

Problematic subject headings

Making our catalogue more equitable, diverse and inclusive at the University of Bristol

Liz Cooper

Discovery Librarian, Library Services, University of Bristol

Damien McManus



Subject Librarian, Library Services, University of Bristol

Received: 13 Feb 2025 | Published: 17 Mar 2025

ABSTRACT

Subject headings have increasingly been the focus of attention from students, researchers and librarians, particularly in relation to headings which causes distress and offence. By using the Locally Preferred Subjects feature of OCLC's WorldShare Management Services library services platform currently in use at the University of Bristol, we were able to plan and implement a procedure to analyse and replace where necessary harmful subject headings. We describe the workflows and processes associated with this project, focusing on the technical and organisational challenges that we encountered and the solutions we found, and how we will maintain and progress the project in future.

KEYWORDS decolonisation; subject headings; diversity; inclusion

CONTACT Damien McManus  damien.mcmanus@bristol.ac.uk  Library Services, University of Bristol

Introduction - landscape at the University of Bristol

The University of Bristol is a large university in the Southwest of England. As an institution, the university has been putting Equality, Diversity and Inclusion at the top of its agenda for some years. In the Library, working groups have arisen in the past decade to reflect this: for example, the Decolonising Working Group and the Inclusive and Accessible Library Forum.

WorldShare Management Services/WorldCat

Our library services platform is the cloud-based WorldShare Management Services (WMS) from OCLC. WMS is founded on WorldCat, a union catalogue and database, with thousands of subscribing and member libraries contributing their metadata around the world. Our discovery layer, Library Search, has a local layer as well as a worldwide layer. Our users can conduct a search within our local holdings or expand that search to the worldwide layer (WorldCat.org).

WorldShare Locally Preferred Subjects feature

In the Autumn of 2022, OCLC introduced a new functionality to WMS, the Locally Preferred Subjects Feature. This tool arose from a community research project, the Reimagine Descriptive Workflows Project, conducted by OCLC to look at cataloguing practices and workflows across its subscribing libraries. The outcomes of this project led to OCLC's overarching Diversity, Equity and Inclusion (DEI) initiative. One strand of this initiative was to enable libraries to replace harmful and offensive subject terms in the discovery layer at a local level. Users would still be able to conduct searches using their preferred language and terms would only be replaced in the local discovery layer, not globally or within MARC records. OCLC recognised that one size did not fit all: decisions around which terms were deemed harmful or offensive and how these terms should be replaced needed to be made at individual institution level.

In October 2022, we decided to form a cross-team Subject Remapping working group within the library to look at ways to utilise this new tool within our discovery layer. We wanted to develop local practices and workflows, taking technical challenges into consideration and exploring ways in which we could make informed choices about local subject replacements.

The Locally Preferred Subjects Feature has undergone several iterations and has gradually become a more sophisticated tool.

How we decide on subject areas

The libraries at the University of Bristol include extensive printed collections, classified using the Library of Congress classification scheme (LoC), and electronic books and journals. The collection is too large to analyse as a whole, so we agreed that we would start by looking at the subject headings in the collection in our largest library, the Arts and Social Sciences Library (ASSL). This collection represents around 40% of our print holdings. The process could be expanded later to look for related and additional headings in collections in our other libraries.

Our first focus was therefore a relatively small LoC range: HT851 to HT1445 that has enslavement as its focus. We then identified several other targeted areas:

- HV1551-HV3024 (entitled in the LoC schedules "People with disabilities")
- Sequences in the R range, for instance RA418-RA790 (broadly "Public aspects of Medicine")

These were chosen to enable us to identify problematic headings and the suggestion of alternatives.

Identification of alternative terms

Before starting the analysis of subject ranges, we gathered policies, lists of preferred terms and position statements from library, archive, and museum websites. Preferred term lists are especially useful for analysis of some terms, particularly those relating to enslavement. Fortunately, there is a wealth of valuable resources freely available, for instance Nottingham City Council's glossary of terms relating to Transatlantic enslavement ([Nottingham City Council, no date](#)). Some terms relating to subjects such as disability and medicine are more contested. While they are seen as problematic to some, others see them as being acceptable. Over time meanings and contexts change. This can lead to some keywords causing offence or distress. For discussions on this see, among others, [Kenny et al., \(2016\)](#) and [Cho et al., \(2023\)](#). We decided not to change a small number of terms which could be seen to be harmful as we could not produce preferable alternatives. An example is 'Geriatric', pejorative in some contexts but also a branch of medicine. Where problematic terms were identified, we included sources for the preferred terms, logged on a spreadsheet for future reference.

How we extract the data

Metadata specialists in our Collections and Metadata teams designed a workflow for extracting subject headings from our catalogue. Our Collections Librarian is particularly adept at designing ways to extract data from multiple systems and distilling data into a workable format. He designed a workflow for extracting subject headings that could be broken into the following tasks:

- Retrieving the catalogue records for the books in the relevant subject area (in Excel lists)
- Extracting the bibliographic data from the catalogue records (in MarcEdit)
- Extracting the main subject headings from the bibliographic data in the catalogue records (in MarcEdit and Excel)
- Sorting the subject headings for analysis (in Excel)
- Presenting the data for use by the wider team (in Excel)

The Collections Team regularly ran reports to extract a list of records of holdings in ASSL into Excel lists, so this metadata was always up-to-date and readily available to work on. Excel lists included basic bibliographic data (author/title), Library of Congress class marks and system numbers. A range of agreed class marks within the spreadsheet could be selected from the classification column, the OCLC numbers in the rows from that range (system numbers for MARC records in WorldCat) could then be highlighted, copied and then pasted into notepad to be worked on.

In MarcEdit, we could use the Batch Search option in the Z59.50/SRU Client to search for the OCLC numbers in our list and retrieve the bibliographic records from our instance of WorldCat. Within MarcEdit, we could then use MarcBreaker to convert the retrieved records from (.mrc) into a format that would be easier to work on (.mrk).

From the latter file, we extracted the 650, 651 and 653 subject fields and copied them to a clipboard in MarcEdit to paste into a new workbook in Excel.

The raw data in the Excel workbook could then be refined (removing pipes and dollar signs, duplicates etc.) and the recurrence of certain terms across the specific subject area revealed. In this way, a spreadsheet containing headings and subheadings from the following schema was created: LCSH, LoC, FAST, MeSH and ILOT. We chose these as they are the main English language schema. This resulting spreadsheet of controlled subject headings could then be shared with the wider working group.

How we identified problematic subject terms and found alternatives (mapping, sources)

After harmful headings had been identified within the spreadsheet, we searched for them in the list of all subject headings in our records in ASSL (in Excel), found identical ones present in other parts of the whole library print collection, and replaced those at the same time.

How we uploaded new terms (upload spreadsheet in Config)

When the WorldShare Locally Preferred Subjects Feature was initially rolled out, OCLC envisioned all institutions utilising one universal template (Google sheet) to share suggested subject terms. In practice, institutions tend to download and localise the master Google sheet to reflect their locally preferred subject features. Some suggestions made by other institutions might be adopted universally, but equally, one size does not fit all. Many OCLC institutions are based in the United States. A preferred subject term in the United States could have a different nuance in the United Kingdom or elsewhere, or might not be a term that is recognised, and vice versa. Even within the same country, different institutions might choose different preferred subject terms. Furthermore, the maximum number of lines available in the master template is 2000. If all institutions used the same template, preferred subject terms would likely exceed 2000 lines very quickly.

For the reasons above, we created a local instance of the master template. We have access to a WMS and configuration sandbox. This enabled us to test how the feature behaves in the discovery layer and make adjustments where necessary. Being able to test the tool proved particularly useful when using wildcards and truncations.

After each class mark area was interrogated, we added new problematic subject terms to this sheet, and re-uploaded the sheet into the WMS Configuration dashboard. Uploading the sheet was incredibly easy and the changes surfaced in the Discovery layer almost instantaneously. Replacing the subject terms did not alter bibliographic data in MARC records at local or global level; it merely replaced the term in our local instance of WorldCat Discovery, with users clearly able to see the problematic terms that this feature replaced.

Latest search feature: local search expansion

In the latest iteration of the Locally Preferred Subject Feature, OCLC have introduced additional functionality, specifically for local search expansion. Users can expand their search using their preferred language: either using a replaced subject term or using their (problematic) preferred term. This iteration requires the use of a new template. It allows terms to be ranked as 'inclusive' or 'exclusive'. 'Exclusive' terms are not displayed to the user by default. This aspect of the new feature could involve a lot of extra staff time and endless decision-making. It could also be more time-consuming because wildcards and truncations do not work in this feature. For this reason, we decided to take the practical approach and implement this new feature in its simplest form. We are mirroring what we have done with the original template: listing our replacement terms and the terms that they replaced.

How this works in WCD

What do our users see when they conduct a search in WorldCat Discovery? In the first iteration of this tool, users could conduct a search using their preferred language. When we implemented the original feature, users conducting searches using a subject term that had been replaced received a pop-up alert to inform them that the term had been replaced. The message emphasised that the original term was still being included in a search of the catalogue:

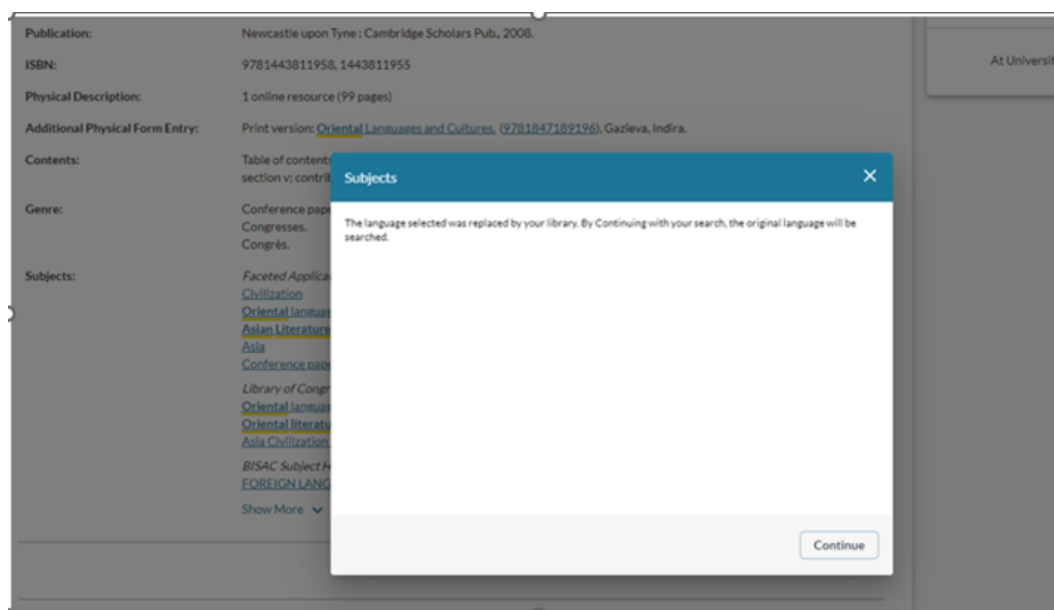


Figure 1: Original pop-up from the Locally Preferred Subject Feature in WorldCat Discovery.

This image is a screenshot/capture taken from OCLC's WorldCat Discovery, which is copyrighted by OCLC. All rights are reserved by them, and this image was used with their gracious permission.

In the subsequent iterations of the original tool, subject searches continue to include both problematic and replaced subject terms. Users are given the option to search using both the replaced search term and the original (problematic) term. They are then

given the choice to continue the search with corresponding problematic subject terms in the underlying bibliographic data, or just with their preferred term.

With the latest functionality, the process is almost the reverse: if a user searches with a term that has been identified as problematic and replaced in the local discovery layer, they get the option to expand their search to include search terms that correlate to their term (less harmful replacement terms) or just with their preferred term. For example, if a user searches using the term 'slavery' and finds a resource they want to view, they can click on the resource title to open the full record:

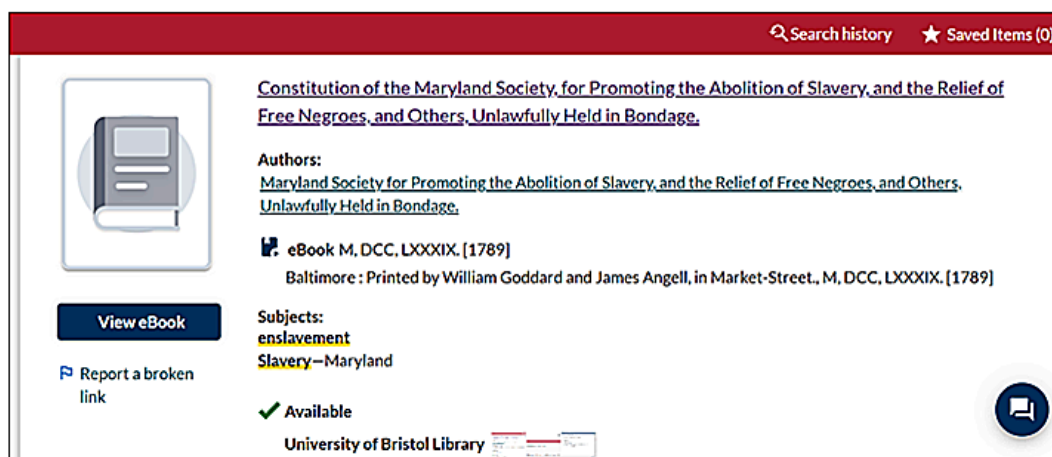


Figure 2: WorldCat Discovery record with highlighted subject terms. *This image is a screenshot/capture taken from OCLC's WorldCat Discovery, which is copyrighted by OCLC. All rights are reserved by them, and this image was used with their gracious permission.*

If they click on the replaced subject heading 'enslavement' (underlined in yellow in Figure 2 above) to expand the search in that direction (in lower case as a consequence of replacing the term as part of a string), they arrive at this page:

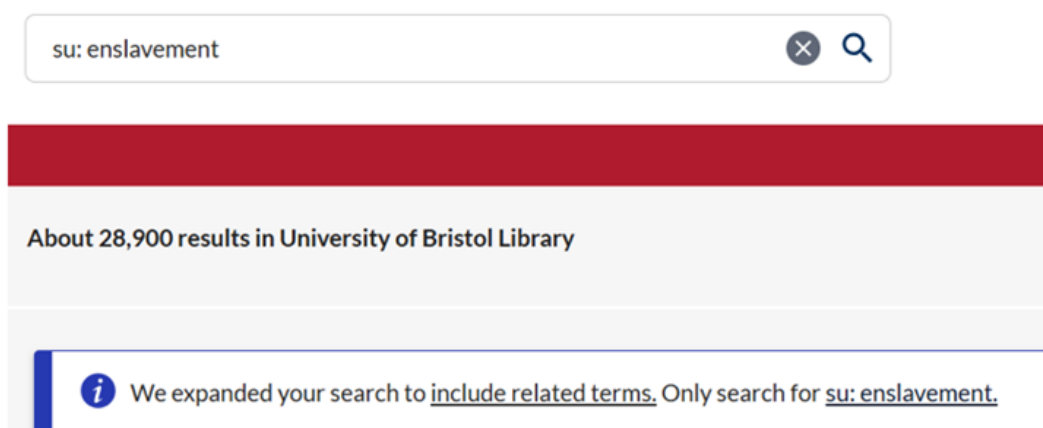


Figure 3: Expanded subject search in WorldCat Discovery. *This image is a screenshot/capture taken from OCLC's WorldCat Discovery, which is copyrighted by OCLC. All rights are reserved by them, and this image was used with their gracious permission.*

As we replaced the term 'slavery' with 'enslavement', the new functionality gives the option to click on the 'include related terms'. Once clicked on, there is another option to expand the extra heading 'sensitive terms'. This is how we can tell the difference

between the usual search expansion and the kind enabled by the problematic language workflow of the new feature:

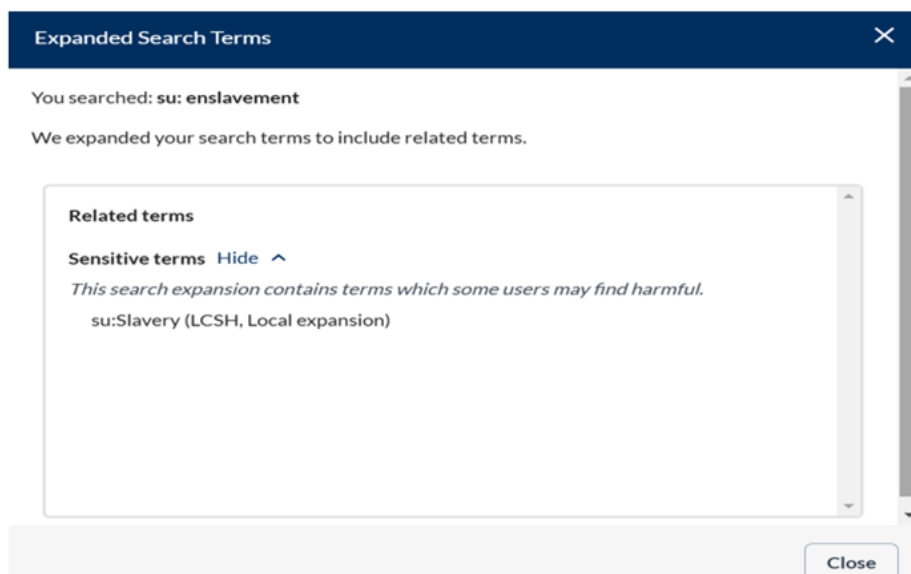


Figure 4: Expanded search terms pop-up in WorldCat Discovery. *This image is a screenshot/capture taken from OCLC's WorldCat Discovery, which is copyrighted by OCLC. All rights are reserved by them, and this image was used with their gracious permission.*

How we promote the project

We have a web page dedicated to our subject headings review on our library website: <https://www.bristol.ac.uk/library/find/subjectheadingsreview/>. This gives information on what we are doing and why. Users are invited to complete a form if they have suggestions for subject heading replacements they encounter. The submission can be anonymous. So far, we have not received any completed forms.

What we have learned

The process outlined above is now firmly in operation, with three of the group members working on downloading subjects for review, and the other members (mainly subject specialists) working through the terms. We have learned the following:

We have only found a handful of harmful subject headings that we have needed to replace. Subject areas that we have thought would be knotty, have proven to be relatively straightforward. Perhaps this reflects the reparative work that both the Library of Congress and OCLC (with FAST headings) are doing to update subject headings. These terms are being added to bibliographic records in WorldCat at global level by OCLC.

Some terms are highly contested and their meanings can vary depending on the context.

Wildcards do not currently work in the original template where subject terms have subheadings. We will feed this back to OCLC. It would be too onerous to create lines in the templates for all the instances with subheadings, so we are hoping that OCLC will address this in a future iteration.

We have not had any feedback from users. We are aware that many users are searching for known items. For those conducting subject searches, they are more focused on retrieving relevant search results rather than on the subject terminology itself.

Ongoing and future work

Once we have exhausted all the subject areas in ASSL, we aim to turn our focus to our other libraries whose collections are outside the arts and social sciences. This might prove to be a bigger challenge for collections which are not classified with LoC.

Cultural, political and scientific contexts are constantly shifting and evolving, which means that our project will be iterative. Accordingly, we will need to devise a method for verifying subject headings which have been amended and those that remain current and acceptable.

Talking to the wider UKI WMS User Group is helpful. A small number of members of the Subject Remapping Group meet with colleagues from three other institutions who also use WMS. These meetings are a good opportunity to share knowledge and best practice about workflows and replacement terms. As Bristol is a comparatively large institution there is the staff capacity to undertake the project we have described. Smaller libraries may not have as many staff available to analyse headings in a systematic way, but they are able to upload master templates to ensure that they are improving subject terminology on a smaller scale.

Conclusion

Our subject headings project at the University of Bristol continues to be rewarding and challenging. Perhaps the most significant challenge has been how and where to start analysing such a major collection, and how to record and justify the changes we propose as we continue with our work. Other major considerations include staff capacity for the project, expertise in programs such as Excel and MarcEdit, and having a robust method for tracking changes and capturing the sources used to inform our decisions. It is encouraging to discover, so far at least, that relatively few headings need attention, but this does not detract from the importance of this work. We would be very keen indeed to hear from colleagues at other institutions who have undertaken similar projects or who are considering similar processes.

Acknowledgements

Special thanks and acknowledgement to OCLC for their support of WorldCat, and its associated products and services. OCLC® and WorldCat® are registered trademarks/service marks of OCLC, Inc. All rights are reserved by them.

References

- Cho, M. K. et al (2023) Words matter: The language of difference in human genetics. *Genetics in Medicine*, 25(3), article number 100343. Available at: <https://doi.org/10.1016/j.gim.2022.11.011>
- Kenny, L. et al (2016) Which terms should be used to describe autism? Perspectives from the UK autism community. *Autism*, 20(4), pp. 442–462. Available at: <https://doi.org/10.1177/1362361315588200>
- Nottingham City Council (no date) *A Glossary of Terminology for Understanding Transatlantic Slavery and 'Race'*. Available at: <https://nottinghammuseums.org.uk/wp-content/uploads/2021/06/3a.-Slavery-and-Racial-Terminology-Glossary-Omitted-Terms.pdf> [Accessed: 15 January 2025]