Emerging Topics: Financial/Business Models Workstream Workshop
Paul Courant, with help from Kevin Guthrie, 28 Feb. 2012

We have said all along that the development of business models for the DPLA would require more information about exactly what the DPLA will produce. I propose that at our workstream meeting on March 12-13 we consider a number of possible activities that the DPLA is at least somewhat likely to engage in, and start work on possible business models for a subset of these. The Content & Scope Workstream and the Technical Workstream have both provided starting points for our discussions.

The Content and Scope Workstream, at its meeting in early February, focused largely on metadata, assuming a model in which the DPLA would link to many collections, with implications for both metadata standards and the licensing regimes for metadata. Jeremy York is a member of the Content and Scope workstream, and although he was not present at the February meeting, he knows the issues. Blane Dessy was at the meeting as an observer, and wrote up some notes that he sent to our listserv: https://cyber.law.harvard.edu/lists/arc/dpla-finance/2012-02/msg00003.html.

Additionally, the technical workstream has outlined a vision of how the parts of the DPLA might interact with each other and with a technical platform. According to that vision, the "DPLA platform" must interface with three components: (1) the portal for end-users, (2) a community of software developers, and (3) the external sites and services that may contribute to or use the content in the DPLA. These requirements look to have implications for managing what looks to be a fairly complicated open-source software project, and for providing tech support to content providers who contribute via the platform. That's my very terse expression of implications that derive from http://dp.la/dev/wiki/Overview.

To build on the other workstreams' progress, I attach part of a brief document that I sent to the Secretariat some weeks ago outlining deliverables from our project over the next year. I propose that we flesh out this list, and begin to outline the work that will have to be done in order to make substantive progress.

There seem to be at least three types of activities that will require sustainable financial or business models in order to function well:

1. **Support of central structures (human and physical) required of a DPLA**

   There will surely be some operational successor to the Secretariat, providing governance at some scale – metadata standards, stewardship of collective interests, vetting and auditing of contracts with contributors and vendors, and the like. There will also be (see above) management of an open-source software platform, including support for contributors of content (as distinct from code) who will use that platform. There will be heat, light, rent and payroll. The less the better, no doubt, but the amount won't be zero and it will behoove us to think about how to support this type of activity. There is no "business" here, as such. Rather, there is a set of cost centers that support the whole. The role of our workstream with respect to these is unclear, but we, or someone, must identify the activities and costs. For now, this is a placeholder.

2. **Development of collections themselves**

   An example that I think well worth working through would be government documents. This was the topic of a "Beta Sprint" proposal, discussed at the October Plenary that was developed by the University of Minnesota in collaboration with the rest of the CIC. A rich digital collection of U.S. government documents is a natural for the DPLA. There are elements of such collections in various places that could probably be pulled together, but information on what is where and how to get at it is fairly chaotic. Moreover, a great deal does not exist in electronic form, and the digitizing process (and associated metadata development) will be costly. Thus we have two natural projects in the realm of government documents that will require business models that do not yet exist – drawing from existing digital collections and creating (digitizing) new ones.

3. **Support for the metadata work identified by the Content & Scope committee**

   This one is simpler, but not simple. For example, there is a great deal of public domain material, currently held in the HathiTrust and in other locations, that could be found via the DPLA acting as portal. This will require coordination and standards. It will also require, if it is to work, delineating the value proposition for attaching existing collections (HathiTrust public domain is just an example) to the DPLA.

Looking ahead, there are a few issues that should be in the back (or front) of our minds as we work through topics above. Playing out the example, if the majority of the content is in HathiTrust, what value does DPLA add that HT

is not already providing? What is there that DPLA can/should do that HT cannot do? Really identifying what the value proposition is for someone going to this portal would be good work, work that implicates our workstream and several of the others. One also has to wonder how DPLA will be better or more effective than Google, or probably more accurately, Google Scholar, for at least some purposes. Gaining a better understanding of the impact of Europeana would be very useful here—I am thinking about the demand side, where people are actually using it, not the supply side, where people cheer the inclusion of content in a repository, or point out that they have linking relationships.

The goal of our March meeting will be to augment the work we did last October
(http://dp.la/wiki/Financial/Business_Models_Meeting_Notes:_October _20%2C_2011) to produce a reasonably complete list of types of business that will have to be supported for the DPLA to succeed, and then to pick an exemplary handful of these to work through in more detail. We may also want to consider some business (e.g., articulation with e-reader lending in public libraries) that may or may not be part of a successful DPLA. The presence or absence of this particular example, of course, will be highly consequential. Note that our role as conveners is to bring together other members of the community, as well as ourselves, to do the work. Note also that the Berkman Center will provide some resources to help with research projects.

Workstream deliverables (from memo to Secretariat, Feb. 1, 2012, somewhat amended)

Meaningful deliverables for our workstream are contingent upon getting a list of activities, "products" and services offered by DPLA. We need a reasonable estimate of what we are harvesting from existing sources and what new things we will sponsor (for example, digitizing court records). Assuming that we will not sell content, will we sell customized expertise? Will we be willing to work with other entities that might offer services and help to support DPLA? We need to work closely with Content and Scope and Governance to agree on DPLA activities and guidelines for interaction with relevant commercial and noncommercial markets and then figure out how to make it work. The following examples will require more information from the other workstreams.

- Once we get an estimate of the cost of digitizing [fill in your favorite] at various levels of output, find funding for the endeavor. (Who is charged with obtaining cost estimates?)

- Compile a list of governmental and nongovernmental institutions of interest, including the Library of Congress, National Archives, state library associations, philanthropic entities that work with libraries, and other institutions that hold material that will become part of a national digital public library. A crucial issue for the success of any business models here will be showing how being part of the DPLA is of value to the contributors and their constituencies.

- Compile a list of sources that we can be a portal for – sources that are currently funded and sustainable. We would also need to compile details about their individual business models, how reliable they are, and what backup and recovery mechanisms they have in place. What incentives are there/could there be for continued provision of access via the DPLA as a portal? How many general types (and motivations) exist for plausible contributors? What are they? What's the payoff to contributing via the technical platform of DPLA vs. simply linking to a DPLA portal?

- Explore possible business models for providing or assisting in the provision of electronic access to current literature. This will require close collaboration with others, including publishers and leadership in public libraries, and may or may not prove to be feasible. Any workable plan will surely involve a good deal of commercial activity, and DPLA governance will need to be involved in determining appropriate role(s) (if any) for DPLA.

The preceding leads to what may be the first important deliverable, which is:

- Develop a model for one or two "products" by summer, get the models up and running, and use what we learn to extend the model to other domains.

DPLA Financial/Business Models Workshop Agenda
Shapiro Undergraduate Library, University of Michigan
919 S. University Ave, Ann Arbor, MI

March 12, 2012

7:00 p.m.   **Dinner**
Pacific Rim, 114 W. Liberty, Ann Arbor, MI

March 13, 2011
Turkish-American Friendship Room, 4004 Shapiro Library, University of Michigan

8:30 a.m.   **Registration & Continental Breakfast**

9:30 a.m.   **Welcoming remarks**
Paul Courant
University of Michigan Library

10:00 a.m.   **Update from Content & Scope Meeting, Governance**
(Others?)
Blane Dessy (LOC) and Jeremy York (UM Library, HathiTrust)

11:00 a.m.   **Coffee Break**

11:15 a.m.   **Business models for working with existing collections (e.g. HT) and Institutions**
Tom Sanville (LYRASIS)

12:30 p.m.   **Lunch (Hatcher 818 Admin Conference Room)**

1:30 p.m.   **Potential Models for working with Publishers**
Peter Brantley (Internet Archive)

2:30 p.m.   **Inventory of DPLA "things to be financed"**
Kevin Guthrie (ITHAKA)

3:15 p.m.   **Coffee Break**

3:30 p.m.   **Group Discussion: Identify first 1-2 Projects, a Timeline, Deliverables. Identify key questions and strategies for addressing them.**
Paul Courant and Kevin Guthrie
Large flip charts, note takers, etc…

4:30 p.m.   **Summary and Concluding Remarks**
Paul Courant and Kevin Guthrie
> *Potentially raise the question of who else should be in these conversations moving forward (not to join as conveners necessarily but who could be brought in to offer/present ideas & possible solutions), e.g. people from the business community who have **written** business plans before?*