



Digital Scriptorium



NATIONAL
ENDOWMENT
FOR THE
HUMANITIES

Digital Scriptorium

DS Catalog Search Interface Review

13 March 2026

*The National Endowment for the Humanities and Digital Scriptorium together:
Democracy demands wisdom.*

Any views, findings, conclusions, or recommendations expressed in this report do not necessarily represent those of the National Endowment for the Humanities.

Introduction

Digital Scriptorium, through its institutional partner at the University of Pennsylvania Libraries, received a Humanities Collections and Reference Resources implementation grant from the National Endowment for the Humanities to continue and expand our work in building the DS Catalog. This work designated through the grant focuses mainly on creation of original metadata descriptions for previously undescribed manuscript objects and increasing the number of records available in the DS Catalog. Additionally, the grant is intended to support projects that encourage reuse of DS data as well as outreach and education opportunities for member institutions to increase their knowledge about DS Catalog workflows and infrastructure. Elements of this outreach will include workshops and presentations about the DS data model, the structuring of DS data, how to use the DS Wikibase and search interface for improved discovery, and how DS carries out data enrichment and reuse workflows and processes.

In our grant application, DS indicated that in the early months of the grant, a review of the search interface (<https://search.digital-scriptorium.org/>) would be conducted and a report would be shared with the DS Board of Directors and the public. This report endeavors to keep the Board and our community informed about grant-related activities and the completion of workplan goals and deliverables. The following report marks fulfillment of a grant deliverable to assess the DS Catalog search interface:

DS Catalog search interface review process

Through an ongoing “bug report” form made available through the search site and email communications with individual users, DS staff have collected user feedback regarding issues and potential improvements. Many of these issues were resolved through redevelopment work completed prior to the grant award period by independent software developer Michael Holloway, whose work was integrated into the DS technical infrastructure in early 2025.

In an effort to collect more data to inform the interface review discussed in the grant application, PDM Coladangelo conducted two focus groups with the DS Advisory Council held on September 22 and November 10, 2025. Questions to the focus groups were targeted around interface features and affordances as well as the nature and structure of the available data. Detailed notes were recorded, compiled, and synthesized by the PDM, generating the following report. The report discusses focus group observations contextualized by decisions made during the DS 2.0 redevelopment project and the structure of the DS data model. The report concludes with recommendations for future development and improvements in the technical infrastructure of the DS Catalog.

Overall interface aesthetics and appearance for usability

Both focus groups found the appearance to be clean and pleasing in appearance. The general design was noted to be similar to other institutional catalogs and was not overwhelming users. The ready availability of facet options made them helpful for users, suggesting possible access points for browsing. It was suggested that facets might be reordered to emphasize certain aspects which would be of value to users and to deemphasize others “below the fold.” Facets recommended to appear toward the top of the list for primary access included the presence of images and more “tangible” or “situational” elements like language, material type, and place. Facets for names (author, scribe, etc.) and titles were recommended to appear closer to the bottom of the facet menu list. The holding institution was suggested for lower display in the menu since the approach of the catalog is to be “agnostic” and to provide a union catalog experience (not emphasizing one collection over another).

Having drop-down options from the search bar was found to be helpful to prompt users on what elements they can search on. It was suggested the search box be more visible or in a different color to emphasize its presence on the page, and perhaps oriented on the right of the page rather than the left.

Facilitating and improving user experience

Design elements

It was suggested much of the useful information appearing below the fold might be getting lost. This information might be provided or accessible to users in

different ways. For instance, we might include additional menus in the header (like a help page) or draft an FAQ or search guidance that can be linked to. It was also suggested that a banner might allow for more flexibility, providing links to pages to learn more about the project, to explore lesser known things (e.g., a changing banner element to make the interface more active), and a separate page for search ideas (e.g., a resource to suggest searches for new or beginning users)

While focus group members found the facets useful, they suggested ways to improve facet display.

Currently, the facet menu automatically opens the first face (in this case, Holding Institution). It was suggested that this be avoided (i.e., close all facets and allow the user to choose when to open facets) or randomized.

Focus group members also wanted to more usefully sort the facet/browse options and to have the system indicate the number of options available. Focus group members preferred the ability to alphabetically sort for certain facets, such as names. They also wished to have page numbering on facets or indicators of number of options per page.

Holding institutions are currently sorted by name; there was curiosity if that was the best method or if they might be sorted by place. *(PDM note: This was an unimplemented area of the original prototype data model. We might consider revisiting the idea of linking institutions to a Place authority record within the DS Catalog Wikibase in order to facilitate this information. It would also require a revision in the search site pipeline to make this information usable for search and display.)*

Some focus group members had concerns about the display of data values in facets where there were different levels of authority control. For example, Material is an area where data values are highly controlled (i.e., using Linked Open Vocabularies), and thus well-suited for faceting. Conversely, it was suggested other facets where some data values are unreconcilable (some Subject strings rendered as Keywords) might be undermining efforts to display that information in facets.

Content elements

Some data elements were found to be confusing or counterproductive to search and retrieval efforts. Overwhelmingly, this was the feedback regarding the display and faceting of title information. *(PDM note: Because title information is so varied between institutions, record formats, data standards, and cataloging practices, it is difficult to aggregate and harmonize these values without creation of an*

authority, which has been done on an ad-hoc basis as records enter the DS Catalog.) A result of aggregating data from different systems and generated through different practices is a lack of clarity on how the Standard Title functions as assigned in DS records. It was generally agreed that the current approach to title standardization needs to be revisited, or perhaps the Title facet needs to be renamed or clarified.

It was also suggested that a FAQ or guidance document could be created to explain Standard Title assignment and function, including rules for search and display within the DS Catalog. This FAQ could help to describe the use of other authority values as well and to provide transparency about the reconciliation and enrichment process, such as what it means when a data value is listed as “undetermined” or “unknown” in a record. *(PDM note: This specific data value, for instance, arises from the presence of an unspecified data value in the institutional record itself. So, if “Place of origin unknown” or “?” appeared in an institutional record, in the DS record it would be reconciled to an authority value of “undetermined” (Q650) in the DS Place authority. If the record had no data value for place information, that information would simply be absent from the DS record.)* It was also suggested the FAQ could include a glossary to better explain data elements (i.e., what is a “Title” or “Keyword” in this context?). The FAQ would be a place to describe DS practices for data enrichment and how those affect search and display.

The facet label “Associated agent” was found to be too jargony and likely something students would not understand. This could be a simple label change like “Other name” (or explained in an FAQ).

The 1st century label in the Century facet needs to be fixed. It currently only reads as the number “1.”

The availability of date or date range information as a facet as discussed. *(PDM note: PDM gave the focus groups context about limitations in both the source data and the Blacklight options (i.e., software “gems”) for date range search and browse functions. During the prototype phase and development of the search interface, no good option was found for sorting date or date range information. As a result, a date facet was abandoned in favor of guiding users to a SPARQL query in the DS Wikibase Query Service that could be used to search date ranges and specific date values. This approach could be better communicated as available in a SPARQL query through documentation.)* Additionally, while records are filterable by century, sometimes keyword options also include century information, which might be helpful in narrowing searches by date range.

Account creation and bookmark features

Focus group members had not created accounts (or did not recall creating accounts) and were not very familiar with the bookmarking feature. When described, they agreed the process sounded easy to create and use.

One focus group member noticed that citation information was missing from the generated Cite template, which may be a pipeline display issue that needs examining.

It was suggested that perhaps a feature could be added to generate a single link to all bookmarks to categorize bookmarks for later browsing. It was also suggested to add a checkbox feature allowing a user to select or bookmark all search result options or all options on a page.

Search and browse features

Suggestions for metadata field inclusion

The focus group indicated a desire or ability to search by the following fields/elements:

- script type
- item type (i.e., codex, leaf, cutting, fragment)
- document type
- a “real title” option
- alternative versions of titles
- date range
- line count
- measurements
- dimensional range

It was recognized that in conversation with the PDM that the availability of metadata elements, and thus options for search fields, would be determined by the structure and parseability of existing institutional metadata. There were concerns that students or those not familiar with the system and the limitations of the data might be discouraged by unfruitful searches (i.e., “every search box is a Google search”). Developing an FAQ or search guidelines could be used as a positive tool to assist discovery by helping users to understand limitations of the data and the difficulty in parsing or representing relatively unstructured data. A guidance document could inform users what is immediately available as structured data and what is not easily parsed. One example of this kind of information would be understanding where dated information and century information comes from in the data enrichment process.

Quick search results display information

The focus group generally felt that quick search results gave enough information to allow users to make additional search choices and record selections. Potential improvements

suggested included format information (e.g., codex, document, fragment, etc.) and number of leaves, if available.

It was also suggested that the FAQ/guidance document could clarify the ordering of search hits based on extant facet options.

Record/metadata display

Focus group participants generally found the metadata record display to be useful, appreciating that it appeared much like a traditional library catalog record. Some had concerns about the visual clarity of the long bar for enriched values, which includes an internal faceted search (cued visually by magnifying glass) and link out to a Linked Open Vocabulary term (cued by a network symbol), which was found to be a positive (if sometimes “arcane”) feature. An FAQ/guidance document would be helpful in providing clarifying information for what is represented in the visual display and contents of the record.

There was some confusion about the location of links to institutional records, which may need to be evaluated for further UX improvements. *(PDM note: The institutional record link appears as a clickable link on the right-hand side of the page and as a copy-and-pastable link toward the bottom of the record.)* It was also noted that the presence of a link to the member directory to get contact information for institutions was helpful.

Finally, there were continued concerns about title extraction, title search and display features, and the assignment of Standard Titles. Comparing a DS record against a MARC record, for instance, revealed that title information, especially for the contents of compound or non-Western manuscripts, is often represented in MARC 505 or ZXX \$t fields. Title information not extracted and republished from these institutional records could be preventing users from finding manuscript objects relevant to their research. *(PDM note: Since the focus groups have concluded and the MDCF has been onboarded, multiple internal discussions by DS staff have concluded that the criteria and process for title information extraction from MARC records and Standard Title assignment during the enrichment process need to be reevaluated. DS staff have taken steps to plan technical improvements to title extraction. The PDM and MDCF will also begin a far more conservative process for Standard Title assignment, more often foregoing the use of Standard Titles and deferring to cataloger-generated titles. This process will take some time to implement as records are refreshed/updated in the DS Catalog and title enrichment is revised in the DS data dictionaries used for automated data reconciliation.)*

Future action items

Based on the focus groups observations and conversations, the PDM recommends the following actions:

1. Reexamination of Standard Title enrichment. This should include the general use or assignment of Standard Titles, including adoption of a more conservative approach to assignment. Deference to displaying cataloger-supplied titles and avoiding Standard Title assignment where specificity of work/genre cannot be determined will guide future practice. Implementation will occur through revision of data dictionaries on an ongoing basis as records are updated/refreshed in the DS Catalog.
2. Expansion of title extraction from MARC fields beyond current practice. This should include additional MARC fields and subfields (e.g., 246, 7XX \$t) where potentially relevant title information can be stored.
3. Changing the order of facets in the left-hand facet menu to reflect focus group recommendations.
4. Correction of minor display errors which may affect the quality of the search experience and confidence in the DS Catalog.
5. Production of an FAQ/guidance document which discusses the nature and context of data presented in a DS record context. Such a document should explain to users where data originates, how it is chosen to be displayed, and how to navigate DS metadata records.
6. Reevaluation of the DS data model with regard to data fields/elements not currently represented in DS records. Aspects to be considered include data of special use to researchers, availability and machine-actionability of such data, and workflow and labor considerations for processing and maintaining enriched data.

Respectfully submitted,

L.P. Coladangelo
DS Catalog Project and Data Manager