network, a single platform and portal for clear, free, and open access for the general public. "The Digital Public Library of America (DPLA) brings together the riches of America's libraries, archives, and museums, and makes them freely available to the world. It strives to contain the full breadth of human expression, from the written word, to works of art and culture, to records of America's heritage, to the efforts and data of science."¹

The DPLA was planned and designed so that similar projects in other countries and regions would be interoperable, most notably the Europeana Data Model (EDM), which is the European version of the DPLA. The DPLA also provides an open environment where developers and users are able to get the Application Programming Interface (API), bulk data, and a growing range of apps for greater interoperability and innovation; the DPLA encourages community involvement and input to build something that serves as many users as possible. The DPLA's stakeholders are cultural heritage institutions that have or want digital holdings for their users, but there is potential for the DPLA's programming specs and applications to serve a much wider audience (either directly or as a model).

The DPLA was first planned in December 2010 through a gathering at the Berkman Center for Internet and Society at Harvard University. The Alfred P. Sloan Foundation sponsored the event that was attended by a number of leaders in cultural heritage institutions, technology projects, and academia. In October 2011 the Berkman Center and various collaborators began scoping, designing, and constructing the DPLA. After 18 months, in April 2013, the DPLA was publically made available. The Executive Director Dan Cohen and a diverse Board of Directors currently lead and plan the DPLA.

A mix of governmental and private institutions with interests and experience in various aspects of the aggregated collections has funded the DPLA, most notably the Alfred P. Sloan Foundation, Arcadia Fund, Institute of Museum and Library Services, John S. and James L. Knight Foundation, and National Endowment for the Humanities. The Sloan Foundation was the original funder and has remained a featured partner throughout the DPLA's development and roll out. The DPLA also receives support of a different kind by encouraging public involvement and innovation in regards to its API and apps. It hosted an APPFest in 2012 and will hold the first Annual DPLAfest in October 2013, completely free and open to the public with the goal of encouraging

¹ "FAQ". *The Digital Public Library of America*. Web. Last Accessed 9/26/2013.
<http://dp.la/info/about/faq/>.

development of new applications, uses, workshops, and thought about the DPLA and digital repositories in general.

The DPLA serves as a single point of access to the vast digital collections freely available in the U.S. at cultural heritage institutions. It does not hold physical copies of these materials in its database, but instead aggregates metadata and links to the content. The compilation of metadata occurs through a series of content and service hubs throughout the country. Content hubs are large digital libraries, archives, museums, or repositories that agree to a one-to-one relationship with the DPLA and submit, edit, and maintain at least 250,000 unique records that resolve to digital objects. Service hubs are regional or state digital libraries that aggregate information about digital objects within their assigned regions or states, while offering services such as digitization, storage, and metadata. They also serve as points of community outreach for raising awareness and use of digital content in local or regional cultural heritage institutions. Service hubs focus on digitization of existing content, but also offer services for born-digital material as well. Content hubs are large digital repositories that submit and maintain their own records, while service hubs offer support and aggregate data about digital objects and metadata for cultural heritage institutions in their given region or state. There are over 450 partners in these hub programs.

Collecting, ingesting, and storing metadata are only part of the DPLA's mission. It also works to make sure that this material is widely available to the general public. To this end, the DPLA has created its own API (Application Programming Interface). This API has a number of philosophies at its core, namely deep data structure, building on what came before, and openness. The DPLA aims not only at providing access to the scientific and cultural record, but also at stimulating innovation and creating new digital access mechanisms through its openness. The API and its data are available to anyone at <http://dp.la/info/developers/codex/>. Rather than reinventing the wheel and potentially blocking support and innovation from outside institutions, the DPLA has incorporated widely used hardware and software standards along with open source technologies. The API and metadata schema support a deep data structure, accommodating the concept that any resource can have multiple representations at varying levels, from physical manifestations to digital formats. The DPLA focuses on information about the digital representation of the resource (top-level) then works its way down to physical representations (Second Level), to more user-familiar descriptive metadata such as subject and description.

The DPLA's general metadata policy reflects its commitment to "make the cultural and scientific record available, free of charge, to all through databases of metadata. The DPLA, for this purpose, has undertaken the task of ingesting, indexing, enriching, and making available descriptive metadata and wishes to make such metadata widely available,"². All metadata in DPLA databases have no copyright or have had that right waived by either the DPLA or its numerous partners. In addition to the metadata schema outlined above, the DPLA offers applications and tools to better use its metadata while adhering to the letter and spirit of U.S. copyright law. For instance the DPLA outlines how metadata is limited in its expression to non-original objective facts about objects as well as in its grammar to the degree that it is a merger, something that is

² "The Digital Public Library of America Policy Statement on Metadata". *The Digital Public Library of America*. Web. Last Accessed 9/26/2013.

severely limited by the number of ways an idea can be expressed to the point where the idea merges with the expression; this frees metadata from the protection of copyright law and its protection of original expression. Even if some of the metadata records provided do fall under copyright, in order to have them fully aggregated, the DPLA asked that the right be waived under the Creative Commons Zero (CC0) Public Domain Dedication. The ultimate purpose of the DPLA databases is free and unencumbered access to metadata. They provide a space where “users are free to harvest, collect, modify, and/or otherwise use any metadata contained in the DPLA”³.

Invoking the CC0 Public Domain Dedication means that the metadata can be used without attribution. The DPLA encourages users to give such attribution anyway.

... in the spirit of building a community of practice that recognizes contributions, the DPLA believes that giving attribution to all the sources of the metadata, including the Data Provider (cultural heritage institution that owns the object shared) and the Data Aggregators (institutions that aggregate metadata from multiple cultural heritage institutions, e.g., The Digital Library of Georgia, as well as the DPLA) would be helpful and appropriate given the goals of the DPLA and its partners.⁴

In addition to recommending attribution, the DPLA also offers guidelines to keep data current and avoid misrepresentations or misattribution in records, as well as a disclaimer of their risk. It publishes metadata as is and assumes no responsibility for outside use of published data. The best practices guidelines reinforce the DPLA’s commitment and mission to provide an open and supportive environment for the nation’s scientific and cultural record that will facilitate innovation and participation by both content providers and users.

The metadata profile for the DPLA is unique, but builds heavily on the European Data Model (EDM) experience and input from hub providers. Both EDM, the European version of the DPLA, and the DPLA metadata schemas allow for integration of provided metadata, while at the same time respecting the original institutional description. The DPLA’s metadata profile combines a number of the most interoperable and open metadata standards including Dublin Core namespaces, RDF (Resource Description Framework and Schema - the key to the semantic web), OAI (Open Archives Initiative) object reuse and exchange, and a very simple and a mappable authority list, Basic Geo Vocabulary.⁵ There are 6 total classes in the DPLA schema, with three core classes, `edm:WebResource`, `ore:Aggregation`, and `dpla:SourceResource`, along with three other classes, `dpla:Place`, `edm:TimeSpan`, and `dcmitype:Collection`.

Ideally, the DPLA would like to aggregate all digital holdings in every cultural heritage institution in the country and have access to similar projects in other countries or regions. The current reality is a little different. The DPLA is only in the first stages of achieving this goal. Since being rolled out in April 2013, the DPLA has aggregated over 4 million records and has 7 digital collections featured including “Activism in the U.S.”, “Leaving Europe: A new life in America”, and “Boston Sports Temples”. The DPLA

³ “The Digital Public Library of America Policy Statement on Metadata”. *The Digital Public Library of America*. Web. Last Accessed 9/26/2013.

⁴ “DPLA Data Use Best Practices”. *The Digital Public Library of America*. Web. Last Accessed 9/26/2013.

⁵ “Metadata Application Profile”. *The Digital Public Library of America*. Web. Last Accessed 9/26/2013. <http://dp.la/info/developers/map/>.

currently has several search options: subject, map, and timeline. There are several applications available on the DPLA website and more available if one searches. The DPLA is also trying to offer digital services as part of the FCC E-Rate program that does such things as provide high-speed internet content to modernize schools and public libraries.

Reviews of the DPLA have generally been positive, with the different search capabilities being considered particularly exciting.⁶ It is still in its trial phase and only has 11 content hubs, ARTstor, Biodiversity Heritage Library, David Rumsey Map Collection, The Harvard Library, HathiTrust Digital Library, National Archives and Records Administration, New York Public Library, The Smithsonian Institution, The University of Illinois at Urbana-Champaign, University of Southern California, and The University of West Virginia, and 6 service hubs, Digital Commonwealth, Digital Library of Georgia, Kentucky Digital Library, Minnesota Digital Library, Mountain West Digital Library, and South Carolina Digital Library. There is still a lot of work to be done, but the content hubs have increased to 11 from the initial 6. Its ultimate goal is a minimum of 40 state/ or regional service hubs, with many more content hubs. With its combination of digital services, strategic partnerships with some of the larger library systems, museums, and archives in the country, metadata aggregation, and policies of openness to foster community involvement and innovation, the DPLA is positioning itself to be a leader, role model, facilitator, and counter to corporate distribution models. The DPLA is still in its infancy, but even if it fails it lays the groundwork for similar future work.

⁶ “I love the concept and how there are APIs for other institutions to develop their own apps using DPLA’s content. So Great!” said Jen Cwiok, digital projects manager for the American Museum of Natural History. “I think that the site is well organized and well designed. I immediately wanted to dive in and play around with it. The map and the time line are innovative methods for retrieving and displaying data. The browsable subjects were also helpful in determining what content is like.” Page 32. Eylene, Matt. “The Buzz Was Mostly Positive After the Digital Public Library of America’s Launch in April”. *Library Journal*. Vol. 138 No. 09. (2013): 30-32. *Lexis Nexis*. Web. Last accessed 9/27/2013.

Bibliography/Webography

Anne K. Library Discount Network Abate. "Building the Digital Public Library of America : The Hubs Pilot Project". *Against the Grain*. Vol. 25 No. 3 (2013): 58-59. *EBSCO Host*. Web. Last accessed 9/28/2013.

Cottrell, Megan. "A Digital Library for Everybody". *American Libraries*. Vol. 45 Issue 3 / 4 (2013). *EBSCO Host*. Web. Last accessed 9/28/2013.

"Digital Public Library of America (DPLA) Launches Today". *D-Lib Magazine*. Vol. 19 No. 5/6 (2013): 7-7. Web. 1 Oct. 2013. Last Accessed 10/1/2013.

DPLA. Digital Public Library of America. Web. 1 Oct. 2013. Last accessed 10/1/2013.
dp.la

Dillon, Cy. "Planning the Digital Public Library of America". *College & Undergraduate Libraries*. Vol. 19, Iss. 1 (2012): 101-107. *tandfonline.com*. Web. Last accessed 9/27/2013.

Eyline, Matt. "The Buzz Was Mostly Positive After the Digital Public Library of America's Launch in April". *Library Journal*. Vol. 138 No. 09. (2013): 30-32. *Lexis Nexis*. Web. Last accessed 9/27/2013.

"National Archives to Help Launch Dpla Pilot Project". *Advanced Technology Libraries*. Vol. 42 No. 4 (2013): 3-4. *EBSCO Host*. Web. 1 Oct. 2013. Last accessed 10/1/2013.

O'Leary, Mick. "The Digital Public Library of America Opens Its Doors". *Information Today*. Vol. 30 No. 7 (2013): 20-21. *EBSCO Host*. Web. 1 Oct. 2013. Last accessed 10/1/2013.

Peek, Robin. "Focus on Publishing". *Information Today*. (2002): 24. Web. Last accessed 9/27/2013.

Rodzvilla, J. "Digital Public Library of America." *CHOICE: Current Reviews for Academic Libraries*. Sept. 2013: 37. *Academic OneFile*. Web. 1 Oct. 2013. Last accessed 10/1/2013.