Marsha Taichman
April 4, 2017

Cornell University Library's Digital Collections Portal:
Ideas about Assessment and Outreach

What is the Digital Collections Portal (DCP) and why was it created?

In November 2015, a team of developers from Cornell University Library IT and Digital Scholarship and Preservation Services (DSPS) launched the Digital Collections Portal, what I will refer to as the DCP throughout this report. The DCP took several years to develop and collections from the library's Digital Collections Registry are being migrated one at a time with the end goal of using the portal as a way to aggregate most of the library's collections.[1] The DCP Team consists of staff from Digital Scholarship & Preservation Services (DSPS), Cornell University Library Information Technology (CUL-IT) and Library Technical Services (LTS). The DCP provides access to digital collections, including those being migrated from DLXS (Digital Library eXtension Services) to Hydra, as well as selected Cornell collections from Shared Shelf Commons. The intention is to incorporate all of the library's collections into the portal so that they may be discovered and cross-searched in a seamless manner. Features include faceted searching and browsing, a IIIF viewer for image zoom, a map interface for discovery of items with geolocation data, and image downloads.

Over the past few years, a team of CUL-IT and DSPS staff have been working to migrate the CUL's digital collections into a new content management and discovery platform. For a number of years, we have relied on a digital collections application called DLXS provided by the University of Michigan Library. Because the platform will be discontinued, Hydra was selected as an open-source repository system for providing more integrated access to our digital collections. After a successful pilot with the Southeast Asia Visions Collection, the remaining DLXS collections will be migrated one by one into the Hydra platform.

The charge of the Digital Collections Portal Development Group was:

- Design and develop a unified interface for searching and accessing digital collections created by the Digital Consulting and Production Services (DCAPS) and other units as appropriate.
- Utilize tools from the Hydra stack in the portal's development that will allow the reuse of the collection data in other applications easier.
- Collaborate with the Discovery and Access (D&A) team to improve discovery of these collections in the catalog and make discovery of the cataloged digital items possible through the portal.
- Participate in the Hydra community and work with other Hydra developers on campus to share and improve technologies, including the Linked Data and Sufia teams.
- Engage a range of stakeholders (e.g., end users, curators, reference staff) to expand understanding of requirements, use patterns, and user support needs for searching and accessing digital collections.
- Participate in different CUL forums to share the group's accomplishments, work plans, and overall goals – seek feedback.

---

[1] There may be collections from Euclid, arXiv and eCommons, as well as others, that will not be migrated, depending on what system they are in.

The near-term development goals include:

- Continue the migration of specified former DLXS collections to the portal.
- Harvest metadata from Shared Shelf to allow access to appropriate Shared Shelf image collections via the portal.
- Harvest metadata from Kaltura to allow access to appropriate Kaltura audio/video collections via the portal.

Quantitative Data on the DCP

I wanted to see what needs and populations that the DCP was fulfilling.

Questions that I had at the outset of the fellowship included:

1. Where are the DCP users from?
2. How long are users staying on the site?
3. Where are users coming to the DCP from directly?
4. Which pages do users visit? Do they consult collections as well as individual images?
5. What are users doing on the site? What do they download?

I discovered how to find this information for various time periods (my methods can be seen in Appendix 3). I thought that it would be useful to set up a protocol for DSPS to measure these stats on an ongoing basis (monthly) after this fellowship is over.

Since the portal was launched in beta, there have been close to 103.2K visits worldwide (103,169 exactly).[2] The average visit lasts 3 minutes and 2 seconds. More than half (53%) of the visits have bounced, meaning that the users are leaving the portal after viewing only one page. For the users that stayed on the site, an average of 6 actions were performed each visit, including page views, downloads, outlinks and internal site searches. There have been 9,747 downloads and 8,230 unique downloads.

Users are coming from a range of referrer websites, both expected and unexpected. One might anticipate referrals to come from site such as Google, Facebook and the Cornell Library website. But *Saveur* magazine used one of PJ Mode's maps in an article found here: http://www.saveur.com/french-food-map, which received numerous hits.

I learned from the developers that this demographic and functional information was not particularly useful to them. What they would use Piwik for is to gather information about specific data if and when they wanted to make changes to the portal. For instance, if they were going to remove a facet on the search screen, it would make sense to research how many people use this feature. For this reason, I spent most of my time during this fellowship focused on gathering qualitative feedback on the DCP.

---

[2] These figures reflect data from August 11, 2015 to February 28, 2017.

Qualitative Feedback on the DCP

I interviewed a series of faculty members from Arts & Sciences (2)[3], and Architecture, Art and Planning (3), as well an MFA student from the department of Art who works with our collections (1). Library staff members (2) were interviewed at length about the portal. Three of the professors had websites built by DSPS to support and promote these collections that they had curated. They were able to compare the display of their own image collections in the DCP, Artstor and Shared Shelf.

The DCP was praised for the following: an aesthetically pleasing interface, the amount of metadata provided for each record, the benefit of different views (list, gallery and map), and the filters and faceting that help organize searches. Several people said that their experience with the portal was favorable to Shared Shelf Commons, which was almost unanimously disliked. A few people who use Shared Shelf for cataloging wondered if the Digital Collections Portal would ever include a cataloguing tool, as they'd had such issues with faceting and making sure that numerous views of the pieces of the same artwork could be separately inventoried, as Shared Shelf is not a relational database.[4]

Some users still miss Luna for the high-resolution images, the image groups that could be built, and the familiarity of its interface.

Suggestions from users include:

The DCP must be easier to find. This will happen naturally when the portal is more integrated with the catalog. In the meantime, finding ways to promote the portal itself, via features on the library homepage, integration into Libguides, bookmarks, a digital sign near the circulation desk at Olin Library, as well as any other libraries that have digital displays are a few ways that might have impact.

The homepage, while visually compelling, is intimidating for some users. Having more information immediately about how to use the site would make the experience less daunting. It was said that **having collection descriptions sooner might encourage use**, meaning before one clicks through to individual records. Perhaps this is incentive for the developers to create splash pages for the collections as they are being input into the portal.

From the homepage, users can go to Browse Collections (https://digital.library.cornell.edu/collections). It would be helpful to have thumbnails of collection images here to give people a visual referent of what they would be consulting. Additionally, having a section with images from the Rare & Manuscript Collection with a separate taxonomy would be useful.

One user said that she wanted **individual images from the portal to open in new tabs** to better organize searches.

---

[3] Numbers in brackets represent the number of people consulted from each population.
[4] For an example of the latter, please search "farnese" in Shared Shelf.

It would be useful not to have to scroll down to see the metadata for each image, which may vary from device to device. This is a complicated proposition, however, as I think one of the strengths of the portal is the size of the images in the record.

Link directly to object record rather than the collections in Artstor or Shared Shelf.

The date of the image should appear in every single record. There are currently records where this is not the case, as seen here: https://digital.library.cornell.edu/catalog/ss:297987.

Image size as part of each record's metadata is crucial. Not only is it important to have the size of the download possible, but the size of the object being photographed. There are photographs from the Icelandic collection that are smaller than the scans being displayed in the portal. One user browsed the Icelandic and Faroese Photographs of Frederick W.W. Howell and observed that the images appeared much larger than they actually are on his computer screen. It seems clear that it would be useful to list the actual dimensional of the image displayed within the metadata, in addition to the file size/resolution of the download.

It may be wise to consider **3 dimensional viewing for objects in the portal**. One user, a recipient of an Arts & Sciences grant who had not spent much time with the DCP, claimed that while she appreciated the look of the DCP and found it easy to use, her collections were not available with the same functionality there. In this instance, in addition to the 2 dimensional objects imaged, there are also 3 dimensional views that can only be seen and experienced via the collection website. How, if at all, will these features be incorporated into the DCP?

There are certain value added features that the portal offers that cannot be found with Artstor or Shared Shelf Commons. One that was mentioned was the mapping function, which is only available for some collections. A user pointed out that in certain collections where the mapping would be extremely useful it has not been set up. **More mapping for individual collections would be extremely beneficial, especially when geographic location is integral to the discrete images.** An example of this is with photographs from the Historic Glacial Images of Alaska and Greenland collection. If this mapping is too labor-intensive for the developers, the user suggested that the work could perhaps be crowdsourced.

Let's use the same collection of images from Alaska and Greenland as an example of an individual collection. When you click on Browse Collection from this site, the redirect goes to Shared Shelf Commons. A user pointed out that it would be helpful to have both options available to users from the website and from the collection-level record in the catalog. Perhaps instead of linking to the website, links to Shared Shelf Commons and the portal could be included. The user commented that from the library catalog, bibliographic entries can be found that link to many electronic and print journals (such as *The New York Times*) rather than providing one link. It might be good for users to have the option to access the collection in three different ways: through the collection's website, if one exists, Artstor and Shared Shelf Commons.

Collections that might be prioritized for mapping are the Historic Glacial Images of Alaska and Greenland collection, Beyond the Taj: Architectural Traditions and Landscape Experience in South Asia, the Icelandic and Faroese Photographs of Frederick W.W. Howell, the John Reps Urban Explorer collection, and the Billie Jean Isbell Andean collection. These are not necessarily the most consulted collections, but geographic location is crucial to the discrete images.

User testing for entire site would be helpful. The developers intend to do some usability testing after the next phase of integration with the catalog, which is a great initiative. Before this happens, evaluating the UX might be worthwhile. The user who mentioned this had a specific example of a perceived problem in mind. On the item level, in the About the Collection box, it might be good to retitle the link View collection in Artstor/Shared Shelf Commons as View collection in alternate interface (Artstor/Shared Shelf Commons).

Promotion of Digital Collections

One of the goals of the DCP that is not often discussed is its potential as a branding tool for the library.

At CUL, we excel at creating digital collections and refine our workflows on an ongoing basis to increase efficacy. We work hard to build collections, but once they have been created, the promotion aspect is often overlooked, as we move to our subsequent projects. It occurred to me that we would benefit from a kind of protocol for post-production activities that we could use to guide collection promotion. This is not to say that each collection would go through the exact same process, but creating a checklist would be something that faculty members and curators could consult when launching a new collection.

There are many things to consider before building a digital collection, as well as after it is created. Perhaps the most important questions to keep in mind are: Who are your users? How can you create digital objects, including their accompanying metadata that support them? Designing to support users and making users aware that the collections are available is good practice.

With the help of DCAPS and Assessment and Communication, I came up with a checklist for digital collection promotion (Appendix 4). It was my hope that the most recent Architecture, Art and Planning (AAP) grant project related to Sri Lankan vernacular architecture would be completed before the end of this fellowship so that I could pilot the checklist with a collection that I had worked on, but the timing did not work out, as the project is still in process.

Case Study on Collection Promotion: PJ Mode's *Persuasive Cartography*

The most consulted collection by far during 2016 fall semester was PJ Mode's *Persuasive Cartography*. With this in mind, I reached out to Mode to try to understand why this might be the case.

We all know that collection promotion is important, Mode explained. As popular as the website is, he wishes that it was more widely known, but everyone is competing for clicks. Mode collected the maps that were digitized for this site for many years, and kept up a contact list of collectors, dealers and people who took interest in his collecting. When the website was launched, he wrote an email to all of his contacts (comprised of about 60 addresses) and wrote personal messages to others explaining the project. This meant that from the outset, many people were aware of the site, even if they weren't following it closely.

Mode has received much promotion for his collection via blogs. He told me, "In every area of human endeavor, there are people who blog." Accordingly, people who blog are always looking for new

content. With this in mind, he tried to figure out the bloggers who are interested in maps. Once he identified these individuals, he composed an introductory email and tell them about his website, including links in his message so that they would not need to look the site up. Virtually every time Mode sent an individual email, it resulted in a blog entry related to his site. The website *Atlas Obscura* published 7 stories on his maps, and National Geographic online had an article about how maps can be used as data, found here: http://news.nationalgeographic.com/2015/10/151022-data-points-how-make-maps-influence-people/. In the latter instance, the author, Geoff McGhee, found Mode's maps by reading about them on another site (http://bigthink.com/strange-maps/the-map-as-persuader).

There are some blogs where Mode would like his work to be included, such as *Musings on Maps*, and he intends to approach writers when the next installment of the website is available and working smoothly (500 additional maps are being scanned and cataloged). He would like to make headway into the popular press with his next installment of maps on the Persuasive Cartographies website and have articles in *The Wall Street Journal* and *The New York Times*. Since the site launched in 2015, Mode has given 4 map-related talks in Princeton, New York City, Washington and Denver, and he uses these opportunities to discuss the website. He also looks for ways to have his site discussed in scholarly journals, such as *The Portolan* and *The Journal of the Washington Map Society*.

In our discussions, Mode astutely noted that much of the success of promoting his maps has to do with the communities of interest, the involvement of the curator and the inherent interest of the collection, and that there is no way around the fact that some things have more visual appeal than others. He suggest that the library finds a champion for each collection and to have the champion figure out who would be interested in the content globally. DSPS is starting to do this by assigning stewards within the library for each collection as it is being created. That said, we do not always have deep expertise in the collection areas that we are supporting. Perhaps it should be built into the library staff member's role to not only research the content of the site, but also the people who would benefit from using it. It would also be advantageous to get library content into DPLA so that more people would happen upon it.

Talking to Mode about Persuasive Cartography was inspiring, but it raised many issues for me. The project is very much Mode's own, and he has invested a great deal of time and money into making the website a resource that is used. Often, the collections that we build in DSPS are grant projects, and are very closely related to faculty research and teaching. It seems, though, that we don't do much to hold these creators and curators accountable for collection promotion, though they are the ones who are shepherding these digital images into being. As part of the application process for Arts & Sciences and Architecture, Art and Planning grants, it would be useful to have applicants commit to collection promotion. If they could explicitly state where and how this potential collection could be used, the library could assist with making those connections. We could hire research assistants in different subject areas, as we have with students who do metadata entry for collections, and have these assistants find out which communities and professional organizations would benefit from these new collections and what kinds of outreach (such as emails, blog posts or articles) would be more productive in attracting audiences.

For what is the purpose of building a visual collection if it will not be used?

Recommendations Related to the DCP and Collection Promotion:

Here are some recommendations in addition to the specific suggestions discussed in the qualitative section of this report.

- **Forming a DCP Advisory Group[5]** will allow the developers to have ongoing feedback from people in public services about any issues that arise related to the portal. This group also has the potential to be a sounding board for any changes that DSPS hopes to make to the portal in the future.
- **Integration of images within the catalog is imperative.** This is something that Discovery & Access is working on, but I cannot overemphasize the importance of this element of the catalog. I would argue that the current catalog is incomplete as it stands with images being so marginal in a range of other materials that need to be discoverable.
- **Make the DCP more readily discoverable**. Increase its visibility in the library catalog, and figure out how we can best promote it as an institutional repository.
- One faculty member suggested that **DSPS staff meets with all new faculty hires** to discuss their research and how the Images for Teaching and grants initiatives might benefit them. Getting people interested in our services as soon as possible will be advantageous for us as well. The logistics of this seems complicated, and there would need to be coordination with individual departments, but it would be a worthwhile endeavor.
- **UX testing would benefit the DCP**. I discussed this with the developers, who believe that the right time for usability testing is when the next layer of integration happens within the catalog. My thought is that we should test early and often, but I appreciate it could be hard to keep up with the changes that might need to be made.

---

[5] Such a group was successfully created by Jenn Colt and Melissa Wallace.

Appendices

Appendix 1: Examples of Similar Collections Online Identified by the Developers

Columbia Digital Collections: http://library.columbia.edu/find/digital-collections.html

Digital Commonwealth: https://www.digitalcommonwealth.org/

Duke Digital Collections: http://library.duke.edu/digitalcollections/

Library of Congress Digital Collections: http://www.loc.gov/library/libarch-digital.html

New York Public Library Digital Collections: http://digitalcollections.nypl.org/

North Carolina State University Libraries' Rare and Unique Digital Collections:
http://d.lib.ncsu.edu/collections/

Northwestern Images and Audio-Visual in Blacklight:
https://images.northwestern.edu/
https://media.northwestern.edu/

Oregon Digital Collections: http://oregondigital.org/catalog/

Stanford Library Digital Collections: http://collections.stanford.edu/

University of British Columbia Open Collections:
https://open.library.ubc.ca/

University of Nevada, Las Vegas Digital Collections: http://digital.library.unlv.edu

University of Wisconsin-Madison Libraries Digital Collections:
http://uwdc.library.wisc.edu/collections

Virginia Commonwealth University Digital Collections: http://dig.library.vcu.edu/

Yale Digital Collections: http://web.library.yale.edu/digital-collections

Appendix 2: Questions Included in Qualitative Discussion of the DCP

General questions that I wanted to address:

- What are patrons using the site for?
- How was their experience of the portal?
- Did they find what they needed?

Specific questions for all users interviewed:

- Where do you go to look for images?
- Do you use the portal as a way to find images? Why or why not?
- What are you using the site for?
- How was the experience?
- What is your impression of the overall look and feel of the site?
- Is there enough information accompanying the images?
- Did you find what you needed?
- Is the portal and the collections within it supporting your teaching and/or research?

Appendix 3: Methods of Gathering Quantitative Data in Piwik

1. Where are the DCP users from?

   Go to Visitors > Visitor Map (Continent, Country, Region, City, Browser language, Providers).

2. How long are users staying on the site?

   Go to Visitors > Engagement.

3. Where are users coming to the DCP from directly?

   Actions > Entry pages.

4. What pages do they visit? Collections and individual images?

   The top three most consulted collections in the DCP between September 1 and December 2, 2016 were Persuasive Maps: PJ Mode (1,674 views), the NYS Aerial Photographs (810 views) and the Rare & Manuscript Collection's Hip Hop Party and Event Flyers (242 views). See Appendix 4 for the breakdown of additional collection views.

   In order to find this information, I went to the Piwik dashboard and selected the facet Visitors. I then went through every collection that had been input into the DCP up until that point and created a segment (a word or phrase that could be found in every URL of the collection being searched) for each collection.*

   For Persuasive Maps: PJ Mode, the settings set across the Piwik header were:

   Website: PORTAL_JS
   Date Range: September 1 - December 2, 2016
   Visits: [search with the segment "maps"]

   Creating a segment: Add new segment > Name [maps] > Add the condition that the Page URL must contain maps.

   Perform the search.

   * Many thanks to Adam Chandler for helping me to establish a search strategy to find this information.

5. What are they doing on the site? What do they download?

   Actions > Downloads. There are many places to get information on what users are doing on the site.

Appendix 4: Most Consulted Collections in the DCP between September 1, 2016 and December 2, 2016

Alfred Montalvo Bolivian Digital Pamphlets Collection: 174
Alison Mason Kingsbury: Life and Art: 27
Andrew Dickson White Architectural Photographs Collection: 197
Beyond the Taj: Architectural Traditions and Landscape Experience in South Asia: 76
Billie Jean Isbell Andean Collection: 15
Campus Artifacts, Art & Memorabilia: 86
Cornell Coins Collection: 45
Cornell Gem Impressions Collection: 16
Cornell Squeeze Collection: 15
Hip Hop Party and Event Flyers: 242
Historic Glacial Images of Alaska and Greenland: 39
Huntington Free Library Native American Collection: 150
Icelandic and Faroese Photographs of Frederick W.W. Howell: 22
John Clair Miller: 21
John Reps Collection - Bastides: 54*
John Reps Collection - Slides: 37*
Mysteries at Eleusis: Images of Inscriptions: 13
New York State Aerial Photographs: 810
Persuasive Maps: PJ Mode Collection: 1,674
Punk Flyers: 62
Ragamala Paintings: 119

* I ran the search for "reps" and came back with 86 results when I expected to get 91 (54+37).
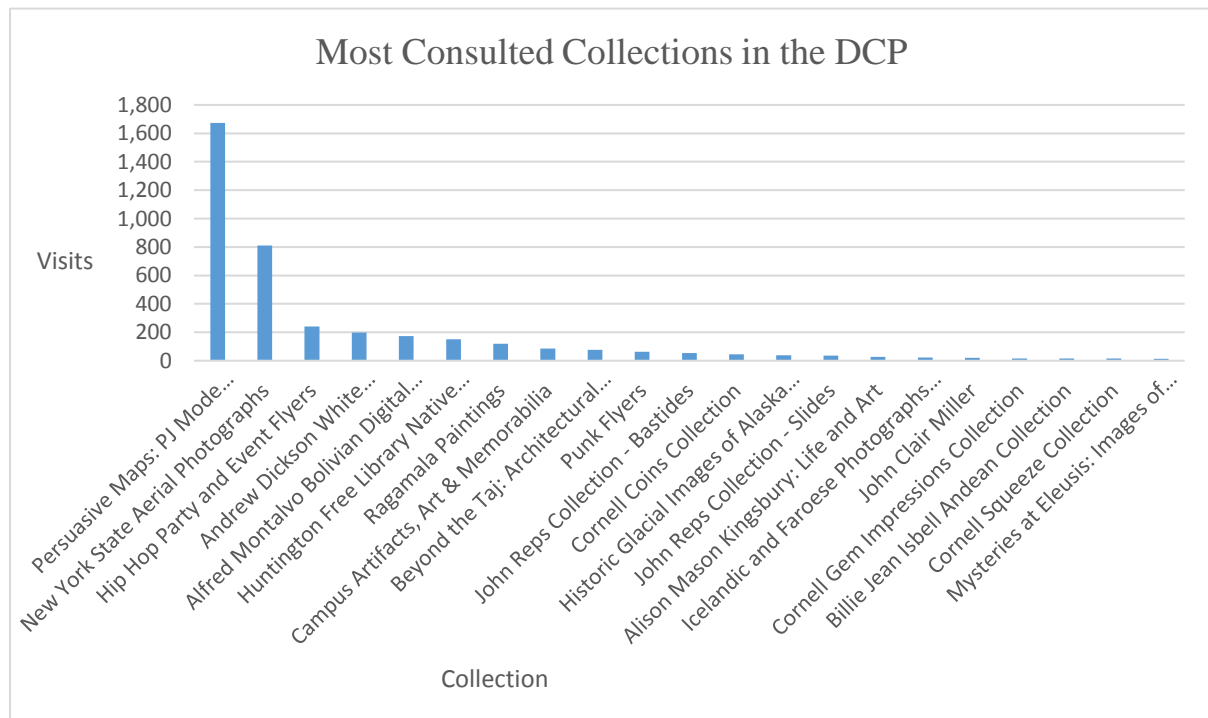
Appendix 5: Checklist for Digital Collection Promotion

Consult with Assessment and Communication about a month prior to release of collection. They will help you through the following steps:

1. Identify audiences likely to be interested in the collection
2. Craft effective message to highlight benefits of collection
3. Identify channels to reach relevant audiences
4. Consider ownership of collection and collaborate with owning college or department as appropriate
5. Consult on copyright issues as needed

Generally the most useful communication channels are:

Appropriate **print and online news sources** can be reached via Assessment and Communication, such as:

- Cornell Chronicle
- Cornell Daily Sun
- Ezra Magazine
- Discipline-specific publications
- Special interest publications

What **listservs** should receive messages about the collection? Reuse core messages created for new sources.

- All Cornell collections should be publicized to CU-LIB as they are created
- Announce to library liaison listserv
- Enlist help from faculty members responsible for the collections and leverage their professional networks
- Research discipline-specific listservs

Make sure that the project goes on the **Cornell Digital Collections page**: https://digital.library.cornell.edu/.

Consider if **promotional material** (bookmarks, flyers, rack cards, etc.) is needed and distribute in relevant departments.

Social media – in addition to unit social media accounts, CUL social media accounts can push messaging out to larger audiences (again, consult Assessment and Communication).

Include links to collections in **Wikipedia** under relevant subject headings.

Blog post (a day in the life of a collection, a collection description, etc.).