

# Encoding Reparative Description

## Promoting Archival Metadata Revision through Data Visualization

DCMI 2024 (DublinCore Metadata Initiative, October 2024 in Toronto, Ontario)

Jesse A. Johnston\*  
Max Eckard+  
Gideon Goodrich+

\* University of Michigan School of Information  
+ Bentley Historical Library

**M** SCHOOL OF INFORMATION  
UNIVERSITY OF MICHIGAN

**M** BENTLEY HISTORICAL LIBRARY  
UNIVERSITY OF MICHIGAN

### Background

The past few years have seen increasing efforts in **reparative metadata** initiatives in libraries and archives. These projects can address past inequities, erasure, or incorrect representations of resources. While necessarily grounded in slow, human, relationship-based work, reparative metadata projects can benefit from computational approaches. Our project demonstrates how we have used automation tools to process metadata from archival finding aids to develop reports and visualizations that may aid in identifying issues in the metadata and addressing them through reparative, ethical, and community-sensitive ways.

**Archival metadata** is usually hierarchically structured, imitating the series arrangement of many collections, and frequently includes **narrative and semi-structured text**. Archival best practice in the US follows content and transcription rules of *Describing Archives: A Content Standard* (DACS) and according to the structure and tags defined by *Encoded Archival Description* (EAD), serialized in XML.

Given the importance of context and relationship in reparative archival description projects, and the large amount of text-based metadata in finding aids, our project developed computational tools that use the EAD structure to analyze and inform reparative description efforts at the Bentley Historical Library (Bentley). We report on these tools, our work to make them widely reusable, and how they may be used to expand reparative description work.

### Tools: ArchivesSpace & Python

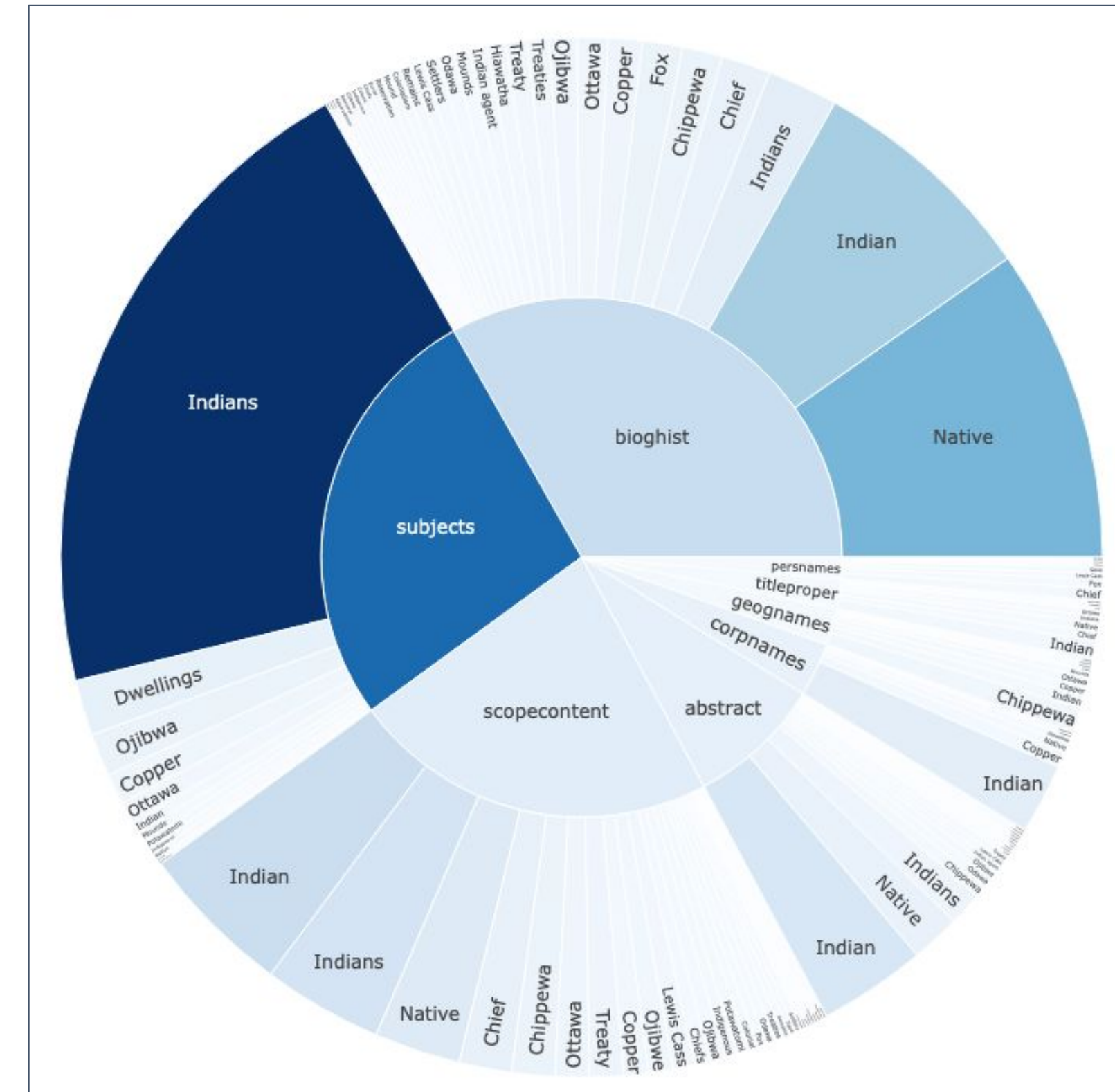
- ArchivesSpace (ASpace is an open-source finding aid management system.
  - ASpace 1.0 released in 2013, financially supported by The Mellon Foundation
  - Maintained by Lyzasis and a community of more than 450 archival organizations
  - Offers programmatic data access through an API and JSON-based serializations
- Code tools were developed using Python in Jupyter notebooks, ASpace JSON outputs
  - Inputs include term lists in text files and ASpace JSON
  - Python libraries include requests (API interaction via HTTP), pandas (data manipulation and analysis), and plotly (visualization)
  - Available on GitHub (scan QR code):



### Research Questions

- How can aggregated metadata reports and visualizations support the process of metadata repair and remediation?
- What reports and visualizations would be useful for archivists to identify, analyze, and take reparative actions to remediate metadata issues in a collection of finding aids (archival metadata)?

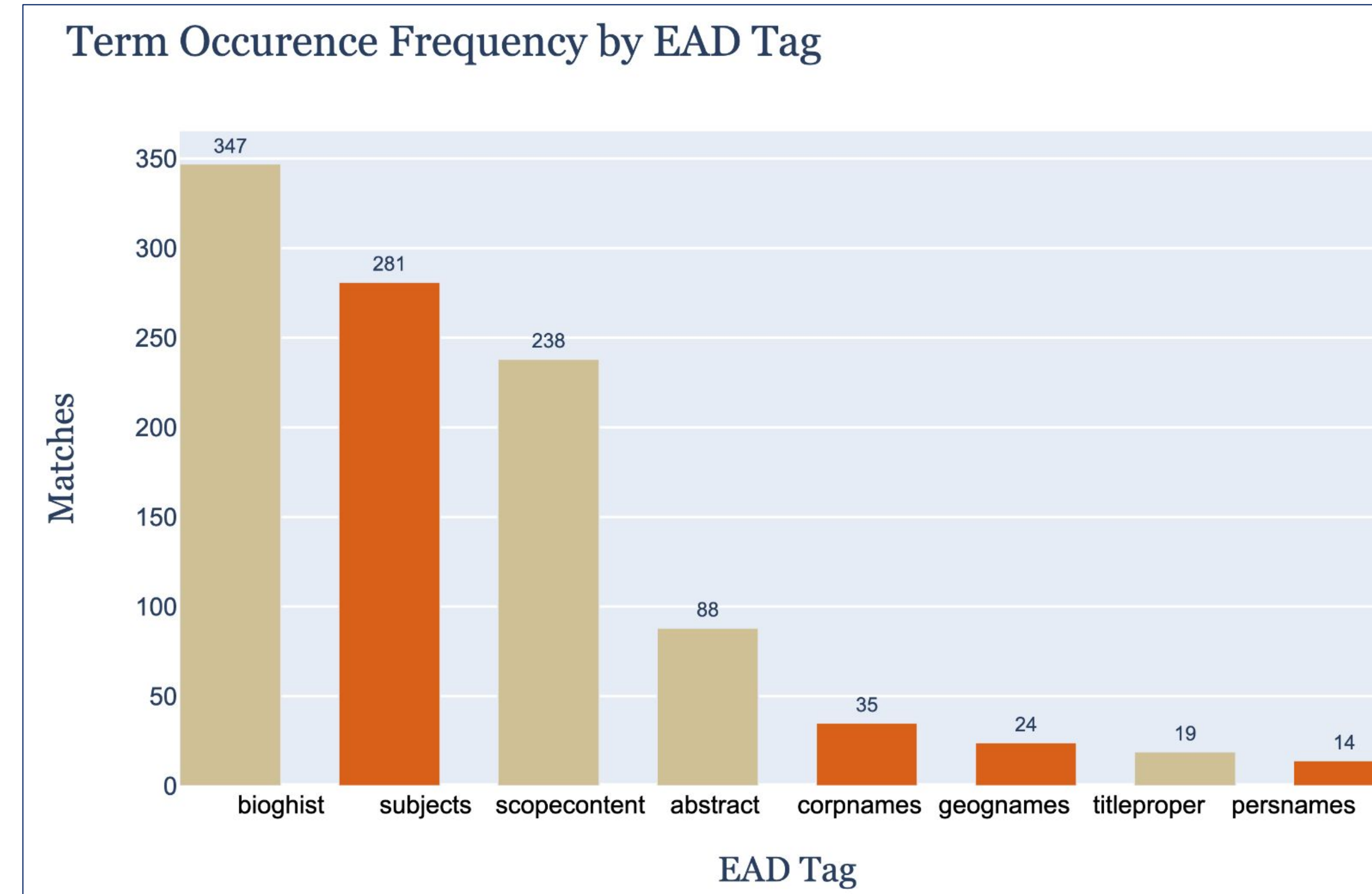
### Selected Visualizations



**Above:** Circular "sunburst" visualization. The central ring indicates the portion of total term matches by EAD metadata tag, while the outer ring indicates the proportion of term matches. Darker colors indicate higher term count (the darkest indicates 216 instances of the noted term in subject tags).

### Data

- 3,601 metadata records for Bentley finding aids
- EAD-tagged but gathered via the ASpace API in JSON
- parsed according to EAD tags and regular expressions to match specific character strings
- Matched according to a term list from Bentley's inclusive and reparative metadata team
- Outputs include graphic visualizations sampled here and key-word-in-context reports (below)



**Above:** Bar chart displaying term matches by EAD tag. The colors illustrate which terms are locally controlled (beige) vs metadata fields (orange) that use external vocabulary encoding schemes. Useful for illustrating different reparative pathways.

### Developing Use Cases & Workflow

The Bentley's Reparative and Inclusive Description Committee developed a new terms list and list of collections related to their Native American/First Nations materials. Terms were selected following consultation with Eric Hemenway (Little Traverse Bay Bands of Ottawa Indians), as well as analysis of current reparative vocabularies projects (like the University of Manitoba's "Changes to Library of Congress Subject Headings Related to Indigenous Peoples"). The resulting list included harmful words, phrases, and euphemisms, as well as terms that may indicate information regarding burials or culturally sensitive materials. Using this list, the code allows us to flag metadata for reparative review based on terms in context and prioritize reparative work, rather than make changes on an ad-hoc basis.

Term	Occurrence (evid_ID)	Tag	Collection	Context
Anishnaabe	umich-bh-2018025	scopecontent	the John P. Murphy collection	Researchers should note that the collection includes materials depicting Anishnaabe peoples, families, homes, and culture from a white colonial lens.
Burial	umich-bh-08811	bioghist	Joseph Best letters papers	On the Island of Mexico he executed three prisoners' burial rounds. He three himself' too, into the practice of anthropology, excavating burial mounds and ceremonial constructions.
Burial	umich-bh-93389	bioghist	Amos S. Owen papers	The location is a sacred site for the Anishnaabe, a place for burials and worship, which was seen as a curiosity and tourist attraction by white visitors.
Burials	umich-bh-2018025	bioghist	the John P. Murphy collection	Archaeology; (Monastic Archaeology; (Monastic Burials)
Burials	umich-bh-0381	subject	George R. Fox (1860-1882) Papers	By 1882 it had become clear to Winster that NARS could no longer remain a captive of OIA. If it was to be an effective organization.
Captive	umich-bh-68485	bioghist	Robert M. Warner Papers, 1940-2007	O'role and the Agreements files regarding emergency managers in various Michigan cities. Chief of Staff materials on public health policy, and the Legal Division materials related to the state's interactions with Michigan's Native American tribes.
Chief	umich-bh-2020007	scopecontent	the Rick Snyder papers.	Although chief executive during the deepest economic crisis experienced in Michigan since the Depression, Governor Miller was generally popular among both the people of the state and his political peers.
Chief	umich-bh-8924	bioghist	William G. Milliken papers	Notes, and delegations of Shaw and Sank Fox (Warriors) from among the Mackinac flow, including the Shaw Fox chief's book.
Chief	umich-bh-90936	scopecontent	the Cheung Best minute book	He also served as special assistant attorney general for the state of Michigan and chief appellate defender for the city of Detroit (1968-1969).
Chippewa	umich-bh-2014100	bioghist	Carl M. Levin papers, 1988-2015	Sean Carter and Black Bear Bands of Chippewa Indians.
Chippewa	umich-bh-2014100	bioghist	the Charles Bennett papers	Part Indian trader at Marquette, Michigan, and interpreter with the Chippewa Indians of Benzie County, Michigan.
Chippewa	umich-bh-98126	bioghist	the Theodore R. Bell papers	the Indians of Michigan.
Indians	umich-bh-82850	scopecontent	the The Indians of Michigan	the Indians of Michigan.
Indians	umich-bh-82850	scopecontent	the The Indians of Michigan	Brief history with comments on activities of his parents who served as Indians of North America; Indians of North America; Missionaries.
Indians	umich-bh-82850	subject	the The Indians of Michigan	the The Indians of Michigan.
Indians	umich-bh-851201	scopecontent	the Peter Dougherty papers	Reports of missionary activities, 1821-1822, of Rev. Peter; missionary teacher to Indians at Ottawa, and Bear Creek (now Pelee, Emmet County), Michigan, also Ottawa, 1835-1842 and May 1850, recording trips from New York City to Detroit, Michigan; located and through via the Erie Canal and the Great Lakes, various descriptions of Ojibwa, Mackinac Island, and other towns and places, in Michigan, and of Indian customs and his missionary activities.

**Above:** color-coded spreadsheet supporting reparative metadata work, based on a CSV keyword in context report.



**Left:** visual representation of reparative workflows currently in development at the Bentley.

### Project Activities & Methods

- Focused on advancing reparative description for archival metadata
- Rooted in an ongoing project, "ReConnect/ReCollect" (RCRC), led by PIs Ricardo Punzalan and Deirdre de la Cruz researching Philippine collections at the U-M
  - From 2021 to 2023, RCRC researchers identified harmful terminology that may be found in archival metadata, specifically finding aids.
  - Ella Li developed Python code to query EADs in XML; this code identified, counted, and recorded the location of the identified terms.
- The present project redeveloped the code to query ASpace's API and parse EAD metadata in JSON.

Following code redevelopment, we have worked with the reparative description team at the Bentley to gather feedback and identify useful reports and visualizations that can usefully support their reparative metadata work.

### Future Directions

We see this work as critical to move reparative metadata analysis beyond "find and replace" and the use of term lists at large. These approaches do not recognize the importance of dialogue, nor the varied relationships and context-dependent language that must be considered in reparative description work. Nonetheless, the complexity of archival metadata that should be addressed can benefit from tools that support reparative metadata work at a wider scale. We envision this work continuing in a few different directions:

- Testing and refining the existing code, including developing a command line tool with more flexible inputs and outputs and/or an ASpace plugin that supports live analysis and data updating
- Developing a methodology for triaging finding aids that need review based on number of "hits"
- Developing a utility that makes it easier for archivists to identify terms related to those in the list, potentially as part of future plugin
- Differentiating analysis based on modes of use (e.g., analysis of term usage, vs term 'correction')
- Developing reparative pathways for locally controlled or narrative fields, vs fields that use external vocabularies
- Getting feedback from community members on visualizations and their utility for illustrating how often and where terms occur

The University of Michigan School of Information (UMSI), one of the first iSchools, delivers innovative, elegant and ethical solutions connecting people, information and technology.

The Bentley Historical Library collects the materials for and promotes the study of the histories of two great, intertwined institutions, the State of Michigan and the University of Michigan.

The Bentley Historical Library and University of Michigan School of Information acknowledge that coerced cessions of land by the Anishnaabeg and Wyandot made the University of Michigan possible, and we seek to reaffirm the ancestral and contemporary ties of these peoples to the lands where the University now stands.