

Digital Collections in Primo User Testing Report

Orbis Cascade Alliance

Unique & Local Content Team

Digital Collections in Primo Project Group

2018 May 29

Project Group Members

Ann Lally, University of Washington, Chair

Sara Amato, Willamette University

Philip Vue, University of Portland

Dan Oswald, Lewis & Clark College

Jodi Allison-Bunnell, Alliance (ex officio)

Executive Summary

Following up on the work done by the [Digital Collections Working Group](#) in 2016-2017, which evaluated how users responded to digital objects in comparator library catalogs, the Digital Collections in Primo Project Group (DCP PG) collected data from five user tests designed to test user responses to digital objects in our own Primo environment. These tests were performed in the Primo Premium Sandbox and in card-sort activities on the [OptimalSort](#) platform. Five Alliance libraries participated in the testing, for a total of 27 test subjects, most of whom were undergraduates. Although there were some limitations to getting usable data from a test outside of a live Primo instance, the results still provide enough information to move this project along.

Overall, the tests showed:

- Users did not react negatively to seeing digital objects in their search results.
- Online/physical availability was usually a secondary or final consideration, i.e. searches rarely began by limiting to online items only.
- Subjects preferred nested facet terms to cut down on long lists in facets. We were able to find agreement on a set of high-level terms to condense many form and genre terms for use in nested facets.

Results of this testing will inform the completion of the Shared Objectives for Digital Collections in Primo ([already presented and discussed in draft form last February](#)), which will in turn inform whether to move this functionality to production in FY19.

Background and Context

[Full Report of User Testing from Digital Collections Working Group](#), 2017 Spring (predecessor)

The DCP PG's user testing follows up on the [Digital Collections Working Group](#), which tested end-user experiences with digital resources in other Primo library catalogs with digital objects included. This group has been the first to test end-user encounters with metadata from the [Alliance Harvester](#) in our own Primo environment.

As of this writing, the harvester has about 40,000 digital objects in it for Primo, and there is a pipe into the Alliance's Premium Sandbox for testing. The [Shared and Unique Metadata in Discovery](#) ad hoc group worked through a list of potential issues and solutions as their first pilot project. The Normalization Rules Standing Group has continued work on resolving [issues with search, display, and function in Primo](#)

Test Focus Areas

The following test focus areas were recommended by the predecessor Digital Collections Working Group, as detailed in the [Digital Collections in Primo Project Group's charge](#):

1. Organization/Hierarchy

Recommend Top Search Box

Determine if digital objects go into a special search scope, along with other materials, or both.

Do we want digital objects to behave the same way as commodity content, or differently because they are special?

Recommend Facets

Determine how to effectively categorize and label digital objects from DAMs or IRs in a search interface that includes all manner of library materials.

For example, a box of print images available in an archives (metadata only available online) vs. a digitized image available online - both could be in a resource type facet of images, but they are distinct in their modes of access.

Should digital objects go into a category under a 'type/genre' facet or be in a facet by themselves?

What is the relationship between facets for 'Resource Type', 'Availability' and 'Physical Location'?

2. Implementation/Usability/Findability

Feasibility

Determine if results from above categorization exercises can be implemented in Primo, and if they provide a satisfactory user experience.

Using results from the above, explore implementation in the premium sandbox, noting if system and/or data limitations exist in implementation of desired features. (Note also the [Trello board that the Normalization Rules SG and the ULC Program Manager work on digital collections issues](#)). How do end users (Research Faculty, Teaching Faculty / Graduate Students, Undergraduate Students, and Digital Humanists) interact with these types of materials?

Guidance provided by the predecessor group includes:

- Ensure path to the digital object is clear.
- We have determined that our end users are not particularly interested in which institution holds digital objects. However, we know that Institution identity is important to the institutions. We will need to determine what is acceptable to the members.
- Each individual institution should have control over whether digital items are included in search results and the display of the digital objects in the local-only environment

Testing methodology

Design of the user tests

There are many different potential consequences of introducing a whole new subset of resources into Primo. How will users react when they encounter a slew of new resource types in their results? Indeed, much of our early discussion was devoted to teasing out what was most significant, and what was even testable.

Large Card Sort: The vocabulary and organization of facets were essential to test, since items in the harvester would be piping in with varied and often user-unfriendly metadata terms. Additionally, they could potentially clutter the lists in the facets by sheer number. Based on the recommendation of the [Digital Collections Working Group](#) in 2016-2017, we chose a card-sort to determine user preferences for labels and categorizing groups of genre terms. We consulted with Holli Kubly from the [Discovery and User Experience Team](#), who had experience with card-sorts, who agreed that this would be a good choice for our use case. We used her recommended product, OptimalSort, that has a very user-friendly interface and an exhaustive set of reports. We conducted a short pilot-test of the OptimalSort software with some of the terms, and adjusted our card-sort based on those results.

The card sort activity was designed with 3 parts.

1. Pre-test Questionnaire: Aimed to gather information about participant's expectation when searching for digital objects in the library catalog (Primo)
2. Card Sort Activity: Participants are given the opportunity to categorize and group digital objects terms as they see fit, and
3. Post-test Questionnaire: Aimed to gather information about participant's search behavior

Originally, we had identified 112 terms for the card sort activity. These terms were based on the Type metadata schemes identified in the [Alliance Dublin Core Best Practices](#), genre terms assigned to digital objects in the harvester, and included terms from Mountain West Digital Library commonly used subset of the entire AAT genre terms.

These were a lot of terms to potentially sort through, so we did some informal pilot-testing to see how test subjects might react to those large numbers. It also gave us a chance to see the OptimalSort software in action so that we could provide good instructions, and some results to see how the reporting would look.

We created 66 cards for the live test after thoroughly analyzing pilot-testing results. The 66 cards in the live test contained terms that were unique, singular by nature, and broader terms representing more specific terms from the 112 terms identified. For example, we consolidated specific terms like featured films, motion pictures, television programs, and video recordings to a broader term Moving Images. All the DCMI type terms required by the [Alliance's Dublin Core Best Practices](#) were originally included as cards. However, DCMI type terms *Collection*, *Event*, *Service*, and *Physical Objects* were excluded from the live test as pilot-test participants found these terms confusing to sort. Each card also contained term definition from AAT to help clarify any confusion during the sorting activity.

Sandbox Tests: We used the [Alliance's Primo Premium Sandbox](#) as the closest thing we had to a live test. During and after the [Digital Collections in Primo Forum on February 8, 2018](#), the primary concern that we heard from Alliance partners was how users would react to digital object records appearing in their search results. We could test this in the Premium Sandbox. We wanted to record what users did when looking for commodity items, like physical books or media items, and also what they did when looking for digital items. We designed a test where the user was tasked to find a handful of materials (some physical, some online-only) for an imaginary class project.

Small Card Sort: Lastly, we found that we had one important question from our objectives that was not quite being touched on yet. We wanted to know whether users thought online items belonged in their own separate 'bucket' (not included in results by default), or if all items belonged in the same 'bucket' (included in results by default). So we designed a final short card

sort designed to see if users group items by type (i.e. CD grouped with mp3), or by online/physical availability (CD not grouped with mp3).

The tests were designed to be able to be as comprehensive as possible in a small time frame so that they could administered by one tester in a single session. They were also designed to be able to be administered to all four of the user personae that the Alliance considers the primary end users: Undergraduate, Faculty/Research, Faculty/Teaching, Digital Humanist.

Recruitment

Because of the approach of end of the school year, [recruitment occurred in a compressed time frame](#). The ULC Program Manager finalized [the testing packet](#) and sent out a request for potential testers to sign up in a google form on April 20, 2018, and a follow-up was sent on April 30, 2018. Additionally, the group's chair and the Program Manager briefly spoke about the project on a Discovery & User Experience open call on April 17th to encourage participation. Nine institutions signed up to conduct tests.

Testing

All Alliance members were provided [the testing packet](#) via the Discovery, ULC, and Systems community lists and encouraged to conduct at least one test. Testing occurred between April 23 and May 4, at 5 of the 9 institutions that signed up. Twenty-four of the 27 subjects tested were undergraduates, so we did not meet our goal of getting all the user personae tested. Additionally, only 16 of those subjects completed both of the card sorting activities.

Results / Analysis

Primo Sandbox Tests

Results from this part of the test can be found here: [Testing in Primo results](#)

First, **users were positive (or indifferent) to seeing digital objects in their search results (i.e. it was no paradigm shift for them to see these things in there)**. Indeed, it was a much bigger obstacle to *not* see results than to see too many, as long as users felt like they were on the right track and had a path forward via the filters.

However, in designing this test, it became clear that **there were some major limitations to trying to get usable data from a test of this sort outside of a live Primo instance**.

Frustration occurred with UI issues, such as unclear paths to access the objects, or that the sandbox was set up with the more useful filters like Resource Type buried 'below-the-fold'.

There was also frustration when there weren't very many results for broad searches, especially

when limiting to digital resources. It was also very difficult to simulate realistic searches. The searches we designed were 'discovery' searches rather than 'known-item' searches.

In the section where we asked them specifically to find digital types, only 11% of users used advanced searching options from the start, and 59% of subjects never noticed the option to search just "digital collections." These numbers could lend some weight to the fact that the majority of our test takers didn't make use of most of the faceting or sorting tools available in the sandbox because they didn't notice them.

There were many comments about not knowing if things were available online or not, which was going to be an issue for the majority who were entirely dependent on the "Resource-Type" facet, as there was inconsistent use of digital indicators for some labels, like "e-video and e-audio", but not for others. Fortunately, this was the focus area for the next section of our user-test.

Large Card Sort

Results from this part of the test can be found here: [Raw Optimal Sort Results](#).

Here are detailed explanations of the tabs of the worksheet:

Tab 1. **DCTypes/AAT terms - Original list.** Contains the original list of Type metadata identified based on Alliance Dublin Core Best. List of terms include the required DCMI Type vocabulary and preferred Mountain West Digital Library commonly used subset of the entire AAT genre term.

Tab 2. **DCTypes /AAT terms - Grouped/Consolidated.** This tab contains the color coded grouping of terms used for the live card sort activity. Terms with similar meaning are grouped by color with the bolded term being the preferred term used in the live card sort test.

Tab 3. **DC Types.** Contains all DCMI Type vocabulary. These are the the required types identified in the Alliance Dublin Core Best Practices guideline.

Tab 4. **Currently in Optimal Sort.** Contains the terms used to create the cards in the live card sort activity. Terms are separated by columns:

- a. In Optima Sort - AAT terms used in live card sort test
- b. DC Types - DCMI type vocabulary used in live card sort test
- c. Currently in Primo - Type terms from harvester currently in Primo Premium Sandbox.
- d. In Primo, no in our test - Terms from harvester currently in Primo Premium Sandbox not used in live card sort activity
- e. DC Type not in our Test - DCMI Type vocabulary not used in live card sort activity

Tab 5. **OptimalSort Dendrogram Results.** Dendrogram results generated by OptimalSort illustrate data clusters and assumptions about terms and relationship. For example, the terms postcard and correspondence are grouped by users as correspondence, and is a Text type digital object.

Tab 6. **Suggestions based on OptimalSort Results.** Our analysis of card sort raw data and suggestions for facets and nested facets based on the card sort results. See [Diagram of Suggested Facets](#) document for facet visualization.

Tab 7. **Pre/Post Test questions.** Pre and post test raw data with diagrams and suggestions on user search behaviors and digital object organization in Primo.

As one might expect, users tended to group like items into broader categories. Terms used for these broader categories were similar enough to suggest potentially useful facet names for these groupings. For example, users grouped terms like *conceptual drawings*, *drawing (visual works)*, *schematic drawings*, *sketchbooks*, and *working drawings* into a common subtype (nested facet) category labeled *Drawings*. Another example is the grouping of the terms *cartographic materials*, *maps*, and *plat maps* under a subtype (nested facet) *Maps* category.

These commonalities led to six high-level facets (*Images*, *Maps*, *Text*, *Video*, *Audio*, *Software*), as well as twelve nested facets: *Drawings*, *Image*, *Maps*, *Articles*, *Books*, *Business Records*, *Correspondence*, *Legal/Government Documents*, *Manuscripts*, *Newspapers*, *Text*, and *Theses*.

These commonalities suggest having the following high-level facets in Resource Type, with more specific terms nested underneath:

- Audio
- Images
- Maps
- Software
- Text
- Video

A detailed mapping of the more specific terms can be found in the [Diagrams of Recommended Facets](#), which is a digest of the card sort's results. Note that the terms are a mix of DC types, ATT types and other terminology currently found on digital objects in the Primo sandbox.

These commonalities should prove useful during the normalization process for general facet terms for the many granular and sometimes eccentric terms that live in the harvester. We have a start on consistency in the digital object metadata on which we can build to provide a coherent search experience.

Pre/Post Questions

Results from this part of the test can be found here: [Raw results of the pre/post questions](#)

These questions asked users to imagine how they would feel in six scenarios in which they would be confronted with digital objects. We wanted to ask users about their preferences directly, and compare with the test results.

Users had contradictory answers that did not match the behaviors witnessed in the tests. While an interesting study in perception versus reality (or clarity of the questions asked), we did not find that the results here were conclusive enough to make assertions.

Small Card Sort

Results from this part of the test can be found here (starting in row 96): [Raw Optimal Sort Results](#).

The dendrogram (a tree diagram showing taxonomic relationships) that OptimalSort provides for this second card-sort test showed a **high level of agreement from users for grouping items by resource type rather than physical/online availability**. Although the test results were pretty small, this result was strongly reinforced by users specifically noting in their post-test comments that they felt that this was the question they were being asked via this card sort activity. In response, they explicitly said that they would prefer to keep like-types together (e.g. all video together, all text together), regardless of physical or online availability: They didn't care if it was physical or digital, as long as they could get the content they wanted. **They wanted to see all their options, for as long as possible; availability was usually a secondary or final consideration**. Although more testing in a live environment with more than just undergraduates may prove that this sentiment is not as clear as it might seem, it was still probably the most surprising overall trend in our testing, and definitely something to note.

Other observations

Examples of nested facet terms for Resource Type came up in our exploration of other library catalogs, such as [Brigham Young University](#), and multiple users independently made remarks during testing that suggested a similar idea to ease the visual clutter of long vertical lists. These results could be used to determine nested facet terms as well. See also: [Diagrams of Recommended Facets](#).

Appendixes

[Testing packet](#)

[Testing in Primo results](#) (raw, reviewed to remove any identifying information)

[Diagrams of Suggested Facets](#)

[Raw Optimal Sort Results](#)

[Optimal Sort Results, Large Card Sort](#)

[Optimal Sort Results, Small Card Sort](#)