One Step at a Time

Using Targeted Pilot Projects to Achieve Meaningful and Scalable Metadata Reparation

Savannah Lake, Joseph Nicholson, and Jenn Brosek

THE J. MURREY ATKINS LIBRARY SUPPORTS TEACHING AND RESEARCH AT the University of North Carolina at Charlotte, a public research university serving nearly 30,000 students, 41 percent of whom come from underrepresented or underserved backgrounds. The library's resources are accessible to users through the catalog, a digital collections repository, an institutional repository, and ArchivesSpace, all of which are tightly integrated, but use different metadata standards and search interfaces.

Evaluating the metadata in these systems through a diversity, equity, inclusion, and accessibility (DEIA) lens became a top priority when the library began preparing a ten-year strategic plan with diversity and inclusion as core values. In 2021, the Atkins Library assembled the DEIA Descriptive Practices Working Group to devise remediation strategies and recommendations for inclusive description in metadata workflows. The group's members included librarians and technical staff from across the library, including Special Collections and University Archives, Collection Services, Public Services, Digital Scholarship and Innovation, and Administration.

The group immediately confronted a serious dilemma: how to craft a metadata reparation strategy broad enough to address significant problems in several systems with different metadata needs and practices, all managed

by staff from different units. The problem seemed insurmountable until the group decided to forgo a single comprehensive approach in favor of smaller pilot projects focused on tightly defined metadata issues. Smaller-scale, iterative projects have proven to be a viable approach to remediation efforts, generating impactful results and learning opportunities while laying the groundwork for more far-reaching approaches in the future.

Recent years have seen a groundswell of inclusive metadata initiatives. A 2017 OCLC study found 70 percent of survey respondents planned to change metadata descriptions to align with DEIA goals.[1] Similarly, the Program for Cooperative Cataloging included diversity and inclusion as an "integral component" of its 2018–2021 Strategic Directions document.[2] Such work reflects a growing recognition that the standards by which metadata quality has been traditionally evaluated have largely neglected the ways in which an increasingly diverse group of users may not see themselves accurately represented in resource descriptions. This chapter seeks to offer further detail on the development and management of reparative projects themselves, demonstrating how institutions can approach wide-ranging metadata issues through smaller, targeted "pilot projects."

THE "PILOT PROJECT" FRAMEWORK

The pilot project framework began with the realization that the needed metadata remediation work was too vast and complex to be handled through a single approach. Additionally, group members were too inexperienced in reparative work and tools to know how to carry out large projects affecting millions of records in multiple systems, at least at first. Participants needed the opportunity to develop remediation skills before doing reparative work on a wider scale, under conditions allowing them to experiment, make mistakes, and work their way through metadata quandaries. Group leaders also understood the initial remediation efforts could turn into a protracted fishing expedition without a firm structure and time constraints.

Pilot projects—smaller initiatives focused on a specific metadata issue that could be remediated within a year or less—seemed the approach best calibrated to the strengths and limitations of the group's members. The working group planned to launch and complete the projects during the first year, assess the results, and then begin new, larger projects during the second year to build

on skills developed during initial remediation efforts. While not fully adhering to agile or scrum strategies, which have seen success in metadata remediation, the pilot projects similarly draw on the value of incremental work, for both creating manageable workloads and adjusting future projects based on previous results.

To begin formulating pilot projects, group members identified potential reparative issues by assessing records through simple keyword and subject searches. Once identified, the group used more rigorous tools, such as indication rules in Alma and XSLTs, to quantify the extent of the issues. Next, group members submitted pilot project proposals. After gathering feedback from other participants, group members were ready to launch their projects.

PILOT PROJECTS

The group's first year of work saw significant progress in several projects. Perhaps unsurprisingly, the projects making the most headway involved subject terminology. These projects benefited from batch remediation strategies as well as lists of problematic LCSH (Library of Congress Subject Headings) terms already compiled by the cataloging community. Ultimately, the first year of reparative work spanned the catalog, the institutional repository, and the digitized special collections, redressing offensive language describing multiple communities.

LCSH Remediation in Alma and Primo

In November 2021, the Library of Congress announced it would change the headings "Aliens" and "Illegal aliens" to "Noncitizens" and "Illegal immigration" and make parallel changes to a related cluster of headings describing immigrants and immigration, among them "Alien detention centers," "Children of illegal aliens," "Illegal alien children," and "Women illegal aliens." The failure of automated back-end authority control functionalities to update the invalid headings in all bibliographic records in the Atkins Library's catalog spurred two group members to propose a pilot project to perform the edits themselves. After identifying 843 records with the invalid headings by using Alma indication rules, the group members employed a normalization process in batches to update the records. A handful of records with headings the indication rules and normalization procedure failed to detect had to be updated manually.

Normalization Rule in Primo

Targeting LCSH "Illegal aliens" headings only, the LCSH pilot project did not succeed in detecting FAST (Faceted Application of Subject Terminology) equivalents for the invalid headings lingering in some records. Nor did it seek to provide more inclusive substitutes for the new Library of Congress term "Illegal immigration," which has been widely criticized as problematic. Additionally, this effort did not seek to implement non-LCSH substitutes for offensive LCSH terms. To address these shortcomings, group members launched a related pilot project using normalization rules in the Primo VE discovery layer to replace problematic headings with more inclusive terms, bypassing record-by-record changes in underlying MARC records.

This approach involved creating two Primo VE normalization rules: one to perform the term substitutions in the Primo VE discovery layer and another to enable users to retrieve records by performing catalog searches for both substituted terms and deprecated terms in the underlying MARC records. Working group members needed the assistance of Ex Libris staff to reindex the Alma database to force the changes in the discovery layer. To date, the approach seems to have been successful, though it does not affect electronic CDI (Central Discovery Index) records appearing in Primo VE independently of Alma. Despite this shortcoming, Primo VE normalization rules are the group's primary method for addressing problematic vocabulary terms in its catalog, at least for now.

FAST Remediation in Digital Repositories

Three group members led a pilot project to remediate problematic subject terms in the library's institutional repository and its repository of digitized special collections materials, which both run on Islandora. To carry out this work, group members first reviewed the literature and existing resources on offensive FAST terminology. Then group members conducted searches within the repositories to identify problematic terms in repository records, compiling a list with proposed changes to share with cataloging staff for feedback. Once everyone had agreed to the changes, group members updated the records as well as the project documentation. The project began with terms related to the LGBTQIA+ community and Indigenous peoples, and it will continue to evaluate terms related to other communities and identities.

While this project benefited greatly from the work already done by the wider cataloging community, it did face challenges. Unlike the library catalog,

Islandora does not allow for differences between the front-end display and back-end indexing. Consequently, updating offensive terms means removing them entirely from the search system. This can present a discovery problem to those researching these topics; however, the group felt the benefits of using respectful language outweighed this impact on searches, especially as users can always consult with library staff on search strategies.

Additionally, because LCSH are more widely used than FAST, some resources suggested replacements for offensive LCSH terms, but not their FAST equivalents. This occurred for terms with geographical subdivisions, for example. Ultimately, group members implemented localized solutions, thoroughly documenting such decisions so staff would understand their full context and would be able to make changes in the future as needed.

Future Projects

Currently, as the DEIA Descriptive Practices Working Group members wrap up the first year's projects, they began investigating and exploring new projects. These include assessing race-conscious editing of oral histories, centralizing documentation on the group's efforts, creating a formal process for responding to comments received through the form where library users can report offensive language, and launching additional subject remediation projects.

FINDINGS

Approaching large-scale metadata remediation through small-scale pilot projects yielded both tangible outcomes and manageable workloads for staff. Importantly, the pilot project framework ensured the working group meetings did not just become discussion groups, but generated real changes in the library's descriptive practices. Although incremental, the reparative results are substantive and form a foundation for continued success in future pilot projects. Additionally, these projects generated connections between staff from previously siloed parts of the library, helping facilitate a more comprehensive metadata approach across multiple platforms.

The time and skill limitations of group members curtailed the success of some projects. Working on a volunteer basis, group members had to balance the demands of their pilot projects with the duties of their regular positions. While several pilot projects were successful, it seems clear that

understanding remediation work as a piecemeal activity performed in between more pressing duties will achieve limited results. Broadly speaking, there was general enthusiastic and ideological support from the library, which included reparative work in its ten-year strategic plan. However, the library could not provide additional funding or positions for these projects. A broader transformation of the Atkins Library's metadata will require a more holistic approach in which remediation work is a fundamental component of technical services workflows.

CONCLUSION

Though limited in impact, the pilot project approach enabled Atkins staff to take important first steps in reparative metadata work. Providing a training ground for testing out remediation strategies on a small scale, the pilot projects helped group members identify scalable approaches for future projects and build the skills they will need to carry them out.

In the next phase of its work, the group plans to create a comprehensive strategy for metadata remediation at Atkins that makes use of some of the same incremental approaches developed during the pilot projects to achieve broader aims. Partnering with communities and stakeholders external to the library, such as vendors, users, and the larger descriptive community, will be integral to this strategy, as each group has a crucial role in changing the metadata landscape. Looking ahead, group members plan to transform the Atkins Library's piecemeal pilot projects into a durable and transformative reparative metadata program.

NOTES

1. OCLC Research, "Equity, Diversity, and Inclusion in the OCLC Research Library Partnership Survey," www.oclc.org/research/areas/community-catalysts/rlp-edi.html.

2. Matthew Haugen and Amber Billey, "Building a More Diverse and Inclusive Cataloging Cooperative," *Cataloging & Classification Quarterly* 58, no. 3-4 (2020): 382, https://doi.org/10.1080/01639374.2020.1717709.