

Title: Building the Duchamp Research Portal

Format: Short paper

Author: Margaret Huang

Affiliation: Philadelphia Museum of Art

Contact information:

margaret.huang@philamuseum.org

+1 215-684-7641

Philadelphia Museum of Art

PO Box 7646, Philadelphia, PA 19101-7646

[Slides here](#)

[Slide 1 - Introduction]

Hello everyone, thank you for viewing this presentation. My name is Margaret, Marge, Huang and I am the Martha Hamilton Morris Archivist at the Philadelphia Museum of Art. I was also the project manager for the Duchamp Research Portal.

[Slide 2 - Project Overview]

The Duchamp Research Portal (DRP) is an online platform that unifies the digitized archival holdings at the Philadelphia Museum of Art, the Centre Pompidou, and Association Marcel Duchamp to make a significant portion of Duchamp's primary source materials accessible and discoverable through a single search interface built on Linked Open Data and IIIF.

Marcel Duchamp was one of the most influential artists of the modern era. He changed the way we think about art. Duchamp's work—which includes paintings, sculptures, prints, and readymades—represents a legacy that informs our understanding of the development of modern and contemporary art.

Making up for the geographic distances between the participating institutions, this online resource digitally aggregates Duchamp-related archival materials along with his artworks and renders them accessible, discoverable, and usable for the Duchamp research community, scholars, and the broader public in a way that has never been possible before. The materials include correspondence, drawings, technical plans, photographs, exhibition ephemera, newspaper clippings, scholarly articles, and more.

[Slide 3 - Project Overview + Specs]

Here is an overview of our project. It represents not only an aggregation of primary source Duchamp materials but also an aggregation of people, technologies, and workflows over the course of many years. I want to recognize that these types of projects don't just happen but require concerted and focused efforts, close cooperation and coordination, and resources both in terms of funding and human energy.

[Slide 4 - Project Timeline: 2014-2016]

The DRP began with a planning phase, funded by a National Endowment for the Humanities Collections and Reference Resources planning grant, from 2014 to 2016.

The main objective during this time was to investigate how the DRP could be implemented. Activities of the planning phase included: establishing an Advisory Board, conducting a user survey, performing collections surveys to document processing, digitization, cataloging, and copyright issues, and producing a summative white paper and proof of concept model. We also established the PMA, Centre Pompidou, and Association Marcel Duchamp as our founding partners.

[Slide 5 - Project Timeline: 2017]

In 2017, we received an implementation grant from the NEH.

In our first year of implementation, we focused on: selecting our development firm (Design for Context) and our digitization vendor (Backstage Library Works). We also started analyzing sample data sets to work toward establishing a common data model.

We used this spreadsheet as an index for our artwork and archives data. The partners reviewed this together to see what types of metadata we were each creating, how we used specific fields, and where our gaps and overlaps were.

[Slide 6 - Project Timeline: 2018]

Significant strides were made in 2018. We used the data analysis work to establish target data models for both art and archives, which was crosswalked conceptually between DACS and CIDOC-CRM, set technical specifications for the DRP, began item level metadata creation, finished digitization at the PMA, and had a beta site with 27,000 images and their associated metadata by the end of the year to share with our partners.

[Slide 7 - Project Timeline: 2019]

2019 was dedicated to finishing metadata creation, gathering feedback and refining the beta site, which included conducting an accessibility audit and adding bilingual capabilities to the site. Creating metadata for the DRP is where I really got to know the materials first hand since I laid eyes on every single piece of paper and photograph that the PMA contributed to the DRP. Here are just a few of my favorite images of Duchamp throughout his lifetime.

[Slide 8 - Metadata Workflow]

I'm going to deviate from the timeline for a bit to focus specifically on our metadata workflow. To make the images in the DRP searchable and usable, a lot of descriptive metadata was created. We re-used existing metadata from our finding aids and enhanced them with item level description. Most of the descriptive work was done in Excel because features such as data filtering and formulas or even the simple ability to copy and paste helped a lot in building

efficiencies. Here is an example of the column headings in our spreadsheet. We then imported the spreadsheets into our collections management system, ArchivesSpace, using a plugin developed by Bobbi Fox who is now at Tufts University. We then export it as XML and commit the file to our github repository. The data then goes through a linked data transformation. The linked data is loaded into the triple store, which we then query using SPARQL to produce JSON-LD documents. Finally, we take the JSON-LD documents and produce simplified versions that the page builder uses to produce HTML documents for the site, and by the search index for the browse and search pages.

[Slide 9 - Project Timeline: 2020]

2020 was dedicated to ingesting the remainder of the Partners' images and metadata, search testing, and rights negotiations. As of this writing in October 2020, the site is ready for launch but due to changes in staff and complications arising from COVID-19, there have been delays in finalizing copyright and intellectual property policies and statements. However, we hope to launch by the end of the year.

[Slide 10 - Project Documentation: drp.museumbrowse.org]

In the meantime, you can find a technologies overview, data models, and other documentation available at drp.museumbrowse.org.

[Slide 11 - Preview: duchamparchives.org]

Now that you know a little bit more about the steps that got us here, I'm going to walk through a preview of the DRP.

Here on the home page, you can toggle between English and French versions of the site. We also have randomized images to entice users to explore the materials in the Portal.

The part of the site under "Museum Collections" presents Duchamp's artwork, which you can browse and filter by type, date, artist, and institution.

The other part of the site, under "Documents", presents the Duchamp related archival materials. Here you can also browse and filter by format, collection, date, people, and institution. You can even combine filters so for example, you can see all correspondence from the 1950s.

If you click on an archival object, you can deep zoom on the image and get more descriptive information which is also available as JSON-LD. You can also explore its relationship with other objects that might be related by something like creation date or format.

You can also navigate the archival material through the finding aid structure. Here you can read narrative parts of the finding aid such as biographical and historical information. You can also drill into the hierarchy. So you can click on a folder title and see all the materials in that folder.

Back to the archival object page, you can also see how you can use the breadcrumbs of the folder, series, and sub-series to see the materials at any point in the hierarchy of the arrangement.

You can also search across both the museum and archival collections and jump between the two using the relationships tab. So here I've searched for Duchamp's artwork *Etant Donnes*. This is an archival condition photo related to the artwork. In the relationship tab, this photo has the *Etant Donnes* object record related to it. And then if you look at the relationships on the object record for *Etant Donnes*, you'll see all the archival documents the artwork is depicted or referenced in.

Duchamp's artwork is the anchor of this site but we hope that by building these relationships and presenting it alongside related archival documents, users can reveal and stumble upon new stories and insights into the life of Marcel Duchamp and his circle.

[Slide 12 - Conclusions]

Lastly, here are some final conclusions I want to share about this project.

Linked data projects might seem intimidating but they are achievable!

There are many short, non-technical videos on YouTube that were especially helpful for getting the concept of linked data to click for me. Linked data work can be iterative and a great first step is to add URIs where possible in your collection management systems or adding institutional IDs to Wikidata.

Work towards consistency, not perfection.

Perfect solutions are hard to attain but working towards consistency allows for repeatable and automated actions to fix or improve things.

Set up a team based on "A Role-Based Model for Successful Collaboration in Digital Art History"

The team model proposed by this article suggests that successful projects should include the following roles: humanist, technologist, data steward, catalyst. We followed this format and it worked well for us.

International copyright is complicated.

Allow ample time, staffing and budget for rights work when collaborating across different legal jurisdictions. We brought in our museum counsel and rights manager to help us.

[Slide 13 - Thank You]

Thank you for your interest and please do not hesitate to reach out to me with any questions at margaret.huang@philamuseum.org.