Leveraging Preservation Practices During Large Scale Relocation and De-duplication

Miriam Nelson
Preservation, Ohio University, Athens, Ohio, USA

Katy Mathuews
Collections Assessment & Access, Ohio University, Athens, Ohio, USA.

Miriam Intrator
Archives and Special Collections, Ohio University, Athens, Ohio, USA.

Damon DeBorde
Metadata Services, Ohio University, Athens, Ohio, USA.

Copyright © 2017 by Miriam Nelson, Katy Mathuews, Miriam Intrator, Damon DeBorde. This work is made available under the terms of the Creative Commons Attribution 4.0 International License: http://creativecommons.org/licenses/by/4.0

Abstract:

Ohio University Libraries has embarked on a project funded by an LSTA grant which will identify and relocate 40,000 of its scarcely-held and locally significant holdings. Currently stored in the Southeast Ohio Regional Library Depository, these items will be moved to a vendor based storage facility 80 miles from Ohio University’s main campus. The goal of this project is to pre-emptively address the long-term retention and preservation of the Libraries’ most distinct collections in the face of budget reductions for the state depository system.

In this paper we will discuss how the grant focus on unique collections surfaced gaps in the Libraries’ collection management strategy, the resultant limitations on data driven analysis, as well as the preservation and retention implications of identifying and moving large numbers of scarce and sometimes unique monographs so far from users. The primary focus of the paper is the Ohio University Historic Book Project (HBP) which grew in response to the challenges posed by the grant project and the anticipated funding reduction.

Keywords: Preservation, Relocation, Special Collections, Storage, Retention

INTRODUCTION

Ohio University Libraries (OUL) is a member of the Association of Research Libraries, the Center for Research Libraries, and a founding member of OhioLINK. OUL facilities include the Vernon R. Alden Library, the Music and Dance Library, and the Southeast Ohio Regional Library Depository (SEORLD). (ARL, n.d.; CRL, n.d.; OhioLINK, n.d.) The Libraries’ world-class collections support the University’s tradition of academic research and contribute to its strengths, its history, and its international focus. As such, OUL is deeply committed to the long-term preservation and access of its distinct collections.
The Southeast Ohio Regional Library Depository is one of 5 Harvard style high density storage facilities in the state. Opening in 1999 SEORLD sought to resolve the competing demands of growing collections; the competition for space in the heart of campus; and the goal of long-term retention and preservation of University collections. The SEORLD operating budget and capital improvement costs are funded by the state of Ohio and supplemented by a grant from OhioLINK, itself a primarily state-funded organization. Operating budgets for the depositories have been flat for the past 7 years. Recent budget shortfalls, however, have resulted in the first year of funding cuts, with a projection of further reductions. These budget constraints have threatened the sustainability of the regional depository system. In order to proactively adapt to possible reductions in funding, OUL, with support from OhioLINK, has embarked on a project to analyze its collection with the goal to reduce the need for auxiliary storage and eliminate reliance on the SEORLD facility. The project is funded by an LSTA grant and will identify and relocate monographic holdings unique within the depository system, the state of Ohio, the Midwest region of the United States, and the nation. The goal of this project is to address the long-term retention and preservation of the OUL’s most distinct collections. (IMLS, n.d.) These scarcely-held items will be moved from SEORLD to a vendor based storage facility 80 miles from Ohio University’s main campus.

In this paper we will discuss how the focus on unique collections surfaced gaps in OUL’s collection management strategy, which does not explicitly deal with “medium-rare” material as a distinct category. We will also address the preservation and retention implications of identifying and moving large numbers of scarce and possibly unique monographs so far from users while also eliminating the mediation of OUL staff in the retrieval process. The primary focus of the paper is the Ohio University Historic Book Project (HBP) which has proceeded in tandem with the LSTA grant project in order to address the aforementioned gaps and leverage this opportunity to move OUL’s policy closer to best practice.

THE CHALLENGE
To identify the Libraries distinct collections, particularly those held within SEORLD, OUL contracted with OCLC’s Sustainable Collections Services (SCS) to perform data-driven collections analysis. SCS provides decision-support through the identification of potential withdrawal and retention candidates at both a title and edition level. After SCS matches OUL’s bibliographic data against target data (i.e., WorldCAT, Hathi Trust, and comparator library collections), bibliographic holdings, and circulation data are available in GreenGlass®, SCS’s web-based, self-service collection analytics tool. Utilizing GreenGlass® OUL identified a target number of 40,000 items or 6% of the SEORLD holdings to be considered for permanent retention. Based on the above analysis this number represents a compromise between retaining the most unique items in the collection and the need to dramatically decrease dependence on the SEORLD facility. Though GreenGlass® is a powerful analysis tool, it cannot account for copy-specific characteristics or interventions such as bookplates, signatures, inscriptions, marginalia, special bindings, or gift notes. In short, while the analysis is able to identify title and edition level duplication, it is generally blind to many of the characteristics of local copies that interest researchers, collectors, rare book librarians, and preservation professionals. Compounding this problem are uneven practices that have governed processing, description, and preservation over the course of OUL’s 200-year history. For example, at times new bookplates were pasted down over old with up to 4 or 5 layers found on a single item and at
other times bookplates were ripped off entirely. Past binding practices, liberal use of book tape, and highly interventive repairs have frequently destroyed any trace of interventions that may have been of interest. Further, while many books still retained their unique characteristics, few have been described at a sufficient level of detail to be discoverable via the integrated library system (ILS) or any future SCS analysis. Because of this significant gap in our ability to rely solely on the available data, it was imperative to develop a scalable process to physically search for and identify as many items with significant interventions as possible. The goals governing the search, known as the Ohio University Historic Books Project (HBP), were: to keep significant items from being subject to large scale de-duplication (an explicit goal of the LSTA grant project); to identify items that should be integrated into the rare book collection; to provide adequate description of their unique characteristics; and to develop a retention statement that would exempt these items from future weeding and serve as a commitment to our partner institutions. An emergent goal of this effort has been to leverage the Libraries’ grant project and the findings from the HBP in order to formally integrate preservation best practices into the complete lifecycle of our collections.

**METHODOLOGY**

The design and scope of the HBP was influenced by Book Traces @ UVA, a project of the University of Virginia Library with support from the Council on Library and Information Resources (CLIR). Book Traces @ UVA developed through the University of Virginia Library’s collaboration with the existing Book Traces project founded by Andrew Stauffer, Associate Professor of English at UVA. Professor Stauffer, a staunch advocate for the research value of what are often referred to as “medium rare” books, recognized that the artifactual value of these resources was under threat due to mass digitization and subsequent de-duplication of general circulating collections. Professor Stauffer demonstrated that a wealth of data could be collected through the serendipitous exploration of the library stacks. Subsequently Kristen Jensen, of UVA Library developed a process to allow for a systematic search and statistical analysis of the data collected. Jensen generously shared this data collection process with us so that we could test the UVA protocol within OUL’s collections. The protocol focuses the search on collections published before 1923 (before U.S. copyright) and employs a Google Form for flexible data collection. We saw both iterations of Book Traces as a possible framework by which we might test the presumption that all material held in SEORLD could be de-duplicated or relocated without copy level assessment. It was also a way to surface questions about security and preservation as they related to the relocation process. With approximately 700,000 items to evaluate, the focus on pre-1923 publications was particularly appealing. In addition to significantly limiting the scope of the search, collections out of U.S. copyright represented an at-risk class of item, as many are widely available through the digitization efforts of Hathi Trust and Google Books. These items were also most likely to represent the early history of the Ohio University Libraries own collections, which were created by the joining of two 19th century literary societies and gifts from early presidents, deans, professors, and other founding figures of Ohio University. We hoped that some of the Libraries earliest acquisitions would be found in those 19th century collections.

The first step was to identify how many items could potentially fall into this developing category of historic and/or locally significant books, as it related to feasibility and the probable impact on the 40,000 item limit mentioned above. Though we knew there was potential for error in limiting our list by publication date, due to variations in entries in that MARC field, in the end it became the most practical filter as it was accessible both in Sierra (our ILS) and GreenGlass®. A preliminary list generated from the ILS
was used to create a sample from which we would produce a time test and small survey of what we might be likely to find. We decided to begin with a sample of 200 books which would be pulled from the shelves in order to test the protocol shared with us by Kristin Jensen. Physically pulling the books would allow us to quickly compare the recorded data to the book in hand and assess whether or not the data collected accurately represented the physical evidence.

Limiting to a pre-1923 publication date we found 25,523 items that would require physical inspection. Our sample quickly revealed that the time spent per book, an average of 3 minutes, and the quality of the data we were recording did not correlate favourably thus would not scale effectively given our available resources. While we could commit 40hrs/week of FTE to the project, at that rate it would have required 32 weeks to complete the project. Constrained by the LSTA grant, we did not expect to have quite so much time. Because our primary goal was to identify and relocate items as quickly as possible, we knew that we must limit time spent in the identification stage. Additionally it was the preference of both our Head of Metadata Services and the Special Collection Librarian that enhanced description be performed with the book in-hand. Therefore we would either need to route all items through Metadata Services before relocating or rely solely on standardized local notes applied through bulk edits, leaving more detailed description for a later stage. This additional high-touch process was also a limiting factor.

One of our initial hopes in conforming to the Book Traces @ UVA protocol was the possibility of contributing OUL’s findings to any future collaborative projects. However due to the lack of a sizable and formally recognized area of bibliographic study comparable to that of UVA it was agreed that we would have to streamline the data collection to focus on what was defined as locally significant interventions. The focus on local significance was validated, as many of the items we were finding were complementary to the collecting areas within the existing Special Collections. While we knew that it would not be possible, or desirable to relocate all such items to the rare book vault, we felt strongly that monographs with connections to the history of the University or the region represented the most compelling argument for permanent retention. We were also compelled to lower the publication date threshold from pre-1923 to pre-1900. Doing so lowered the search field to 16,853 items a number we were confident we could accommodate in our timeline.

Table 1: Collections limited to pre-1900 in the general collections stacks and the Annex.

HBP’s modified protocol relied on Excel to streamline data collection and facilitate any bulk edits performed by Metadata Services. The spreadsheet utilized the dropdown list function to maintain a controlled vocabulary. These vocabularies were used to populate the cells in
3 columns, “primary intervention”, “secondary intervention” and “bookplate”. An item had to display a characteristic within the primary intervention column in order to be flagged as a HBP book. Further interventions could be described in the secondary intervention column, the bookplate column, or in the free notes. This process created a hierarchy that allowed us to limit the time spent per book and privilege those interventions with local significance. We were also able to include a directive to immediately relocate items to Special Collections when books with particular associations were identified (See Table 2).

Table 2, Left: Separate sheet for controlled vocabulary Right: Data entered into the primary sheet. Book found with the name “Henry True” or “H.A. True” were flagged to be moved to special collections as the represented an early gift of high significance to OUL’s history.

**FINDINGS**

Our shift to a more targeted focus on OUL history and pre-1900 imprints proved fruitful in terms of the success we saw in identifying books with significant associational value to the University and its Libraries. The inadvertent weeding or deaccessioning of these books would be a true loss to our current and future ability to research, write, understand, and teach our own history and that of the individuals, organizations, and institutions that have been vital to its growth and development.

Key among these were the **Athenian Literary Society** and the **Philomathean Literary Society**. Both were student organizations associated with the University established during the 1800s; both developed and kept libraries for their members. The two libraries agreed to merge with OUL in the late 1880s to form a single library. Over the course of our search we have identified a wide variety of bookplates and inscriptions used by these societies to mark and identify their books. As they are some of the earliest books added to the library that are still in the collection today, we have deemed them not only important to retain, but also integral enough to our history that we are moving all of them into special collections. In addition, the records of the two societies, including accession logs and other materials documenting their use, are held in special collections. We hope to one day be able to determine which books came into which library and when, thereby gaining deeper understanding of the origins of our current library.

The next category of focus also regards provenance. There are certain names, collectors, marks of ownership, and donations that are already present and familiar in special collections. We decided that when those same names are found in the general collection, they should be relocated. These include books from the collection of **Henry Ayer True** of Marion, Ohio.
True was particularly interested in 19th century Juvenile literature. He often marked in his books, signing his name, noting where and from whom he purchased or was given a book, and often identifying authors, many of them female, who at the time wrote anonymously or using pseudonyms. William E. Peters was a prominent figure in late 19th and early 20th century Athens Ohio and at Ohio University. Peters (1857-1952) was a land surveyor from 1878 to 1888, Athens County surveyor from 1888 to 1891, and a practicing attorney in Athens from 1891 to 1939. He also did an enormous amount of descriptive, cartographic, and photographic research on Athens County, leaving us with unique records of the history of the area. He collected books related to Athens County and Southeastern Ohio more broadly, including textbooks owned by local Ohioans. Many of his books are in a distinctive black binding with the title, author, and “Peters” stamped in gold on the spine. Peters also often left his ownership stamp on the title pages of his books. Due to Peters’ significance to this region, the books he used for his research are of great interest to us and of potential interest to future researchers. Therefore any Peters books were also flagged for relocation into special collections.

RELOCATION IMPACT
Having shown that there are a significant number of items within SEORLD deserving of special consideration, the next step was to ensure their storage and handling better conform to best practices for their long-term preservation. To achieve this we have considered the following questions: Should new labels be applied or old labels be removed? Should these items universally receive enclosures for added protection? Is this collection a high priority for preservation quality environmental conditions? What level of description should these books receive and how will that impact their security? Will better description of their unique value make them targets of theft? The answers to these questions do not just impact the books identified by the HBP, but will also govern how OUL will treat similar items going forward, as well as influence how the 40,000 monographs identified by the LSTA grant project will be treated. Ultimately these procedures will constitute Ohio University Libraries Medium-Rare Book Policy as well as its procedures for collection relocation.

Before procedural questions could be addressed we had to determine if the HBP items would join the collections at the vendor facility. A preservation risk assessment proved the best strategy to determine the most appropriate storage. Because we could make some assumptions about the quality of paper and the average condition of the HBP items, we could logically consider them all to have a similar level of inherent vice and would, thus, rank the condition-based-risk to be high. Similarly the HBP items themselves are, by definition unique, and their local significance ranks their research value as just below that of our special collections. These characteristics place them at risk from damage from poor environmental conditions and theft, as well as mark them as a high priority for preservation resources. While the likelihood of theft is variable, the threat of poor environmental conditions is perennial. Improvement to the environment would certainly have the greatest impact on the long-term preservation and accessibility of these items. While such conditions remain elusive on campus the most logical approach was to recommend that all HBP items be sent to the vendor facility which provided the security of a closed stack system and preservation quality environmental controls.

The next step of the project required us to consider what preventative measures must be taken to ensure that our processes are consistent with the value of the items and their associated risk. Because these books were already processed as a part of the general collections they
have been stamped, barcoded, and labeled. It is not feasible to remove these markings and in
the spirit of provenance it may not be consistent to do so. Therefore the recommendation
of the HBP was to utilize special collections marking procedures as outlined by RBMS for
any additional processing required. (ACRL-RBMS, n.d.) Corrugated clamshell boxes were
also recommended for all HBP books in order to ensure safe storage in a high density
environment and during transport. Because the value of these monographs is tied to their
local significance, it is unfortunate to see them leave our campus. However, we have an
opportunity to prioritize their enhanced description so that our users might discover the
interventions through the catalog, by their own research, or through the help of our subject librarians.

Processing, enclosures, and enhanced description will certainly cause delays between
retrieval and relocation. As we proceed it is only logical to suspend access to these items
during this period. However, utilizing the Excel spreadsheet we will be able to do this easily
through a bulk edit either making them available again as they are accessioned into the
vendor facility or once all items have been received. While enhanced description might
increase interest in these items, we must proceed with the assumption that they will continue
to be low use. Should interest increase, we may consider digital imaging to create surrogates
of the interventions, including only as much contextual information as necessary. At
that point we may also wish to consider a modified circulation policy to restrict use.

COLLABORATOR PERSPECTIVES
Special Collections
From the perspective of special collections, the goal of this project was to attempt to identify,
describe, stabilize, and protect any and all valuable books in Ohio University’s Libraries.
Definitions of value vary wildly and can be highly subjective. Our primary
targets were books with important associational value to Ohio University. We were on the
lookout for interventions that provide evidence of provenance related to the history of the
University. In addition, we were including limited edition books, fine press or fine bindings,
or books having content or authorship of particular interest to the University’s teaching,
learning, and research missions, and particularly those related to Southeastern Ohio, women’s
history, and African-American studies.

Moving books into the rare book vault has presented challenges regarding workload for
metadata services and of course, space. There is very little room to grow in the rare
book vault’s current arrangement so this project has pushed us to create space, which is an
ongoing process. The primary ways in which we are doing this is by moving select items out
of special collections that can and should be circulating (this includes current periodicals,
contemporary secondary and tertiary sources that are freely available and easily replaceable,
and stores of rarely-used antiquarian book dealer catalogues that can be accessed online or
through inter-library sharing and exchange programs such as OhioLink and InterLibrary Loan).

Despite the challenges, we are greatly encouraged and excited by what we have discovered
and will be able to save from any near- or far-future deaccessioning and weeding projects. In
addition, we are getting to know the collection in new ways. This deepens the knowledge of
the resources we hold, the stories they tell and the histories they represent, improves
the instruction, research, outreach, and promotion services the Special Collection Librarian
is able to provide to students, faculty, and the general public.
**Metadata Services**

For metadata services, the goal of this interdepartmentally collaborative project was, simply, to rise to the cataloguing challenges presented. From a metadata point of view in particular, the HBP was focused simultaneously at the collection-level and the item-level. Correspondingly, metadata services pursued methods to enhance records through bulk editing procedures in addition to item-by-item processing. The first step was crafting an Action Note (i.e., MARC field 583) indicating OUL’s commitment to retain. A 583 field (containing the action “committed to retain,” the date, and the project name “OUL Historical Books Project”) has been added to records individually or through the bulk record editing tools in the ILS. This first action metadata services considered as withdraw-proofing: workflows were updated to include filtering out any record with a 583 field with the above wording from withdraw candidates for any/all weeding projects. A 710 MARC field with the name OUL Historical Books Project was also added. This afforded a quick method, whether to researchers at the OPAC or staff on the back end, to access all items touched by the project.

While 583 and 710 fields indicate a commitment to retain and offer additional access, the fields do not express why the item is retained. To capture and share the various interventions discovered, several other MARC fields and standards were adopted. In particular, MARC fields 561, 590, 655, and 700 (Ownership and Custodial History, Local Notes, Genre/Form Terms, and Personal Name Added Entry, respectively) were utilized. These fields relay the notes recorded in the spreadsheets as cataloguing metadata, attempting to enhance the usual goals of description and access. For example, the 655 field utilized Rare Books and Manuscripts Section (RBMS) Controlled Vocabularies to follow established standards and the 700 field will record authoritative names and RDA Relationship Designators, offering not only enhanced access points for researchers but establishing OUL records as consistent with the current best practices in the field. While these descriptive fields can be challenging to implement clearly and consistently, time and logistics emerged as the most important factors to metadata work as the project unfolded.

The refinement of the HBP upper publication date limit from 1923 to 1900 narrowed focus and, effectively, increased the ratio of significant discoveries to routine items. As spreadsheets grew, it became clear that not the discovery but the cataloguing of items would be a probable bottleneck. Backlogs, for better or worse, are relatively commonplace in cataloguing departments; yet, the items are often near at-hand and awaiting ingestion rather than potential deaccession. As discussed in the Relocation Impact section above, important materials identified through the HBP will end up at off-site storage. Logistically and emotionally, that is a challenging backlog. As ILS tools were employed to bulk add and edit the MARC 583 and 710 fields, the same tools are being investigated for adding and editing MARC fields that carry the intervention information (again, 561, 590, 655, and 700). The adoption of established best practices ameliorates this process as autographs, inscriptions, bookplates, etc. are described with standardized terms from controlled vocabularies and MARC bibliographic standards.

At present, metadata services continues investigating patterns and trends with the significant items, anticipating which items may need hands-on work and which items can be enriched successfully remotely. The project has spurred the refinement of technical skills and the enhancement of workflows. Yet, the department is even more pleased with playing a role in protecting our unique materials and striving to assist colleagues in moving toward making collection management decisions at scale.
**Collections Assessment & Access**

This project had very practical applications for collections assessment and access staff, both at the project-level and as a means to inform broader practices. Collections assessment and access staff used the Excel spreadsheet created within this project to physically examine items and record intervention information. This method was helpful as it allowed staff to move efficiently through the stacks with a laptop to record information. The true value of this analysis, however, allows for efficient planning of collection relocation. The material not selected for storage at the offsite facility will be withdrawn or relocated to the main library’s general stacks or special collections. Identifying the number of historically significant books that may be relocated to the general collections allows collections management staff to proactively create the space necessary to store the items. This is particularly important given that books may be located in a wide variety of call number ranges, with some subjects containing more books than others. For example, many historic books identified in this project are located in literature and history call number ranges. Having an idea where our historically important materials are classified helps us plan small- and large-scale shifting projects to allow space for relocation of materials in specific call number ranges. The information identified in this project also helped to open conversations about implementing a systemic deselection system to provide a more proactive approach to storage space.

**CONCLUSION**

Growing out of the larger LSTA grant initiative the Ohio University Historic Books Project gave us a unique opportunity to change course, adopt better practices, and at the same time identify previously unrecognized treasures. More consistent physical processing and more complete description will not only benefit our users and make our collections richer, but it will also help us adopt a less aggressive general collections preservation strategy that can match our existing resources. Large scale relocation and de-duplication projects must utilize data-driven analysis to make informed decisions with efficiency. Through the Ohio University Historic Books Project we were able to include targeted copy level assessment to verify one of the key goals of that analysis; the long-term retention and access of Ohio University Libraries distinct collections. Most important of all was the cross-departmental collaboration, which empowered us to advance our practices and more confidently balance collection decisions that must be made at scale.

**Acknowledgments**

The authors wish to acknowledge the substantial influence of Book Traces @ UVA a project of the University of Virginia Library with support from the Council on Library and Information Resources (CLIR) as well as Kristin Jensen, designer of the Book Traces @ UVA protocol. We also wish to acknowledge the significant contribution of Jeff Fulk, Sandy Gekosky, Patrick Koga, and Macie Penrod whose skilled work and dedication to the Ohio University Libraries made the project possible.

**References**

