Alter Heritage: a Web App to Gather Expert Knowledge on Inclusive Cultural Heritage Metadata

Andrei Nesterov[1,*], Laura Hollink[1] and Jacco van Ossenbruggen[2]

[1]*Centrum Wiskunde & Informatica, Science Park 123, 1098 XG Amsterdam, The Netherlands*
[2]*Vrije Universiteit Amsterdam, De Boelelaan 1105, 1081 HV Amsterdam, The Netherlands*

Abstract

In this demo paper, we present the web application Alter Heritage. Its goal is to support researchers in collecting domain experts' knowledge about inclusive cultural heritage metadata. While stakeholders in the cultural sector have been formulating strategies to mitigate biases in metadata, there is little empirical knowledge on how the metadata can be revised for inclusivity by users. With Alter Heritage, researchers can study what kind of alterations users make in artefacts' metadata to make it inclusive, how these alterations differ per artefacts and users. We develop the app's core metadata editing functionality based on requirements elicited from existing practices and guidelines in the cultural heritage domain. The data gathered via Alter Heritage will contribute to understanding of the specific alterations that make digital artefacts' metadata more inclusive.

Keywords

Metadata, Cultural Heritage, Inclusivity, Annotation, Domain Expert Knowledge, Web Application

1. Introduction

What users learn about cultural heritage on the websites of GLAMs (galleries, libraries, archives, and museums) largely depends on metadata. Thanks to the metadata, users can search through artefacts and get contextual information about them: dates, places, keywords, textual descriptions, and relations to other artefacts. Apart from useful information, metadata can also transmit cultural biases. For example, subject terms from controlled vocabularies may be offensive towards marginalised communities.[1] Titles and descriptions assigned by archivists are not free of prejudiced views either, which users might perceive as racist, sexist, homophobic, or ableist.[2]

To address biases in the metadata of artefacts, stakeholders in the cultural heritage domain have been developing strategies.[3, 4, 5] Their efforts have resulted in different documents: policies of cultural heritage institutions, academic research papers, or guidelines from formal and informal groups advocating for better representation of their heritage (such as Anti-Racist Description Working Group [6]). Despite numerous documents, there is little empirical knowledge about how experts can change metadata to reduce biases. This led us to to develop Alter Heritage, a web app for research purposes. The tool allows annotators to make changes in (biased) metadata and captures their input: for example, whether they edit text in certain metadata fields, hide them from the view or remove them entirely, add new fields, provide explanation notes or warnings to signal about biases. Using the app, we can study how annotators alter metadata of selected cultural heritage artefacts, whether there are differences between annotators, as well as between alterations of different artefacts.

We select 14 *domain documents*, such as the guidelines and strategies mentioned above. From these documents, we elicit requirements: which functionality the web app must offer to annotators in order to realise the suggestions on making the metadata inclusive. We implement the functionality for 38 user actions extracted from the documents, design the web app interface, and run a pilot with three users.

Table 1

An example: formulating the web app functionality based on a guideline from one of the domain documents (Metadata Best Practices for Trans and Gender Diverse Resources[8])

Suggestion	Action	Use case	Scenario	Functionality
"correct offensive or inaccurate language"	correct language	annotator removes offensive terms	by replacing terms in a textual field	edit a textual field
			by removing the whole textual field	remove a textual field

While multiple tools have been developed specifically for the cultural heritage domain to gather annotations (see related work in [7]), Alter Heritage is a research tool to study how annotators change metadata with the goal of increasing its inclusiveness. Apart from some common functionalities of existing tools, such as adding new tags or comments to artefacts, Alter Heritage incorporates other functionalities that primarily relate to dealing with offensive content: for example, hiding parts of the metadata on a web page or adding content warnings. To annotators, Alter Heritage offers an interactive environment for editing digital artefacts' metadata. For researchers, Alter Heritage allows setting up annotation studies with any type of artefacts, as long as their metadata can be put into the app's data scheme. The data collected with Alter Heritage will contribute to a deeper understanding of what makes cultural heritage metadata inclusive.

The Alter Heritage demo is available online[1] as well as its code with open licence and documentation.[2]

2. Requirements Elicitation

We use the domain analysis approach to elicit the web app requirements from the domain documents.[9]

First, we select domain documents about inclusive metadata strategies. We find them with keywords[3] using the Web search engines, academic repositories, and applying the snowballing method: following the references in the initial set of documents. While some documents are broad and general, others contain concrete actions. We targeted the latter ones and selected 14 documents in total (see the list in the documentation). Our aim was not to collect an exhaustive list of documents, but to collect a diverse set of actions of how metadata can be altered to increase its inclusivity.

Second, from the documents, we extract phrases with actions on making metadata inclusive. In total, we identified 48 actions. For each action, we formulate use cases: which user performs an action and how. For these use cases, we have only one type of user – an annotator. We specify use cases with 41 scenarios that illustrate how annotators might perform actions in practice using the web app.[10]

Third, covering all the scenarios, we outline the web app functionality. In Table 1, we provide an example of how we adapted suggestions from the domain documents into the web app functionality. The full table with actions, use cases, scenarios, and the functionality based on them is available in the documentation.

The requirements elicited from the domain documents concern only annotators editing metadata. Apart from them, we formulated additional requirements associated with how researchers would manage the annotation task: loading and displaying metadata, saving and retrieving annotators' input.

3. Functionality and Implementation

When annotators access Alter Heritage, they begin with a consent screen showing their privacy terms. If a consent is given, annotators see the main screen, which includes instructions, an artefact (an image) and its metadata on the left, navigation and questionnaire on the right (see Figure 1). Annotators can perform multiple actions editing metadata. These actions are applicable to both textual metadata fields

---

[1]https://alterheritage.project.cwi.nl/#ekaw24demo

[2]Documentation: https://github.com/cultural-ai/AlterHeritage

[3]For example, one of the queries we used was *"cultural heritage" AND metadata AND (inclusive OR bias)*

Figure 1: A screenshot of the Alter Heritage interface (in Dutch). The annotation instructions are collapsed in an accordion above, the artefact and its metadata are on the left, the navigation and questionnaire are on the right. The green check marks and the box inform a user about the task progress. This screenshot shows alterations made by a pilot user: added yellow note and red warning indicating the usage of the offensive term in the title.

and subject terms: adding a custom field or a subject term, removing, hiding/showing, and adding a note to them. To textual fields, annotators can also add a warning indicating, for example, offensive content. The button "Restore" allows annotators to restore the original metadata.

After completing their edits, annotators click "Submit" to send their changes to the server. Unsaved changes are stored in annotators' browser to allow navigation across artefacts preserving edits. The app notifies annotators when they submit edits and complete the task. Annotators can make more changes even after submission and resubmit them.

For researchers, the annotation tasks can be managed by creating JSON-files. There are 5 JSON-files linked to an annotator: consent, original metadata, edits, responses to questionnaire, and the submission progress. Researchers set ID for each annotator. This ID is used as a hash value in the web app link, for example, *#ekaw24demo*. The web app checks the hash value to read and write into corresponding files suffixed with the same ID. This way, researchers can distribute different sets of artefacts' metadata to different annotators. The initial metadata of artefacts should be put in a JSON-schema, so that the web app could parse and display it. Researchers can indicate explicitly in a JSON-file which metadata fields will be available for editing by annotators. The annotators' edits are stored separately from the original metadata in the same JSON-schema to ease the before-and-after comparison.

In this demo version, we chose not to implement such functionalities as image manipulation (e.g., cropping) and connecting external datasets (e.g., thesauri for pre-defined subject terms) that would

make the annotation process more complex. These features could be realised in future development.

Alter Heritage is built with HTML, CSS, and JavaScript, running on a Node.js[4] server. It utilises the Fetch API to parse JSON-files and POST requests to save annotators' edits.

To test the functionality and interface, we ran a pilot with three users. They edited the metadata of six artefacts, completed the questionnaire, and gave comments during the task. The pilot helped us eliminate errors in interface interactions and understand users' behaviour. As a result, we made the interface more intuitive based on users' feedback.

4. Demo Use Case

During the demonstration of the app, users will have the opportunity to test all the functionality we implemented and see how the data is being processed in real time. The app will load the metadata of six cultural heritage objects we selected for users to annotate. Originally, we developed and tested the interface in Dutch. For this demo, we will provide an English version.

There are various types of biases that can be found in the cultural heritage metadata. In our demo, we focus on one use case: (non)inclusive representation of LGBTQ+ people. We obtained artefacts' metadata that contains outdated and stereotyping terminology referring to LGBTQ+ people. First, we identified outdated and offensive LGBTQ+ terminology in specialised sources, such as Homosaurus[5] and a practical guide for searching LGBTQ+ related objects in archives.[11] Second, we queried this terminology in the metadata of artefacts from the European cultural heritage institutions using Europeana API[6]. We selected six objects with potentially non-iclusive metadata. For example, the historical term *"sodomites"* used in the title of one of the selected objects could be considered pejorative.[7] We expect participants to alter metadata using our app if they find the usage of such terms inappropriate to make the representation of LGBTQ+ related artefacts more inclusive.

5. Conclusions and Future Work

In the cultural heritage domain, it is a known problem that cultural biases are transmitted through digital artefacts' metadata. To mitigate the risks, the stakeholders in this domain have published multiple documents with various strategies. However, to the metadata developers, there is little empirical knowledge available on how metadata can be altered to reduce biases and represent heritage inclusively. We developed a web application Alter Heritage intended for research purposes. The tool's aim is to gather knowledge on how domain experts alter metadata of cultural heritage artefacts to make their representation inclusive. The core functionality of Alter Heritage was developed according to the requirements extracted from the domain documents, such as policies and guidelines. In Alter Heritage, users can make changes in (biased) metadata: edit, hide from the view, remove, or add content. Researchers can collect these changes to analyse which alterations contribute to making metadata inclusive, how changes differ per artefact or user.

In future work, Alter Heritage could be updated to support more metadata editing functionality, which we left out: images manipulation and support of external thesauri for auto filling of subject terms. Another additional feature could be a user-friendly admin panel for researchers, in which they can set up the desired components in the app interface (such as optional consent screen or questionnaire) and import/export metadata and annotations.

---

[4]https://nodejs.org/en
[5]https://homosaurus.org
[6]https://apis.europeana.eu/en
[7]https://homosaurus.org/v3/homoit0000227

Acknowledgments

References

[1] A. Nesterov, L. Hollink, J. van Ossenbruggen, How Contentious Terms About People and Cultures are Used in Linked Open Data, in: Proceedings of the ACM Web Conference 2024, WWW '24, Association for Computing Machinery, New York, NY, USA, 2024, p. 4523–4533. doi:10.1145/3589334.3648140.

[2] M. R. Rinn, Nineteenth-Century Depictions of Disabilities and Modern Metadata: A Consideration of Material in the P. T. Barnum Digital Collection, Journal of Contemporary Archival Studies 5 (2018). URL: https://elischolar.library.yale.edu/jcas/vol5/iss1/1.

[3] A. Chilcott, Towards protocols for describing racially offensive language in UK public archives, Archival Science 19 (2019) 359–376. doi:10.1007/s10502-019-09314-y.

[4] J. Guiliano, C. Heitman, Difficult Heritage and the Complexities of Indigenous Data, Journal of Cultural Analytics 4 (2019). doi:10.22148/16.044.

[5] K. Wright, Archival interventions and the language we use, Archival Science 19 (2019) 331–348. doi:10.1007/s10502-019-09306-y.

[6] A. Antracoli A., A. Berdini, K. Bolding, F. Charlton, A. Ferrara, V. Johnson, K. Rawdon, Archives for Black Lives in Philadelphia: Anti-Racist Description Resources., 2020. URL: https://github.com/a4blip/A4BLiP.

[7] E. Kaldeli, O. Menis-Mastromichalakis, S. Bekiaris, M. Ralli, V. Tzouvaras, G. Stamou, CrowdHeritage: Crowdsourcing for Improving the Quality of Cultural Heritage Metadata, Information 12 (2021). doi:10.3390/info12020064.

[8] The Trans Metadata Collective, J. Burns, M. Cronquist, J. Huang, D. Murphy, K. Rawson, B. Schaefer, J. Simons, B. M. Watson, A. Williams, Metadata Best Practices for Trans and Gender Diverse Resources, 2023. doi:10.5281/zenodo.10076876.

[9] D. Zowghi, C. Coulin, Requirements Elicitation: A Survey of Techniques, Approaches, and Tools, Springer Berlin Heidelberg, Berlin, Heidelberg, 2005, pp. 19–46. doi:10.1007/3-540-28244-0_2.

[10] H. Holbrook, A scenario-based methodology for conducting requirements elicitation, SIGSOFT Softw. Eng. Notes 15 (1990) 95–104. doi:10.1145/382294.382725.

[11] N. Shopland, A Practical Guide to Searching LGBTQIA Historical Records, LGBTQ histories, Routledge, Taylor & Francis Group, Abingdon, Oxon, 2021.