

Challenging legacies at the British Library

Alan Danskin

The British Library established a corporate Anti-Racism Project (2020) designed to encourage participation via six subgroups, with staff recommendations incorporated into “Enacting Change”, the Library’s Race Equality Action Plan (2022). The research and recommendations of the Cataloguing and Metadata subgroup fed into a pilot project proposed as a proof of concept to develop end-to-end workflows, focused on the South Asian Collection and the Caribbean Collection. Work has already been done to align ethical approaches to cataloguing with the Library’s values. During the pilot project, new training modules for an ethical approach to metadata are being developed. The outcome will be to remove barriers to access and embed sustainable best practices.

Introduction

This article is based on the presentation given on 13th May 2022 as part of the ARLIS UK & Ireland Cataloguing & Classification ethics series. It has been substantially updated.

In July 2020, the British Library established a corporate Anti-Racism Project to, ‘make recommendations both in terms of immediate actions and longer-term proposals that we will integrate into our strategy and our culture, to make us a truly anti-racist organisation, in a wide-ranging and sustainable way’. (Liz Jolly, Chief Librarian)¹

The project was staff led and designed to encourage participation by staff, irrespective of grade, role or location to look at all aspects of the Library’s operations, services, policies and procedures. Six subgroups were established to review and make recommendations with respect to:

- Audiences
- Behaviour, Values and Experiences
- Cataloguing and Metadata
- Collections and Content
- Data Research and Insights
- People and Human Resources Policy

Project management was deliberately light touch. A seventh leaders and coordinators group facilitated communication and coordination between the subgroups and the overall Project Manager, who reported to steering group.

These subgroups began work in late 2020/early 2021. The work was conducted mostly via MS Teams as many staff were working from home and included colleagues based in London and Yorkshire.

The subgroups were charged with making recommendations by the end of September 2021. The recommendations were incorporated into, *Enacting Change*, the Library’s Race Equality Action Plan (REAP), published in January 2022.

Enacting change made a commitment that the Library’s senior management team would ‘take ownership of the Subgroups’ recommendations and deliver real change.’

1. ‘British Library commits to being an anti-racist organisation’ press release, <https://www.bl.uk/press-releases/2020/july/british-library-commits-to-becoming-an-anti-racist-organisation>, [accessed 13/02/2023]

Cataloguing and Metadata Subgroup (CAM)

Much of what I am going to cover derives from the work of the Cataloguing and Metadata subgroup. The research they conducted and the ideas they generated have profoundly influenced the REAP.

The members of the subgroup defined its scope and purpose as follows:

- To understand and expose the extent and impact of discrimination and bias in the Library's cataloguing practices and metadata
- To facilitate the development and enhancement of descriptions, interpretations, and presentations of collection items
- To collaborate with colleagues involved in these areas of the Library, and to learn from best practice beyond the Library, to make recommendations to enhance the safety, inclusivity, and accessibility of the Library's collections for all potential users

To this end, members undertook desk research, reviewing literature and actions taken by other institutions to make their metadata more inclusive. Some of this work was undertaken in collaboration with the Audience and Collection and Curation Subgroups.

Broadly speaking, they identified three broad areas for action : metadata, access and discovery, policy and principles.

The group made ten specific recommendations which were accepted by the steering group and are included in the REAP. The group also emphasised the dependency of effective outcomes on resources by recommending, "the creation of new projects and posts" for cataloguing and metadata and allocation of "resource to Technology to support the infrastructure."

Cataloguing and Metadata Recommendations:

1. Establish a comprehensive glossary of problematic and inclusive terminology as a reference for the cataloguing of culturally sensitive material
2. Establish a multifaceted approach to auditing the collections for racist descriptions through resourced projects and business as usual work in cataloguing and curation teams
3. Create a more accessible and transparent feedback process to encourage users to report harmful material
4. Develop processes to enable continuous collaboration with stakeholder communities, while respecting our partners' time and labour
5. Establish a flexible schema for treating problematic terminology that considers terms in context, while not inhibiting the accurate identification of collection items
6. Implement guidance by which users are made aware of terms or images in the Library's content or metadata that may cause offense or distress.
7. Develop more flexible subject approaches
8. Develop the means to search the collections by language, especially lesser-known languages, and to search in original scripts
9. Work with Indigenous communities and experts to establish appropriate access and description protocols for material of relevance to Indigenous cultural heritage
10. Establish ethical principles for cataloguing and metadata from which to update and maintain policies and procedures

There is common ground between the CAM recommendations and recommendations of the Collections and Curation Subgroup to justify a collaborative approach for delivery. We agreed to work together on an end-to-end pilot project.

Pilot Project

It was clear that an audit of the collection (170 million items and counting) and catalogues of the British Library would not be achievable within the three-year scope of the REAP. Instead, a pilot project with a more limited scope was proposed as a proof of concept to develop end-to-end workflows in preparation for transition to business-as-usual within 3 years.

The South Asian Collection and the Caribbean Collection were selected for the pilot project. These collections contain diverse material types, including published and unpublished resources, texts, images and sound recordings. The collections are relevant to communities in the UK and are of international significance. They are sufficiently large for impact of the pilot to be discernible to

staff and users. The scope and diversity of the collections will provide a valuable test for development and validation of policy and procedures.

We have obtained project funding for additional (3 year) fixed term posts: Metadata Lead for Equity and Inclusion; Metadata Coordinator for Equity and Inclusion and 1.6 additional curatorial posts. These dedicated roles will be supplemented by permanent staff, including metadata analysts, cataloguers and curators, where necessary.

The project plan is still being finalised, but the first substantive piece of work will be an audit of the pilot collections, to identify terminology or content that may be offensive, discriminatory or harmful. The audit will also contribute to identification of collection items for more detailed provenance research.

The *Inclusive terminology: guide & glossary for the cultural heritage sector* May 2021, developed by Carissa Chew, while at the National Library of Scotland, will be a valuable tool for identification of catalogue records that contain problematic terminology. The glossary is particularly strong in relation to African slavery and the American transatlantic slave trade. We hope that our work on the Caribbean and South Asian Collection will enable us to develop and expand its scope. We are in discussion with colleagues at the National Library of Scotland regarding the sustainability and future development of this valuable resource.

The audit will enable us to identify catalogue records and collection items that contain terminology or content that may be offensive, discriminatory, or even harmful. A fundamental principle is to preserve the content of items in the collection for future researchers. We cannot change it, but we can and should advise users about the content that they may find distressing or offensive. This could include images or sound recordings, as well as texts, anywhere in the collection. Therefore, it is appropriate to provide general content guidance on the website and in the catalogue. Such advice is commonplace in television and other media and is already widely practiced by other cultural heritage institutions.

Metadata belongs to the Library so, in principle, we can amend any problematic terminology. In practice, catalogue records include information transcribed from the item for discovery and identification. Redacting words in the title of an item may prevent it from being returned by a search query or being identified in a results set. The intentions of the original author and the needs of the user can be respected by separating indexing from display. For example, by using a redacted form for display in the user interface and the un-redacted form for indexing. This is also commonplace in newspapers, film, television, etc.

Most metadata is not transcribed, it has been added by cataloguers employed by the Library or its precursor institutions, dating back to the mid-19th century. There is no doubt that offensive, discriminatory or distressing terminology assigned by a cataloguer should be amended. However, there is an argument for retaining previous versions to preserve the scholarly record. This is not always technically or practically possible in our current systems. The pilot project will explore different options to balance scholarship and efficiency. For example, if the descriptions originated in a printed catalogue there may be no need to retain different versions of the machine readable record for scholars.

Metadata is also acquired from other sources, including metadata that the Library purchases from aggregators and publishers; metadata that we derive or copy from other institutions; metadata created by crowd sourcing; and metadata derived from controlled vocabularies, such as Library of Congress Subject Headings (LCSH), DDC, MARC 21 or FAST (Faceted Application of Subject Terminology). Some of these issues will be addressed through contractual agreements with suppliers, which are out of scope for the pilot project, others may influence copy cataloguing policies and procedures.

The pilot will have to evaluate different approaches for addressing terminology imported from controlled vocabularies. This can include proposing new or alternative terms and identifying more granular or specialised vocabularies to supplement LCSH or FAST, but the challenge is to devise practical means by which problematic terms can be automatically suppressed or mapped to alternative terms in display. The Library's decision to switch from LCSH to FAST was driven in part by the flexibility of FAST's faceted structure and the potential of linked data as a bridge between different vocabularies.

We also aim to make the catalogue more inclusive by improving access in original scripts and by building on the success of the *Languid* project² to explicitly identify endangered and minority languages.

2. Victoria Morris, "Automated Language Identification of Bibliographic Resources", *Cataloguing & Classification Quarterly*, Vol. 58, issue 1, 2020, <https://doi.org/10.1080/01639374.2019.1700201>

In parallel with the planning for the REAP, the internal Content Guidance Working Group (CGWG) has been reviewing examples of content and metadata on the Library's Website to develop a principled approach to content guidance. At present such decisions are largely at the discretion of projects, individual curators, or teams. The CGWG principles and tools will be applied in the pilot project and will feed into training and business-as-usual procedures.

We aim to involve communities in the work of the pilot project. These will include internal staff networks, but we also want to engage with the wider community to identify priorities and obtain feedback. We will encourage users to help us remove barriers to access by reporting content or metadata that is discriminatory or distressing.

Benchmarking

OCLC's research report *Reimagine descriptive workflows: a community informed agenda for reparative and inclusive descriptive practice*, by Rachel L. Frick and Merrilee Proffitt³ was used to benchmark the scope of our planned activity. The report proposes three main areas for change.

Organizational shifts

Allocating resources and changing priorities. It also means a commitment to amend past practices and to sustain these changes over the long term. All of this is inherent in the REAP and explicit in our planning and resourcing for the pilot.

Operational workflows

The outcomes will change what is done and the way it is done. The emphasis is on making the Library a welcoming place to all of our users and to be open to and encouraging of community engagement.

Professional and personal development

Work has already been done to align ethical approaches to cataloguing with the Library's values. During the pilot project, we will develop new training modules for an ethical approach to metadata.

Conclusion

The title of this article reflects the scale of the challenge facing the Library to identify and address discrimination and inequity in the collections and catalogues, dating back over one hundred and fifty years. It is also an acknowledgment that if we are to serve our diverse audiences, now and in the future, we must constantly challenge our assumptions.

The purpose of our Library Catalogue is to make public collections accessible to those who want to use them for research, inspiration, or enjoyment. The outcome of this project will remove barriers to access to our Caribbean and South Asian Collections and will embed sustainable best practices to extend the work to the collection as a whole.

Alan Danskin
Collection Metadata Standards Manager
 British Library
 Boston Spa
 Wetherby
 West Yorkshire
 LS23 7BQ
 UK
 Email: alan.danskin@bl.uk

3. Rachel L. Frick and Merrilee Proffitt, *Reimagine descriptive workflows: a community informed agenda for reparative and inclusive descriptive practice*, <https://www.oclc.org/content/dam/research/publications/2022/oclcresearch-reimagine-descriptive-workflows-a4.pdf>, [accessed 14/03/2023]

References

- Brown, Hugh and Rajakumar, T. *Enacting change: rolling out our Race Equality Action Plan*. British Library Living Knowledge blog, 26th January 2022 <https://blogs.bl.uk/living-knowledge/2022/01/enacting-change-rolling-out-our-race-equality-action-plan.html>
- Brown, Hugh and Rajakumar, T. (2022) *Enacting Change; an action plan for race equality at the British Library*. PowerPoint Presentation. 2022 <https://blogs.bl.uk/files/bl-race-equality-action-plan-jan-2022.pdf>
- Chew, Carissa (2022) *Inclusive Terminology Project: information and resource pack*, https://docs.google.com/document/d/17W4hstjFFIWRdD5gWjx0WDKuqzUHMqat9zylAjqW_X8/edit
- Morris, Victoria (2020) Automated Language Identification of Bibliographic Resources, *Cataloging & Classification Quarterly*, **58**:1,1-27, CCQ DOI:[10.1080/01639374.2019.1700201](https://doi.org/10.1080/01639374.2019.1700201) <https://bl.iro.bl.uk/concern/articles/6c99ffcb-0003-477d-8a58-64cf8c45ecf5>
- Frick, Rachel L., and Merrilee Proffitt. (2022). *Reimagine Descriptive Workflows: A Community-informed Agenda for Reparative and Inclusive Descriptive Practice*. Dublin, OH: OCLC Research. <https://doi.org/10.25333/wd4b-bs51>.