State of the Collections: Measuring the Baseline Discoverability and Accessibility of The Huntington Library's Archival Collections

Introduction

Until 2022, The Huntington Library did not have a comprehensive overview of the size of its archival processing backlog, which collections were in the backlog, or a breakdown of the discoverability and accessibility of its archival collections. Furthermore, the library could not assess the rate of acquisitions coming in annually against collections being opened for research and how or if our backlog was changing over time.

This lack of information inhibited the library's ability to make fully informed decisions related to processing, collection management, acquisitions, and staffing as well as assess levels of demand for unprocessed collections. Library patrons were also unable to fully search the library's archival holdings.

As an institution that has been collecting and describing its collections for over 100 years, the concept of a backlog and how to assess it was complicated by huge shifts in standards, practices, and systems and tools that occurred over decades and piecemeal retrospective projects. Two main limiting factors in understanding the backlog were that no single system had records for all of our collections from which we could create a definitive list and the local standard for expressing collection extents was most often an item count.

Maggie Hughes, mhughes@huntington.org
Archival Processing Manager
The Huntington Library

Project goals

1. Provide information on the processing backlog
2. Provide high level data on the discoverability and accessibility of collections
3. Lay the groundwork to assess and strategically address both
4. Consolidate and update collection location data

Methodology

In order to understand the library's archival backlog and the discoverability of those collections, we needed to know all of the collections in our holdings, if they were processed or not, and if they had online records. In order to get this data, we completed a multi-year survey of the library's main archival stacks locations.

Prior to beginning the survey, we updated the local extent standard to linear feet with a container summary, moved collections that weren't stored in permanent shelving locations, and cleaned and labeled stacks areas when needed. We used this as an opportunity to deprecate redundant shelf lists and establish a definitive shelf list for tracking collection locations. During the survey, staff moved through each stack space, shelf-by-shelf, collecting data. Each collection was added to an excel spreadsheet with core identifying information and a high-level assessment of its accessbility and discoverability. We also verified materials locations in the shelf list, adding or updating it when necessary.

We did not do research outside of the stacks and instead relied on the information at hand as this was first pass and it allowed us to move more quickly. Assessing whether any processing work had been done on a collection was a judgement call. "Some" indicated that it was clear some work had been done on the collection, but it either had unprocessed portions or some fundamental issue(s) that would require staff attention.

DATA COLLECTED

For each collection, the team gathered and noted:
- Core identifying information:
  - Title
  - Accession number (if readily available)
  - Call number (if assigned and readily available)
  - Linear feet
  - Container summary
- Assessment of accessibility and discoverability:
  - Does it have a record in our OPAC? Yes/No
  - Does it have a record in WorldCat? Yes/No
  - Does it have a finding aid online? Yes/No
  - Is the collection processed? Full/Some/None

STAFF AND TIME COMMITMENT

In total, 11 staff members helped to survey, and the team (6-8 people at a time) spent over 40 full work days surveying. Staff surveyed the 3 main stacks locations from August 2019-March 2020. In March 2022, we surveyed 2 additional smaller stacks locations.

Known gaps in data

- Excluded some very small collections by sometimes adhering to a previous local definition of a collection as being 40 items or more.
- There are a few small stacks locations that we have not yet surveyed.

Related projects

1. Project to hand-key legacy paper finding aids and publish them online:
   a. Resulted in 368 new online finding aids.
   b. All have updated core data, including container lists to improve paging efficiency and reduce need for staff mediation.
   c. Helped to establish ASpace as our system of record for collection description.
2. Hired first accessioning archivist in 2021 who implemented accessioning-as-processing workflow to reduce number of new acquisitions going into the backlog. Hope to further expand on this work.

Future projects

- Isolate processed collections without an online finding aid from spreadsheet list and create online finding aids with container-level lists.
- Retroactive accessioning project to give all collections at least a minimal collection-level MARC record to meet baseline discoverability.
- Identify collections in the backlog that can be minimally processed.

Results and Findings

We have a list of all of our collections and can better understand our backlog and its needs to inform future work. We continue to maintain this data -- adding new aquisitions, adding any materials we discover that were missed in the survey, and keeping the status of collections up to date. We can now track the state of our collections and understand how it's changing over time.

The chart to the right shows the breakdown of the collections into 2 broad categories:
- **Visible and accessible**: Has a record in the OPAC, in WorldCat, an online finding aid, and was assessed as being fully processed.
- **Hidden and/or backlog**: Was missing one of more of the accessibility or discoverability elements.

Breakdown of the collections in the hidden and/or backlog category

Legend: 2022, 2023. Categories: No OPAC, No WorldCat, No finding aid, Not fully processed. X-axis: 0 to 2,000.
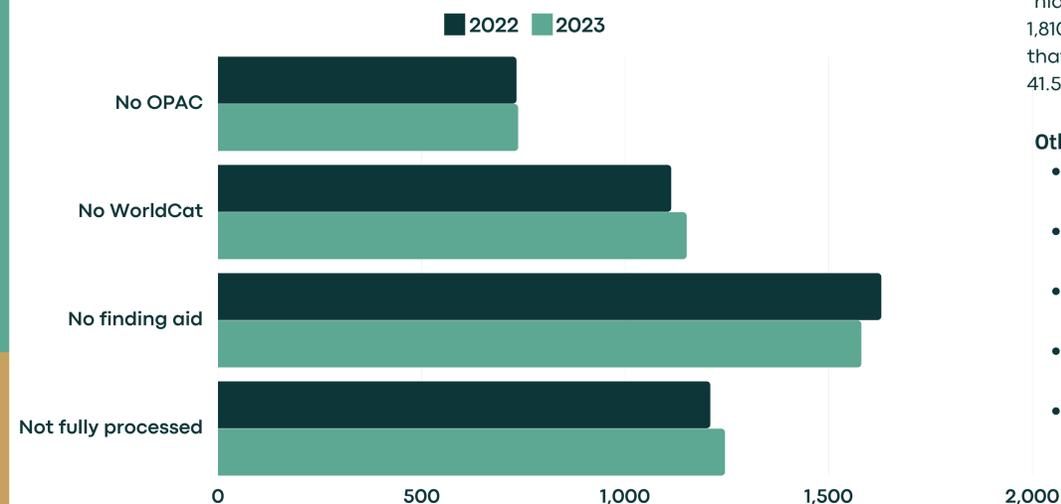
Number of collections in broad categories

Legend: 2022, 2023. Categories: Visible and accessible, Hidden and/or backlog. Y-axis: 0 to 2,000.
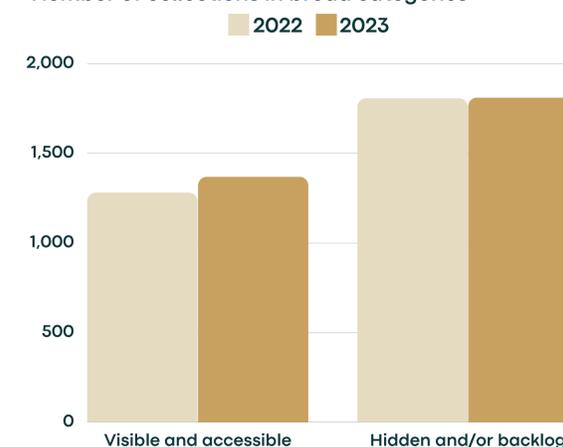
In 2023, while the number of collections that are "hidden or backlog" increased by 4 (from 1,806 to 1,810), the overall percentage of the collections that are visible and accessible increased from 41.5% to 43%.

Other findings of note (as of 2023):

- 64 collections are processed but do not have a record in the OPAC.
- 507 collections are processed but don't have an online finding aid.
- 75 collections have a finding aid but were assessed as not being fully processed.
- 547 collections that are not fully processed are 1 ≥ 5 linear feet.
- 273 collections that are not fully processed are <1 linear foot.

Select References

UNLV University Libraries Special Collections (2015-2016). UNLV Archival Collection Survey documentation. Retrieved July 7, 2023, from https://drive.google.com/drive/folders/0B25SSpmxsyXOLVp3UjZSWFJMR00?resourcekey=0-Lg6cQ-aC2JNmfi719aKW2g

Santamaria D. A. (2015). Extensible processing for archives and special collections : reducing processing backlogs. Neal-Schuman an imprint of the American Library Association.

Weber, Chela Scott, Martha O'Hara Conway, Nicholas Martin, Gioia Stevens, and Brigette Kamsler. 2021. Total Cost of Stewardship: Responsible Collection Building in Archives and Special Collections. Dublin, OH: OCLC Research. OCLC Total Cost of Stewardship.

THE HUNTINGTON
Library, Art Museum, and Botanical Gardens