2013

# Digital Public Library of America

Clem Guthro
*Colby College*, guthroc@hawaii.edu

# Digital Public Library of America

By Clem Guthro

Visionaries, pundits, cynics, and ordinary citizens have, over the years, waxed eloquent over the idea of a digital library that would make all knowledge accessible. H.G. Wells in his collection of essays *World Brain* envisioned a universal comprehensive and accessible "encyclopedia" that encompasses "all that is thought and known" (Wells 1938: 78). Likewise, Vannevar Bush (1945) envisioned his Memex machine as a personal library that certainly approximates much of what the Internet has provided. Major players such as the Library of Congress, Internet Archive, members of the Association of Research Libraries, and various state libraries and cultural organizations have digitized books, photographs, sound records, and films from their collections and have built "digital libraries" on a small scale.

Private enterprise entered the digital library picture when Google announced in December 2004 that it had signed agreements with major universities and libraries (Stanford University, the University of Michigan, Harvard University, the New York Public Library, and Oxford University's Bodleian Library) to scan all or part of their collections. Although this agreement was greeted with skepticism by many (Anthony 2005), it provided the first glimmer of hope that a large-scale digital library might indeed be possible.

Because Google intended to scan in-copyright publications, publishers and authors sued Google over copyright infringement. The Google Lawsuit, as it was commonly known, and its proposed settlement raised another round of discussions over the digital future. Many thought that the settlement provided promise for libraries and the public (Trachtenberg and Vascellaro 2008), while others worried about the

negative consequences of relinquishing stewardship to a monopoly such as Google (Darnton 2009). Judge Denny Chin disavowed the settlement Google had struck with authors and publishers due to copyright violations and unfairness. The decision opened the door for the emergence of the Digital Public Library of America or DPLA as it is commonly called (Darnton 2011).

In October 2010, Robert Darnton, University Librarian at Harvard, convened a meeting at the Radcliffe Institute for Advanced Study, which brought together key leaders from libraries, research institutions, cultural heritage organizations, government, and foundations to consider the possibility of building a national digital library. Using a grassroots approach, funding from the Alfred P. Sloan Foundation, and a hosting commitment from Harvard's Berkman Center for Internet and Society, the DPLA kicked off a two-year planning process with the goal of having a formal organizational structure, content, and a working prototype by early 2013.

A steering committee and secretariat were established to move the work forward in a rapid but thoughtful and coordinated manner. Chaired by John Palfrey, then at Harvard Law Library, the steering committee included luminaries from the academic and public library sphere, think tanks, and foundations. (A complete list of the steering committee can be found at http://cyber.law.harvard.edu/research/dpla/steering). The secretariat, a small group from the Berkman Center, under the leadership of Maura Marx, was charged with the day to day operations and coordination of overall activities and workstreams.

"Workstreams" are broadly defined as an area of business or project activity that needs to be explored and managed. DPLA chose six workstreams: audience & participation, content & scope, financial/business models, governance, legal issues, and technical aspects. Co-chairs and appointed participants defined the work, which was supplemented by a public wiki and virtual participation from the wider public. DPLA also used three plenary sessions that brought together hundreds of librarians, scholars, industry experts, copyright lawyers, technologists, and students to talk, plan, and provide input. These sessions also served as a venue to announce major funding from the Institute for

Museum and Library Services, National Endowment for the Humanities, and the Knight Foundation.

Though it is tempting to see digital libraries predominantly as technology projects, the DPLA is that and much more. The DPLA sees itself as consisting of five major areas: code, metadata, content, tools and services, and community.
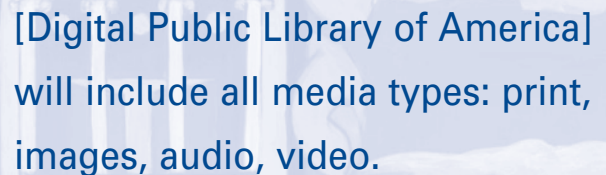
The code and the requisite technology are the underpinning of the DPLA. The code is being built on the principle of openness and interoperability. As much as possible free and open source code is used, and any code developed by DPLA will be released as open source and contributed back to the larger developer community. Metadata is the descriptive data about a person, place, thing, or event, which comes from library catalogs. DPLA will aggregate existing metadata from libraries, archives, and museums to enable users to search and find collections and individual items.

As with any library collection, content is king. DPLA will include all media types: print, images, audio, video. DPLA is beginning to gather content by working with a number of service hubs, mainly state and regional library groups, to aggregate content and to show the public a taste of what is possible. Other service hubs along with large content creators such as Harvard and the Hathi Trust will be added. In beginning most materials will be in the public domain, but DPLA will investigate options for orphan works (works still in copyright, but where copyright owners cannot be found), materials that are in copyright, and e-lending models. It is the intention to make DPLA widely and feely accessible "with no restrictions… with use and reuse governed only by public law" (http://dp.la/about/elements-of-the-dpla/).

Tools and services are critical to the life, nurture, and future of the DPLA. DPLA intends to provide more than content and an interface, but also a robust set of tools and services that will allow users, programmers, and other members of the community to use the content in new and interesting ways and to build additional tools and services that will further its work, reach, and influence. DPLA will make its own tools available in an open manner for reuse and extension and for commercial purposes. It encourages others who develop tools and services built on the DPLA platform and using DPLA content to make these available for others to use.

DPLA is being built by and for the community writ large: library and cultural heritage community, general public, and private and commercial developers. The platform will be participatory to allow the community of users to exploit the rich content and functionality for their needs. DPLA intends that the ongoing development and support of the platform and the DPLA initiative will be community driven.

The creation of the DPLA calls to mind that the United States is late to the game in terms of a national digital library. Major digital libraries exist in France, the Netherlands, Norway, South Korea, and Australia (Singer 2011). In 2008, the European Union created Europeana (www.europeana.eu/portal/), an overarching digital library of European cultural heritage. While others may decry our slow start out of the gate, we have much to learn from the experience of others. Europeana, which is approximately five years ahead of DPLA in terms of development of technology and policy, has been a major help in DPLA's development. DPLA is using lessons learned by Europeana to move forward quickly in terms of local development. It is also building on the openness of Europeana and building in interoperability between the two systems. More importantly DPLA and Europeana have already been building collaborative exhibits that will span the collections of both libraries. The first such exhibit "Leaving Europe: A new life in America" (http://exhibitions.europeana.eu/exhibits/show/europe-america-en) was launched in December 2012. On April

> [Digital Public Library of America] will include all media types: print, images, audio, video.

18–19, 2013, DPLA will hold its official launch at the Boston Public Library. At this point it will move from dream to the first phase of reality. Getting to this phase is a monumental achievement and one to be lauded and celebrated, but it is certainly just the beginning. The challenges will be predominantly funding and policy issues, neither of which are insignificant in

today's economic and rancorous political climate. Fortunately the library community, which is solidly behind this project, is tenacious, hardworking, collaborative, and pragmatic, all desirable qualities for building DPLA's future.

## POLICY IMPLICATIONS

Moving from concept to reality for the DPLA is not primarily a technological issue, but one of copyright and intellectual property (Carr 2012). The public policy implications of copyright and intellectual property in the digital age are not merely legal concerns but important social and economic ones. Legislators wrestle with competing claims of content producers (e.g., authors and publishers), scholars, educators, and the general public on the use and remix of content.

The vision of the DPLA is broader than a collection of public domain works or works with a Creative Commons license that provides for access and use. Because the "right to digitize" resides with the content creator, the challenge of building a robust digital library is not insignificant. Orphan works is an area that is gray at best, and one that needs some serious work by the U.S. Copyright Office and Congress. DPLA is interested in helping make serious progress on the orphan works issue. If this could be addressed in a manner that favored libraries, a significant corpus of works published in the 20th century in the U.S. could be made available. Much of born digital content is governed by license agreements which usually trump copyright law, especially in terms of "fair use" exceptions. (The term "born digital" refers to content created originally in digital format.) DPLA will need to be creative in approaching how "licensed" materials might be included.

### *Whither Maine?*

Although the policy implications outlined above are no different for Maine, the issues facing Maine are matters of public policy around universal access of its citizens to the digital future. There are, perhaps, three major areas where current policy will affect whether Maine will be able to participate in DPLA on equal footing with other states.

A robust, high-speed, broadband digital infrastructure is necessary for the state's economic well-being, which has not been a strategic priority of state government. Most broadband access is provided by commercial companies, which disenfranchises many poorer people in Maine. As digital content is increasingly high bandwidth-dependent, a robust infrastructure must become a strategic priority for the state.

Second and more specific to the DPLA, there is a lack of digital and administrative structure and funding. The DPLA plans to work with state or regional hubs to harvest aggregated metadata and/or content. Maine is one of 10 states that have no statewide "digital library." Elements of it exist in the work of the Maine Historical Society, Maine InfoNet, and a variety of individual small digital library initiatives at the University of Maine, University of Southern Maine, Colby, Bates, Bowdoin, Bangor and Portland public libraries, and the Maine State Library. There is, however, no overarching technical and administrative infrastructure to draw these projects together in a coherent whole that could provide the genesis of the "Digital Library of Maine" or serve as a hub to provide content to the DPLA. Without such infrastructure, Maine content will be noticeably absent, and the citizenry of Maine underserved.

Third, there is a lack of awareness of "digital government" and "digital education," which will be part of the future for most of the U.S. Digital government refers to the digital infrastructure and services that meet the needs of government and help government meet the needs of citizens for both information and services. This is a major initiative of President Obama (www.whitehouse.gov/sites/default/files/omb/egov/digital-government/digital-government.html), but a similar initiative seems to be lacking in Maine. Digital education is the recognition that digital technology and digital content provide new opportunities to rethink and enhance education at all levels. The Maine Learning Technology Initiative is only a first step in this direction. Serious money and policy development around education for all Maine citizens is needed to make the next generation ready for the world they will inherit. DPLA is certainly positioning itself to make a difference in education across the spectrum from kindergarten through university. Its impact

on Maine could be great if we embrace the need for digital government and digital education in a serious and thoughtful way. ❧

**REFERENCES**

Anthony, Robert S. 2005. "Word for Word, Bit by Bit: No Matter the Promise, Opinions are Split on Google's Plans to Scan Library Books." Black Issues Book Review 7(3): 18.

Bush, Vannevar. 1945. "As We May Think." Atlantic Monthly (July).

Carr, Nicholas. 2012. "The Library of Utopia." Technology Review 115(3): 54.

Darnton, Robert. 2009. "Google and the New Digital Future." New York Review of Books 56(20): 82–84.

Darnton, Robert. 2011. "Google's Loss: The Public's Gain." New York Review of Books 58(7): 10–12.

Singer, Natasha. 2011. "Playing Catch-Up in a Digital Library Race." New York Times (Jan 9): BU3.

Trachtenberg, Jeffrey A. and Jessica E. Vascellaro. 2008, "Google Deal Likely to Give Digital Books Broader Use." Wall Street Journal (October 2).

Wells, Herbert G. 1938. World Brain. Doubleday, Doran, Garden City, NY.

**Clem Guthro** is the director of libraries at Colby College in Waterville, Maine, and the vice chair of the Maine InfoNet Board.