DCPAPERS

Inclusive Metadata in Digital Libraries: A Qualitative Study of R1 Academic Institutions

Stephanie M. Luke[1,*,†] and Trevor Stratton[2,†]

[1] University of Illinois Urbana-Champaign, 1408 W. Gregory Dr. Urbana, Illinois, United States
[2] Michigan State University, 366 W. Circle Dr., East Lansing, Michigan, United States

Abstract

A qualitative, multi-institutional study investigating how academic libraries in the United States integrate inclusive and reparative description principles and practices into their digital collections. This paper draws on a series of interviews with metadata professionals from eight R1 institutions to assess how they approach the remediation of legacy metadata in their digital collections. The findings demonstrate that, while inclusive metadata is widely recognized as important, its implementation remains underutilized and unevenly applied.

Keywords

inclusive metadata, digital libraries, reparative description

1. Introduction

In recent years, information professionals have become increasingly interested in confronting the embedded biases of traditional cataloging and archival description. Librarians and archivists have become more aware that the metadata they use to describe objects is inherently imperfect [1]. Metadata represents the views and experiences of those creating and describing resources rather than those who are represented in the materials [2].

The movement to reevaluate how the profession approaches the description of resources first began with traditional cataloging [3]. Criticism of Library of Congress subject headings as inequitable and non-representative resulted in numerous librarians calling for a review and reimagination of how authorized terms are created [4, 5]. The work of reflective metadata practices soon emerged in the archival field as well [6]. The concept of reparative description, or the "remediation of practices or data that exclude, silence, harm, or mischaracterize marginalized people in the data created or used by archivists to identify or characterize archival resources," became increasingly important to archivists wanting to reevaluate their descriptive practices [7]. Archives as a field has begun to adopt a view of cultural competency and humility in its approach to producing metadata [8].

Only more recently have librarians begun discussing the role of inclusive metadata in digital collections [9]. Like library catalogs and archives, digital libraries offer their own challenges to description [10]. The use of the Dublin Core metadata schema in most digital libraries offers flexibility in metadata creation and application [11]. This flexibility, however, can prove challenging to producing consistent description that uses inclusive metadata

* Corresponding author.
† These authors contributed equally.
smluke2@illinois.edu (S. M. Luke); stratt84@msu.edu (T. Stratton)
0000-0002-5713-2915 (S. M. Luke); 0009-0003-0802-3696 (T. Stratton)

https://doi.org/10.23106/dcmi.952536765

principles [12]. These challenges are underscored in the case of aggregated digital collections due to the loss of local context within those environments [13]. As a result of the unique challenges presented by digital collections, maintaining accurate and inclusive metadata is of paramount importance.

In recognition of the significance of this work, the authors sought to evaluate the current state of inclusive metadata practices within digital libraries. They designed a series of interviews with nine institutions in the United States. They aimed to uncover whether inclusive metadata was being discussed in academic libraries and if these libraries were adopting inclusive description practices in their digital collections. The interviews focused on the specific inclusive metadata-related projects undertaken by each institution and the obstacles that stand in the way of making further changes.

2. Methods

This qualitative study is based on a series of semi-structured interviews with metadata professionals who work with digital library metadata at academic institutions classified as "R1: Doctoral universities - Very high research activity" by the Carnegie Classification of Institutions of Higher Education. To ensure that diverse perspectives are represented within this study, each of the nine United States Census Bureau designated divisions of the United States (New England, Middle Atlantic, East North Central, West North Central, South Atlantic, East South Central. West South Central, Mountain, and Pacific) are represented. To further account for the diversity of perspectives, an approximately equal number of public and private institutions (5 public and 4 private) are represented in the study (Table 1).

Table 1
Institutions interviewed in the study

Institution	Region	Division	Type
Institution 1	Northeast	New England	Private
Institution 2	Northeast	Middle Atlantic	Public
Institution 3	Midwest	East North Central	Private
Institution 4	Midwest	West North Central	Public
Institution 5	South	South Atlantic	Private
Institution 6	South	East South Central	Public
Institution 7	South	West North Central	Private
Institution 8	West	Mountain	Public
Institution 9	West	Pacific	Pubic

The interview phase of this study utilized purposive stratified sampling. The authors targeted metadata professionals to participate in this study by identifying a Metadata Librarian (or similar role) who regularly works directly with digital collections metadata at their institution and emailing them to request participation in this study. If individuals meeting those criteria could not be found, unit heads of metadata and/or cataloging departments were contacted via email to determine who would be the best to include in the study. If contact could not be established with a metadata professional from an institution after one week of waiting for a response, or if the metadata professionals were unwilling or unable to participate in this study, the authors proceeded by identifying and making contact at another institution that fulfills the same criteria.

https://doi.org/10.23106/dcmi.952536765

A standardized interview protocol was used to maintain consistency while allowing for flexibility in responses. The interview questions were designed to yield long-form responses to comprehensively capture the nuances from each respondent in the form of detailed qualitative data regarding how their institution approaches metadata remediation and changes in best practices for their digital collections (see Appendix). The authors conducted the interviews virtually to accommodate the geographic diversity and schedules of interviewees.

3. Findings and discussion

At the time this paper was written, the authors had conducted eight of the nine interviews, with the final interview scheduled. Despite the differences in region and type of institution (public or private), the interviews revealed that metadata professionals working with digital libraries are experiencing many of the same challenges in making their platform's metadata more inclusive. All but one of the institutions revealed that they primarily reuse existing metadata from MARC, DACS, and other records, while occasionally creating new or expanding existing descriptions for their digital library platforms. All of the institutions indicated that they often work with subject and description experts like catalogers and archivists and crosswalk metadata to Dublin Core standards. A major challenge that the institutions indicated was that there is a disconnect between their information systems and platforms. For example, an institution might digitize a book and reuse an existing catalog record to create a Dublin Core record on the digital library. However, if descriptions or subject headings are later altered in the catalog record, there is no automatic update to the corresponding metadata in the digital library. This siloing of systems limits the reach of metadata changes to one discovery platform. The interview participants indicated a need for more dynamic and interoperable changes in metadata across systems and platforms.

Only one of the institutions interviewed indicated that they had conducted any kind of audit on their digital library metadata with inclusion in mind. Most indicated that they did not have the time, staffing, or resources to undertake a large-scale audit. However, institutions seemed more likely to have taken on project-based work with inclusive metadata. There was an even split between institutions who indicated that they had worked on projects to improve the inclusivity of metadata, with four answering in the positive and four institutions responding in the negative. Of those institutions indicating that they had participated in projects, one was related to implementing LGBTQ+ alternative vocabularies, another involved changing wording centering on Japanese-American internment, and a third sought to raise the visibility of underrepresented communities. When asked what challenges they have faced when attempting to make metadata more inclusive, institutions noted a variety of issues. Four indicated that time was an issue, four indicated staffing or resources, and three indicated that systems or infrastructure were a primary barrier. This reveals that although metadata professionals may be working at different institutions with collections of varying sizes and coverage, these institutions are faced with similar challenges in making their digital library records more inclusive.

The authors asked the interview participants if their institution hosted a harmful language or content statement. Four answered in the positive and two in the negative. Another two indicated that they formerly hosted a statement but that the statement had been taken down. When asked why they did not host a statement or had taken their institution's statement

https://doi.org/10.23106/dcmi.952536765

down, the interviewees noted a number of issues, but primarily that institutions were fearful of curbing intellectual freedom and that there was confusion over who had the final authority to create and post the statement. The authors did not ask if the current political climate had affected these decisions, but this may be an area for future research.

When asked if they used metrics to track the impact of inclusive metadata practices on their digital libraries, all institutions indicated that they had not. This question elicited discussion about what these metrics might look like and if any such system is possible.

Finally, the authors asked participants if they had noticed whether the digital library or the traditional library catalog tended to more quickly adopt changes in inclusive metadata practices at their institution. Three indicated that the digital library is quicker to adapt, two indicated that the traditional library catalog is quicker to adapt, and three indicated that they have not noticed a significant difference in adoption time of inclusive metadata practices between digital libraries and traditional library catalogs at their institutions. The institutions that indicated that the digital library is typically quicker to adapt noted that the metadata schemas, workflows, and best practices used in digital libraries tend to be more flexible than their traditional catalog counterparts. Additionally, some interviewees mentioned that the number of items in digital libraries tends to be smaller than the number in traditional catalogs, which makes it easier to quickly make changes. Meanwhile, the institutions that indicated that traditional library catalogs are quicker to adopt changes tended to focus on differences between the systems used while noting that the infrastructure for the traditional catalog at their institution is more conducive to change than the system they use for their digital library. Additionally, the institutions that indicated that there is not a significant difference between the two largely highlighted how neither group has made significant progress on this front at their institution, so they both fall equally short in their abilities to adapt to changes.

4. Conclusion and future research

This series of interviews proved illustrative in capturing the state of inclusive metadata practices in digital libraries at R1 academic institutions. The authors found that there was significant diversity in their results. However, the most significant finding was that most institutions are not conducting large-scale audits for inclusive metadata in digital libraries. Instead, if any work on this front is being done, it is done in project-based or one-time changes to digital collection records. The authors also found that many of the most prevalent challenges to the adoption of inclusive practices in digital libraries relate to systems, infrastructure, and staffing.

Based upon the qualitative data that the authors collected, they have begun working to create a survey that will further illuminate the state of inclusive metadata practice in digital libraries. With this survey, the authors are expanding their sample size beyond R1 institutions in the United States. They will seek to generalize the results and assess broader trends by including institutions of varying sizes both in the United States and abroad. They are working to capture this additional data in a more quantifiable manner. They are designing the survey using Likert scales. They will note trends between institutions, highlighting differences between global regions and the nature of the institutions.

https://doi.org/10.23106/dcmi.952536765

References

[1] H. A. Olson. "The power to name: Representation in library catalogs." Signs: journal of women in culture and society 26.3 (2001): 639-668.

[2] S. M. Luke and K. Slover. "Reparative and Inclusive Metadata: UTA Libraries Reevaluates Its Practices." Archival Outlook (March/April 2021): 3, 16.

[3] S. Berman, Prejudices and Antipathies: A Tract on the LC Subject Heads Concerning People, Scarecrow Press, Metuchen, NJ, 1971.

[4] S. A. Knowlton, "Three Decades since Prejudices and Antipathies: A Study of Changes in the Library of Congress Subject Headings." Cataloging and Classification Quarterly 40.2 (2005): 123-145.

[5] E. Drabinski. "Queering the Catalog: Queer Theory and the Politics of Correction." The Library Quarterly: Information, Community, Policy 83.2 (2013): 94–111.

[6] M. Caswell and M. Cifor. "From Human Rights to Feminist Ethics: Radical Empathy in the Archives." Archivaria 81 (2016): 23-43.

[7] Society of American Archivists, Dictionary of Archives Terminology, 2025. URL: https://dictionary.archivists.org.

[8] J. Tai. "Cultural humility as a framework for anti-oppressive archival description." Journal of Critical Library and Information Studies 3.2 (2021): 1-23.

[9] A. Abney, A. Boczar, and S. Jordan. "Conscious Editing-Driven Metadata for Archives and Digital Collections: A Case Study." The Journal of Academic Librarianship, 50.5 (2024): 1-8.

[10] S. Werling and E. Radio. "Comparing Critical Cataloging Procedures and Developing Local Policies in the Context of Digital Library Metadata." Journal of Library Metadata 25.1(2025): 33–55.

[11] J. Weagley, E. Gelches and J. R. Park. "Interoperability and Metadata Quality in Digital Video Repositories: A Study of Dublin Core." Journal of Library Metadata 10.1 (2010): 37–57.

[12] S. M. Luke, S. Pezzoni and W. Russell. "Towards More Equitable, Diverse, and Inclusive Representation in Metadata and Digitization: A Case Study." The Serials Librarian 82.1-4 (2022): 55–62.

[13] S. L. Shreeves, E. M. Knutson, B. Stvilia, C. L. Palmer, M. B. Twidale, and T. W. Cole. "Is 'Quality' Metadata 'Shareable' Metadata? The Implications of Local Metadata Practices for Federated Collections. Association of College and Research Libraries (ACRL) Conference 2005, Minneapolis, Minnesota. URL: https://alair.ala.org/server/api/core/bitstreams/95509f44-3804-4b06-a71e-a49083c7aeb1/content.

A. Appendix

The authors asked the interview participants the following questions.

1. Can you characterize your digital collections for us?
2. Can you describe how you approach metadata creation for your digital collections? Are you migrating metadata for your digital collections from your catalog records and other legacy descriptions or are you creating metadata specifically for your digital objects?

https://doi.org/10.23106/dcmi.952536765

3. Has your institution undertaken any projects to make your digital library metadata more inclusive? If so, can you describe these? What are the challenges that you have experienced in working toward making your digital library metadata more inclusive?
4. Have you implemented a harmful language or content statement? Can you describe the reasons behind your institution's decision to implement or not implement?
5. Have you undertaken any audits of your digital library for inclusive metadata? If so, can you describe these?
6. Does your library use any metrics to track the impact of inclusive metadata practices in your digital library? If so, can you describe any outcomes or trends that you have discovered?
7. Have you noticed any differences between how digital libraries and traditional catalogs have adopted inclusive metadata practices at your institution? Do you feel like either has been more responsive in implementing these practices?
8. Is there anything you want to tell us about this topic that we haven't asked you yet?

https://doi.org/10.23106/dcmi.952536765