



A Knowledge Graph of Contentious Terminology for Inclusive Representation of Cultural Heritage

Andrei Nesterov¹(✉) , Laura Hollink¹ , Marieke van Erp² , and Jacco van Ossenbruggen³

¹ Centrum Wiskunde & Informatica, Amsterdam, The Netherlands
{nesterov,hollink}@cwi.nl

² KNAW Humanities Cluster, Amsterdam, The Netherlands
marieke.van.erp@dh.huc.knaw.nl

³ VU University Amsterdam, Amsterdam, The Netherlands
jacco.van.ossenbruggen@vu.nl

Abstract. Cultural heritage collections available as linked open data (LOD) may contain harmful stereotypes about people and cultures, for example, in outdated textual descriptions of objects. Galleries, libraries, archives, and museums (GLAM) have suggested various approaches to tackle potentially problematic content in digital collections. However, the domain expertise and discussions about words and phrases used in LOD-collections are scattered across different resources and detached from the collections themselves. In this paper, we capture domain expertise about English and Dutch contentious heritage terminology in a knowledge graph. Contentious terms in the resulting graph are then linked to entities from other LOD-resources used in the cultural domain and beyond, including Wikidata and WordNet. We make our design decisions explicit and report on the linking process. The developed knowledge graph makes expert knowledge interoperable, so it can be reused by the cultural heritage community and other LOD-developers to contribute to a more inclusive representation of cultural heritage on the Web.

Keywords: Knowledge Graph · Linked Open Data · Contentious Terms · Cultural Heritage

Resource type: Knowledge graph

License: CC BY-SA 4.0

DOI: [10.5281/zenodo.7456064](https://doi.org/10.5281/zenodo.7456064)

URL: <https://w3id.org/culco#>

Disclaimer. This paper contains derogatory words and phrases. They are provided solely as illustrations of the research results and do not reflect the opinions of the authors or their organisations. The derogatory and potentially offensive content is presented in “***quotes, boldfaced and italicised***”.

1 Introduction

Large collections of cultural objects are made available online as linked open data (LOD) on the websites of galleries, libraries, archives, and museums (or GLAM) [16,29]. Textual descriptions of objects in these collections may be originally written long time ago before digitisation. As a consequence, outdated language in digital cultural heritage may communicate historical stereotypes about people and cultures. In modern context, such stereotypes take forms of racism, ableism, homophobia, and other kinds of discrimination negatively affecting users [7,17]. Moreover, the stereotypes in LOD-collections might permeate applications built on top of such data [11,24].

The risks of problematic language are recognised by the cultural sector. GLAM have been developing approaches for more inclusive representation of objects in their collections [30]. For example, institutions provide explanations about inappropriate terminology in content warnings accompanying online collections¹ or publish general statements on their websites.² There is expert knowledge about problematic terminology that GLAM and other actors have produced, however, this knowledge is often detached from digital collections [18]. While object descriptions in collections are structured and often interconnected in knowledge organisation systems (KOS) used by heritage institutions, the domain expertise and discussions about problematic words in these collections exist in separate publications in different formats. To illustrate, the association Archives for Black Lives in Philadelphia published the document “Anti-racist description resources” [1], which recommends how to describe objects related to slavery, suggesting to use the terms “enslaved” or “captive” instead of “*slave*” when referring to people. At the same time, users do not see such discussions around the term’s usage when they find the word “*slave*” in LOD-collections.³

Curators of digital cultural heritage collections and other LOD-contributors can benefit from machine-readable resources that connect expert knowledge to potentially harmful content in their data. This paper aims to incorporate GLAM professionals’ domain knowledge about problematic terms into a knowledge graph to make the expert knowledge reusable and interoperable.

As a source of expert knowledge, we adopted the English and Dutch glossaries of problematic words and phrases found in museum databases. These glossaries are contained in the publication “Words Matter: An Unfinished Guide to Word Choices in the Cultural Sector” [22]. We refer to the problematic words and phrases from “Words Matter” as “contentious”. The “Words Matter” glossaries give explanations on why a certain term is considered contentious and suggests how to use terms appropriately including synonyms. The consistency of the

¹ A content warning in the Europeana gallery “Black people in European art”: <https://www.europeana.eu/en/galleries/black-people-in-european-art>. Accessed on 10.12.2022.

² The Getty Research Institute “Anti-Racist Statement”: https://www.getty.edu/research/institute/antiracist_statement.html. Accessed on 10.12.2022.

³ For example, the image entitled “slave from “[Across Africa, etc. [With a map and plates.]]” in the Europeana collection: <https://edu.nl/nttgk>. Accessed on 30.11.2022.

glossaries' structure was the main motivation for selecting "Words Matter" as a knowledge source, because it enabled identifying conceptual elements of the glossaries and modelling the relationships between them as a knowledge graph.

We formulated two research questions:

- **RQ1.** How can we model expert knowledge about the usage of English and Dutch contentious terms in the cultural heritage domain?
- **RQ2.** How can contentious terms from the developed knowledge graph be linked to other LOD-resources?

For RQ1, we elicited knowledge from the "Words Matter" glossaries. First, we examined the structure and content of the glossaries in two languages to define their conceptual elements. Second, we verified our conceptualisation conducting structured and unstructured interviews with the experts involved in the glossaries production. Third, we populated the knowledge graph with the original content of the glossaries based on the conceptualisation and interviews.

To answer RQ2, we selected two groups of resources to link them to the developed knowledge graph. The first group includes controlled vocabularies used in the cultural domain: Thesaurus Wereldculturen (NMVW) of the Dutch National Museum of World Cultures⁴ that produced "Words Matter" and the Getty Art & Architecture Thesaurus,⁵ which is used by many institutions. The second group consists of commonly used LOD-resources: Wikidata⁶ and Princeton WordNet 3.1.⁷ In each of the selected resources, we manually found entities that are the most relevant to the contentious terms in our knowledge graph, which we call "related matches". We linked both English and Dutch contentious terms to Wikidata and Getty AAT, only English terms to Princeton WordNet and only Dutch terms to the NMVW-thesaurus.

The main contribution of this paper is a knowledge graph representing domain expertise about English and Dutch contentious terminology used in cultural heritage. The contentious terms are linked to related entities from four LOD-resources frequently used in the cultural heritage sector and beyond. We report on our modelling choices and explain the process of identifying related matches of contentious terms in other LOD-resources. Having expert knowledge in machine readable format would facilitate the development of (semi-)automatic approaches to tackle potentially problematic terminology in LOD-collections making them more inclusive for users.

⁴ Thesaurus Wereldculturen: <https://collectie.wereldculturen.nl/thesaurus>. Accessed on 10.12.2022.

⁵ Getty Vocabularies: <https://www.getty.edu/research/tools/vocabularies>. Accessed on 10.12.2022.

⁶ Wikidata: https://www.wikidata.org/wiki/Wikidata:Main_Page. Accessed on 10.12.2022.

⁷ Princeton University "About WordNet": <https://wordnet.princeton.edu/> Princeton University. 2010. Accessed on 09.12.2022.

2 Related Work

Approaches to Contentious Terminology in the Cultural Sector. GLAM practitioners and researchers have formulated various approaches to tackle problematic terminology in heritage collections to make them more inclusive for users. Two groups of such approaches are especially relevant to our work. First, there are approaches directed at exposing controversies. They include marking offensive terms in objects metadata with special symbols (brackets and quotes), displaying content warnings, and providing appropriate synonyms next to offensive terms [6, 7, 32]. The second group is related to contextualisation: enriching offensive terms with additional information about why they are used to describe objects, who used them and during which historical periods [10, 14, 17].

The knowledge graph we developed aims at connecting contentious terms to their alternatives and explanations from experts. It relates to both making contentious terms visible in LOD and their contextualisation.

Modelling Problematic Language in LOD. Datasets of problematic language are used in various areas of computer science, one of them being hate-speech detection in social media in the field of natural language processing. Hurtlex is a multilingual lexicon with several categories of offensive terms, including “negative stereotypes” [2], the category closely related to contentious terms in our knowledge graph. Although Hurtlex is not available as linked open data, the terms are given identifiers and mapped to their equivalents in other languages. Apart from lexicons, there are ontologies developed to formalise offensive language for its detection. These ontologies are based on categorisations of offensive terms drawn from corpus analysis [20] and conceptualisation of definitions and theories of hate speech [3].

Another direction to systematise offensive terms for their auto-detection is developing extensions to existing LOD-resources. The Open Multilingual Wordnet (OMW)⁸ is enriched with Japanese offensive terms taken from several lexicons [4]. The researchers, who proposed this enrichment, analysed how offensive terms can be categorised looking at Princeton WordNet. They also manually mapped offensive terms to synsets in OMW. In another research project, Princeton WordNet synsets are linked to the terms scraped from social media, including the “vulgar” terms, based on manual annotations [21]. This is similar to our approach of matching contentious terms to synsets in Princeton WordNet.

One LOD-resource that contains such categories of terms as “slurs” and “historical” is “Homosaurus”, a controlled vocabulary of LGBTQ+ terms.⁹ Besides this categorisation, the vocabulary contains textual explanations of offensiveness (literal values of the property “*rdfs:comment*”). A use-case with “Homosaurus”

⁸ Global WordNet Association on GitHub: <https://github.com/globalwordnet/OMW>. Accessed on 18.12.2022.

⁹ Homosaurus. An international LGBTQ+ linked data vocabulary: <https://homosaurus.org>. Accessed on 18.12.2022.

illustrated how additional information about terms' usage can contextualise discriminatory terms in Library of Congress Subject Headings (LCSH)¹⁰ and move to the terminology accepted by the community [15].

There are three key differences between the knowledge graph we developed and existing lists and LOD-vocabularies of offensive terms. First, our knowledge graph is based on cultural heritage domain expertise containing suggestions and alternatives for contentious terms in English and Dutch. In our modelling process, we do not categorise offensive terms ourselves, but preserve the experts' judgments. Second, our modelling allows to mark terms as contentious depending on context. Third, contentious terms are linked to four LOD-resources used both in the cultural sector and on the Semantic Web in general. These links are helpful in gathering more background information about the terms.

The development of a knowledge graph in this paper extends our previous work, in which we constructed a crowdsource-annotated corpus of contentious terms in contexts taken from historical newspapers [5]. The corpus was used for machine-learning based detection of contentiousness. The combination of this corpus and the knowledge graph can be used to improve the detection of contentious terms in heritage collections.

3 Eliciting Knowledge About Contentious Terms

A selection of common contentious terms in the cultural heritage domain is described in the publication “Words Matter: an unfinished guide to word choices in the cultural section”. It is freely available online as a PDF-file on the website of the Dutch National Museum of World Cultures.¹¹ The publication’s goal is to provide guidance on word use to cultural heritage professionals, so that their choices of describing heritage objects “are more conscious and informed” [22]. “Words Matter” provides glossaries of contentious terms in English and Dutch, which we took as a source of expert knowledge.

We elicited knowledge about contentious terms in two steps applying direct and indirect knowledge elicitation techniques described in [8, 31]. First, we analysed the structure of the “Words Matter” glossaries and identified their conceptual elements. These conceptual elements served as building blocks of the knowledge graph schema. Second, we conducted structured and unstructured interviews with experts, who took part in producing the publication.

3.1 Identifying Conceptual Elements in “Words Matter”

Contentious and Suggested Terms. The “Words Matter” glossaries include terms “that are sensitive to particular groups, that can cause offense, that elide important context, and that are understood as derogatory” [22]. We refer to such

¹⁰ Library of Congress. Controlled Vocabularies: <https://www.loc.gov/librarians/controlled-vocabularies/>. Accessed on 10.12.2022.

¹¹ “Words Matter - Publication”: <https://www.tropenmuseum.nl/en/about-tropenmuseum/words-matter-publication>. Accessed on 02.12.2022.

terms as “contentious”. “Suggested terms” are the words and phrases mentioned in “Words Matter” that serve as alternatives to contentious terms.

The nature of the contentious terms in “Words Matter” is heterogeneous, although most of them refer to (historically) marginalised people and cultures of the Dutch colonial period. Some of the terms are archaisms (“**Bombay**” as the former name of the city Mumbai in India), others have sensitive connotations only in specific contexts (for example, the term “**primitive**” when referring to peoples, cultures, styles and art). Many terms in the list may be defined as “one-sided terms” from the framing bias perspective [26] (for example, referring to the Dutch-Indonesian war as “**police actions**”). All terms from the glossaries can be seen as “sensitive lexical items” from the lexicographic perspective [19]. Many terms in the glossaries appear as contentious only in particular senses and contexts, because one term can have several meanings (polysemy) or share the same spelling with a term that have a different meaning (homonymy).

The Structure of the Glossaries’ Entries. A single entry of the glossary has three main parts: a title which is a contentious word or a phrase, a textual part below it entitled “History, use, and possible sensitivities” (which we call a description), and the section “Suggestions”. We present conceptual elements of the “Words Matter” glossaries on Fig. 1.

Apart from the title, contentious terms also appear in descriptions. For example, the entry “**Aboriginal**” mentions other terms that are marked as “controversial”: “**Indian**”, “**Inuit**”, and “**Métis**”. In some of these cases, entries reference other entries with the text “see also”. The section “Suggestions” has individual suggestions in a bulleted list, which include various content:

- A general suggestion that can be applicable to several entries (for example, “The term is appropriate when used respectfully”);
- A word or a phrase that can be used instead of a contentious term (“Asian” for the term “**Oriental**”);
- A synonym that can be used only in some contexts and does not fully replace a contentious term (for example, the term “Moroccan-Dutch” is one of the possible alternatives for the term “**Allochtoon**”);
- An example of how a contentious term or its synonym can be used appropriately in speech; usually, it is a phrase providing additional context in which it is appropriate to use the mentioned terms, for example, the phrase “There was an artistic movement called ‘primitivism’” for the term “**Primitive**”; this kind of usage examples are italicised in the publication.

Differences Between the English and Dutch Glossaries. Most of the entries in the “Words Matter” glossaries were originally written in Dutch and then translated into English. The Dutch and English versions contain 56 and 55 entries, respectively. In the English glossary, five entries have Dutch titles: “**Blank**” (meaning “white”), “**Inboorling**” (“native”), “**Indisch**” (refers to the former Dutch East Indies), “**Jappenkampen**” (“Japanese concentration

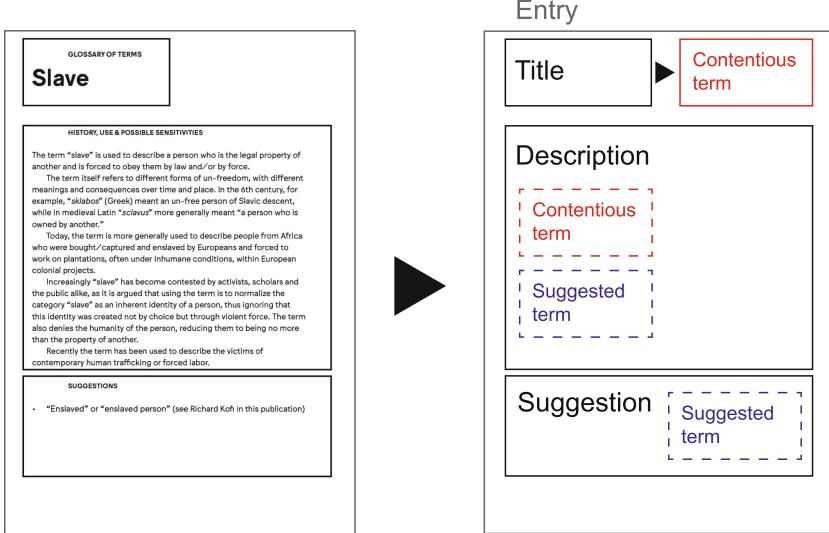


Fig. 1. The conceptualisation of the “Words Matter” glossary entries. A page from the glossary is on the left, and its conceptual elements are on the right. The dashed line shows that contentious and suggested terms are optional in the description text and suggestions.

camps”), and “**Politionele actie**” (“police action”), however, their descriptions and suggestions are in English. As the experts explained to us, these terms are kept in Dutch, because they would lose their meaning and context after translation. For the same reason, the entries “**Inlander**” (“native”) and “**Islamiet**” (“Muslim”) are not translated into English and are unique to the Dutch glossary. In the English version, the entry “**Native**” is not a translation of “**Inlander**”, it has unique description text and suggestions. These differences were important in our decision to separate the English and Dutch entries while populating the knowledge graph (Sect. 4.3).

3.2 Interviews with Domain Experts

We verified our work with two experts from the Dutch National Museum of World Cultures, who were involved in the creation of “Words Matter”. We held a meeting with the experts conducting both structured and unstructured interviews. During the unstructured interview, we presented our preliminary work to the experts and discussed the identified conceptual elements as well as the modeling choices we made while creating the knowledge graph (see Sect. 4.2 for modeling choices). The structured interview consisted of detailed questions about difficult cases, which arose when populating the knowledge graph with the glossary entries (Sect. 4.3). After the consultation, we refined the schema and

content of the knowledge graph. The meeting notes and the experts' responses are documented and published in the resource repository.¹²

3.3 Motivation to Select “Words Matter”

Other cultural heritage institutions and communities also bring awareness to problematic language used in heritage collections. Results of such initiatives are published in different forms such as blog posts,¹³ policy documents, as the aforementioned “Anti-racist description resources”. The website “The Cataloging Lab” lists 57 organisations that published statements about offensive language in various forms.¹⁴ On this website, users can suggest potentially problematic terminology from Library of Congress Subject Headings (LCSH).¹⁵ A similar possibility to users is offered by Triangle Research Libraries Network (TRLN).¹⁶ This organisation has also made available a list of the subject headings remappings: 216 pairs of problematic and suggested words and phrases.¹⁷

An example from the Dutch context is a report of the Cultural Heritage Agency of the Netherlands about “traces of slavery” in art collections [25], which contain list of terms used to search records about slavery. It is stated in the report that some of the search terms are derogatory and offensive, although it is not specified which terms were considered as such and by whom.

Compared to other resources, “Words Matter” provides more comprehensive information about contentious terms. It includes English and Dutch glossaries of contentious terms often found in museum databases not limiting the scope to a specific topic (such as “slavery”). There is background information about contentious terms and suggestions on their usage in a modern context. These glossaries have a consistent structure with relationships between contentious terms and suggestions. This motivated our selection of “Words Matter” as a source of expert knowledge to develop the first version of a knowledge graph of contentious terminology, but in future work, others can be added.

¹² Interviews with domain experts. Meeting notes: https://github.com/cultural-ai/wordsmatter/raw/main/Meeting_Notes_12Oct2021.pdf.

¹³ “California State University Libraries to change the display of the subject heading “Illegal Aliens” in joint public catalog”: <https://libraries.calstate.edu/csu-libraries-change-subject-heading-illegal-aliens/>. Accessed on 02.12.2022.

¹⁴ “List of Statements on Bias in Library and Archives Description”: <http://cataloginglab.org/list-of-statements-on-bias-in-library-and-archives-description/>. Accessed on 02.12.2022.

¹⁵ “Problem LCSH”: <https://cataloginglab.org/problem-lcsh/>. Accessed on 02.12.2022.

¹⁶ “TRLN Discovery Subject Remapping”: <https://trln.org/resources/subject-remapping/>. Accessed on 02.12.2022.

¹⁷ The GitHub repository “marc-to-argot” of Triangle Research Libraries Network: <https://edu.nl/kbaxv>. Accessed on 02.12.2022.

4 Developing a Knowledge Graph of Contentious Terms

The conceptual elements identified in the “Words Matter” glossaries were transformed into classes and properties following the principles we formulated. This section explains the modelling decisions and population of the knowledge graph.

4.1 Modelling Principles

Before converting the glossaries into a knowledge graph, we set three principles to guide the modelling process: (1) preserving the integrity of the original publication, (2) reusing existing LOD vocabularies when possible, and (3) allowing extension and reuse of the developed knowledge graph and its schema in other cases not limited to “Words Matter”.

The first principle stems from the fact that the “Words Matter” publication represents domain expert knowledge generated by a team of cultural heritage professionals and researchers, and any modification of the publication structure and content might influence the integrity of this knowledge. We held interviews with the experts to ensure that the first principle is respected. The second and third principles represent best practices of ontology development in the Semantic Web community [23]. The third principle enables reusing the knowledge graph in future work.

4.2 Modelling Choices: Classes and Properties

We present the knowledge graph schema in Fig 2. Following our second modelling principle, we searched for existing properties and classes in the W3C Data recommendations¹⁸ and the “Linked Open Vocabularies” register.¹⁹

It is important to differentiate between a (SKOS) concept or (Wiki) entity (for example, a Wikidata item Q12773225 with the label “*slave*”²⁰) versus a discussion *about* the term (e.g. a discussion about “*Slave*” being a contentious issue). To avoid confusion, we introduced a new class *ContentiousIssue* for the latter, instead of reusing, for example, *skos:Concept*. In “Words Matter”, a term can be contentious while serving as a suggestion for another contentious term. To model this and other discussions about terms, we assigned each term an URI using the “SKOS eXtension for Labels” schema (SKOS-XL).²¹ We modelled contentious and suggested terms as instances of the *skosxl:Label* class with its *skosxl:literalForm* taken from “Words Matter” text as is.

¹⁸ W3C “All standards and Drafts”: <https://www.w3.org/TR/?tag=data&status=REC>. Accessed on 05.12.2022.

¹⁹ Linked Open Vocabularies (LOV): <https://lov.linkeddata.es/dataset/lov/>. Accessed on 05.12.2022.

²⁰ Wikidata entity “slave” (Q12773225): <https://www.wikidata.org/wiki/Q12773225>. Accessed on 06.12.2022.

²¹ SKOS Simple Knowledge Organization System eXtension for Labels (SKOS-XL) Namespace Document - HTML Variant: <https://www.w3.org/TR/skos-reference/skos-xl.html>. Accessed on 06.12.2022.

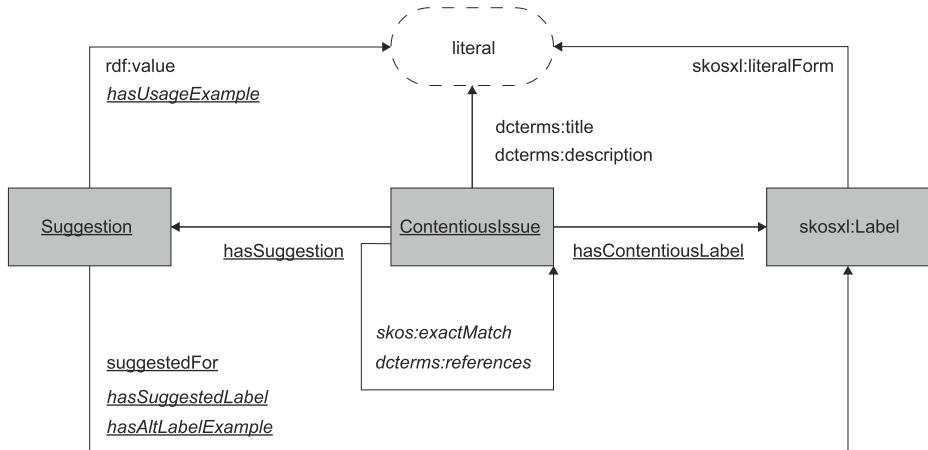


Fig. 2. The knowledge graph schema with custom classes and properties underlined. The italicized properties are optional.

To differentiate between contentious terms as a concept (*skosxl:Label*) and as a name of the glossary entry, we used the *dcterms:title* property from the DCMI Metadata Terms.²² So, each glossary entry has a title as a literal value and the contentious term it describes as an URI.

The instances of the *Suggestion* class represent individual suggestions that are given as separate bullet points of the section “Suggestions”. There is at least one suggestion item in every entry of the glossary. In several entries, suggestions have similar meanings but are phrased differently. For example, the suggestion “Use terms that people find respectful and acceptable for others to call them” (for the term “**Colored**”) is similar to “Adopt the terminology used and accepted as respectful by the people themselves” (for the term “**Aboriginal**”). We decided to count these suggestions as equivalents, giving them the same textual value. It allows inferring which contentious terms share similar suggestions. We preserved the original suggestions in a separate file.²³

We modelled the relationships between the classes *ContentiousIssue* and *Suggestion* with the property *hasSuggestion*. *ContentiousIssue* is connected to *skosxl:Label* with another custom property *hasContentiousLabel*. For example, *ContentiousIssue* entitled “**Slave**” has contentious label with the literal form “**Slave@en**”, and there is one suggestion for this contentious issue with two suggested labels (*hasSuggestedLabel*) “Enslaved@en” and “enslaved person@en”.

Individual suggestions may have different content, modelled by different properties. If an instance of *Suggestion* has a concrete term or a phrase to be used instead, it has the property *hasSuggestedLabel*. Another property *hasAltLabelExample* indicates that a suggestion gives a suggested contextual synonym

²² DCMI Metadata Terms: <https://www.dublincore.org/specifications/dublin-core/dcmi-terms/>. Accessed on 06.12.2022.

²³ <https://edu.nl/t7n7v>.

Table 1. There are more Contentious Issues with Dutch titles, and the English version has more Suggestions and Suggested labels.

	Contentious Issues	Suggestions	Contentious labels	Suggested labels	Total labels
@en	50	87	75	48	123
@nl	61	81	83	41	124
Total	111	168	158	89	247

for a contentious term, while the value of *hasUsageExample* represents how a contentious or suggested term might be used in speech appropriately. Additionally, we explicitly stated for every *Suggestion* for which contentious label (*skosxl:Label*) it is a suggestion by using *suggestedFor*.

Other properties in the schema were adopted from existing vocabularies. The value of the *dcterms:description* property from DCMI Metadata Terms is the text from the “History, use, and possible sensitivities” section, which describes the glossary entry. Although this section gives extensive information about the usage of contentious terms, their cultural contexts, and etymology, we kept it as a literal value. Breaking down this textual information into semantically related parts requires more complex modelling. We model implicit “see also” references in the description text as explicit *dcterms:references* between Contentious Issues. The textual content of each *Suggestion* is a literal value of the property *rdf:value*. In total, we have introduced three custom classes and six custom properties.

4.3 Populating the Knowledge Graph

The knowledge graph schema was manually populated with the original content from the “Words Matter” glossaries in English and Dutch. Terms (*skosxl:Label*), the entries they appear in (*ContentiousIssue*), and suggestions (*Suggestion*) were given URIs. We decided to use meaningless URIs to avoid offensive URIs containing more meaningful terms. Textual content has been tagged “@en” or “@nl” where appropriate.

Because of the differences between the English and Dutch glossaries (see Sect. 3.1), we gave the entries of the two versions separate identifiers. In cases when the entities in both languages were equivalent (if translated), we connected them with the property *skos:exactMatch*.²⁴ For example, the contentious issue “**Slaaf**” is a *skos:exactMatch* of its English equivalent “**Slave**” with a different URI. The number of instances of the populated knowledge graph in two languages is given in Table 1.

Competency Questions. We formulated competency questions to ensure that the knowledge graph follows the structure of the original glossaries. The competency questions check the relationships between the English and Dutch versions

²⁴ SKOS Simple Knowledge Organization System Reference: <https://www.w3.org/TR/skos-reference/#L4858> Accessed on 06.12.2022.

of the glossaries and their parts, such as entries, suggestions, and terms. For example, we checked with a SPARQL-query which contentious terms had suggested terms. There are 10 competency questions, answers to which serve as evaluation of the developed knowledge graph. The questions and the SPARQL-queries we used to answer them are available in the resource documentation.²⁵

Availability and Reuse. The knowledge graph schema is available at <https://w3id.org/culco#>. We documented it following the FAIR practices [12, 13]. The glossary itself can be downloaded in Turtle format at <https://w3id.org/culco/wordsmatter/>. The developed resource is published on GitHub²⁶ and registered on Zenodo.²⁷

5 Linking Contentious Terms to LOD-Resources

The resulting knowledge graph has URIs of contentious labels, which we link to four LOD-resources. This section explains the linking process based on the guidelines we set and gives an overview of the obtained links.

5.1 Selecting LOD-Resources

We link contentious terms from the knowledge graph to other LOD-resources for two reasons: 1) it enriches contentious terms with related concepts, in literal values of which they are used, and 2) it connects the found occurrences of contentious terms in external LOD-resources to their suggested labels and explanations from experts in the knowledge graph. We selected four LOD-resources to be linked to contentious labels in the knowledge graph: controlled vocabularies used by cultural heritage institutions (Wereldculturen Thesaurus (NMVW) and Getty AAT) and commonly used LOD-resources (Wikidata and Princeton WordNet 3.1). In every resource, we searched for a related match (the term derives from the property in the SKOS-vocabulary *skos:relatedMatch*) of every contentious label. A related match is a concept that uses a contentious term in its labels, and the meaning of the term is the closest to the meaning in “Words Matter”. In the case of Princeton WordNet, it is a synset with a contentious term as a lemma. Table 2 lists the LOD-resources selected for linking, their properties used for labelling, and the number of the found related matches.

Wikidata is one of the largest knowledge graphs on the Web with a variety of application areas, including cultural heritage [16, 27]. Princeton WordNet can provide lexical information about contentious terms, including their synonyms, definitions, and examples. Getty AAT serves as a reference resource to many cultural heritage institutions [9]. Having related matches of contentious labels in Getty AAT can be helpful in finding links to more cultural heritage datasets.

²⁵ Competency questions: <https://edu.nl/3ttu6>.

²⁶ Cultural AI Lab on GitHub: <https://github.com/cultural-ai/wordsmatter>.

²⁷ DOI: <https://doi.org/10.5281/zenodo.7456064>.

Table 2. SKOS is used in Wikidata and NMVW for labeling entities. In Wikidata, there are two other equivalent properties for a preferred label. Getty AAT adopts SKOS-XL. Princeton WordNet uses OntoLex vocabulary for written representation of lemmas.

Resource	Properties for labelling	Language	# related matches
Wikidata	skos:prefLabel (rdfs:label, schema:name); skos:altLabel	EN	61
		NL	70
Princeton WordNet 3.1	ontolex:writtenRep	EN	56
Getty AAT	xl:prefLabel/xl:literalForm; xl:altLabel/xl:literalForm	EN	42
		NL	27
NMVW	skos:prefLabel; skos:altLabel	NL	19

The NMVW-thesaurus is used by the Dutch National Museum of World Cultures that published “Words Matter”. We link contentious terms in our graph based on knowledge of experts from this museum to the actual thesaurus that is used to represent its collection.

5.2 Identifying Related Matches

Querying LOD-Resources. We linked contentious terms to external entities in the selected LOD-resources at the level of the *Label* class. For linking, we took only labels that are in the object position of a triple $\langle ContentiousIssue \text{ hasContentiousLabel } Label \rangle$. It means that these labels are marked as contentious in the glossary entry. In total, there are 75 English and 83 Dutch contentious labels.

The literal values of the contentious labels are mostly in a singular form. To find contentious labels in other resources that occur in plural or comparative (for adjectives) forms and other variations (spelling differences), we collected word forms for every contentious label using external datasets. Word forms for English terms were obtained from DBnary using their API [28]. For Dutch terms, we used a HTTP-request service provided by INT, the Dutch Language Institute.²⁸ After the manual inspection of the extracted word forms in both languages, we observed that word forms for some labels were still missing. We added more word forms for 18 labels. As a result of this step, every contentious label was given a list of word forms, which resulted in 154 English and 242 Dutch query tokens. This is expected as Dutch is a morphologically richer language than English.

To find entities that have contentious terms in their literal values, we searched every token in the selected LOD-resources. In Wikidata and Getty AAT, we searched both English and Dutch tokens. Princeton WordNet 3.1 (PWN) were

²⁸ Dutch Language Institute. GiGaNT: <https://ivdnt.org/corpora-lexica/gigant/>. Accessed on 18.12.2022.

searched for English tokens. NMVW is only available in Dutch, so we searched for Dutch tokens in this thesaurus.

Guidelines to Select Related Matches. For most of the query tokens, there was a large number of query results in the resources. For example, the term “black” has more than 134,000 hits in Wikidata. As our goal was to select only related matches, we performed this selection manually. To be consistent during selection, we formulated the following guidelines:

1. A query token of a contentious term is used in literal values of the corresponding properties of the resources (see the column “Properties for labelling” in Table 2);
2. The found token should be used in a similar meaning to the meaning it has in the associated “Words Matter” glossary entry;
3. If multiple entities are found, we pick one entity which is the closest in meaning and scope to what is intended in the “Words Matter” entry. An exception is made for PWN.

In some cases in PWN, it was not possible to differentiate the meanings of the found related synsets, so we allowed more than one related match. This resulted in 24 English contentious labels having more than one related synset in PWN.

The literal values of the resources, in which a query token was found, were taken into account to judge the meaning of a token (guideline 2). For example, apart from preferred and alternative labels (aliases) in Wikidata, we also looked at “Description”.

To ensure high-quality of the related matches selection, two authors of the paper, of whom one is a native Dutch speaker, performed this step independently. The disagreements and mismatches were resolved during a discussion between the authors.

Linking Results. We obtained 275 related matches from four LOD-resources, 159 of which are for English labels, 116 are for Dutch labels. Almost half (131) of these related matches come from Wikidata (Table 2 provides the number of matches per resource). 142 out of all 158 contentious labels in English and Dutch were linked to at least one of the selected LOD-resources. 29 out of 75 English and 10 out of 83 Dutch contentious labels have related matches in all the corresponding resources. In the knowledge graph, contentious labels are linked to the URIs of their related matches with the property *skos:relatedMatch*.

Looking at the occurrences of contentious labels in the literal values of the related matches, we found that 46 English and 50 Dutch contentious labels from our knowledge graph are used as preferred labels of the related Wikidata items. 23 English and 25 Dutch (out of 27 found) contentious labels are used as preferred labels of the related concepts in Getty AAT. Dutch contentious labels in NMVW were used as preferred labels in 7 out of 19 found concepts.

To illustrate the found related matches, we give an example of the label with the literal value “*Slave@en*”. It has three related matches from Getty

AAT,²⁹ Wikidata,³⁰ and PWN (one synset).³¹ In the related concept of Getty AAT, the term is used 4 times in alternative labels. The preferred label of this concept is “enslaved people”, which is similar to the suggestion for this term in our knowledge graph (“enslaved person”). On the contrary, in Wikidata, the related item uses the contentious term as preferred label, which aliases are “enslaved person” and “enslaved”. PWN does not give any synonyms of the term in the related synset. None of the resources contain any information about the term’s contentiousness and its potentially inappropriate usage.

6 Conclusions and Future Work

We constructed a knowledge graph representing English and Dutch contentious terminology often used in museum object descriptions. The resource is based on domain expert knowledge elicited from the publication “Words Matter” [22]. The publication’s consistent structure enabled identifying its conceptual elements that constituted the knowledge graph schema. Two domain experts verified our modelling choices and decisions regarding the knowledge graph population.

In total, there are 75 English and 83 Dutch terms in the knowledge graph that are potentially contentious depending on context. These terms are linked to explanations of their usage, suggestions, and alternatives given by experts. Additionally, we linked contentious terms to other LOD-resources used in the cultural sector and beyond: Wikidata, Princeton WordNet, Getty AAT, and the NMVW thesaurus. The resulting resource has been made openly available with a CC BY-SA 4.0 license following FAIR practices. In future work, the knowledge graph can be used to develop applications that highlight and contextualise offensive and outdated terms in cultural heritage objects’ descriptions, making their representation more inclusive for users.

The publication, on which the knowledge graph is based, originates from one organisation presenting a viewpoint of the European cultural context. Since 2017, when the publication was produced, new discussions about contentious terms have emerged. When using the knowledge graph, its limitations, such as time and scope, should be acknowledged and included, along with other sources, in future updates of the knowledge graph.

We observed that contentious terms in the literal values of the LOD-resources studied in this paper often appear without information about their potential sensitivities. In future work, our knowledge graph can be used for further research into how contentious and suggested terms are used in other LOD resources. A large-scale inspection of such cases could identify problematic aspects of using culturally-sensitive language on the Semantic Web.

Acknowledgements. This work is a part of the “Culturally Aware AI” project funded by NWO (Dutch Research Council). We thank the Dutch National Museum of World

²⁹ <http://vocab.getty.edu/aat/300230899>. Accessed on 18.12.2022.

³⁰ <https://www.wikidata.org/wiki/Q12773225>. Accessed on 18.12.2022.

³¹ <http://wordnet-rdf.princeton.edu/id/10628841-n>. Accessed on 18.12.2022.

Cultures for collaboration and personally Cindy Zalm, Marijke Kunst, Marjolijn van Beelen, and Richard van Alphen. Thanks to Antoine Isaac (Europeana Foundation) for giving recommendations at the first stages of the knowledge graph modelling.

References

1. Antracoli A.A., et al.: Archives for black lives in Philadelphia: anti-racist description resources (2020). <https://github.com/a4blip/A4BLiP>
2. Bassignana, E., Basile, V., Patti, V.: Hurtlex: a multilingual lexicon of words to hurt. In: Cabrio, E., Mazzei, A., Tamburini, F. (eds.) Proceedings of the Fifth Italian Conference on Computational Linguistics (CLiC-it 2018). CEUR Workshop Proceedings, vol. 2253. CEUR (2018). <https://ceur-ws.org/Vol-2253/#paper49>
3. Battistelli, D., Bruneau, C., Dragos, V.: Building a formal model for hate detection in French corpora. Procedia Comput. Sci. **176**, 2358–2365 (2020). <https://doi.org/10.1016/j.procs.2020.09.299>. Knowledge-Based and Intelligent Information & Engineering Systems: Proceedings of the 24th International Conference KES2020
4. Bond, F., Choo, M.Y.H.: Taboo wordnet. In: Proceedings of the 11th Global Wordnet Conference, pp. 36–43. Global Wordnet Association, University of South Africa (UNISA) (2021). <https://aclanthology.org/2021.gwc-1.5>
5. Brate, R., Nesterov, A., Vogelmann, V., van Ossenbruggen, J., Hollink, L., van Erp, M.: Capturing contentiousness: constructing the contentious terms in context corpus. In: Proceedings of the 11th on Knowledge Capture Conference, K-CAP 2021, pp. 17–24. Association for Computing Machinery, New York (2021). <https://doi.org/10.1145/3460210.3493553>
6. Brown, L., Gutierrez, C., Okmin, J., McCullough, S.: Desegregating conversations about race and identity in culturally specific museums. J. Museum Educ. **42**(2), 120–131 (2017). <https://doi.org/10.1080/10598650.2017.1303602>
7. Chilcott, A.: Towards protocols for describing racially offensive language in UK public archives. Arch. Sci. **19**(4), 359–376 (2019). <https://doi.org/10.1007/s10502-019-09314-y>
8. Cooke, N.J.: Varieties of knowledge elicitation techniques. Int. J. Hum.-Comput. Stud. Int. J. Man-Mach. Stud. (1994). <https://doi.org/10.1006/ijhc.1994.1083>
9. Díaz-Corona, D., Lacasta, J., Latre, M.Á., Zarazaga-Soria, F.J., Nogueras-Iso, J.: Profiling of knowledge organisation systems for the annotation of linked data cultural resources. Inf. Syst. **84**, 17–28 (2019). <https://doi.org/10.1016/j.is.2019.04.008>
10. Dodd, J., Sandell, R., Delin, A., Gay, J.: Arts and humanities research board: buried in the footnotes: the representation of disabled people in museum and gallery collections. Technical report, Leicester, RCMG (2004). https://leicester.figshare.com/articles/report/Buried_in_the_footnotes_the_representation_of_disabled_people_in_museum_and_gallery_collections/10077098
11. van Erp, M., de Boer, V.: A polyvocal and contextualised semantic web. In: Verborgh, R., et al. (eds.) ESWC 2021. LNCS, vol. 12731, pp. 506–512. Springer, Cham (2021). https://doi.org/10.1007/978-3-030-77385-4_30
12. Garijo, D.: WIDOCO: a wizard for documenting ontologies. In: d'Amato, C., et al. (eds.) ISWC 2017. LNCS, vol. 10588, pp. 94–102. Springer, Cham (2017). https://doi.org/10.1007/978-3-319-68204-4_9
13. Garijo, D., Poveda-Villalón, M.: Best Practices for Implementing FAIR Vocabularies and Ontologies on the Web, vol. 49, pp. 39–54. IOS Press (2020). <https://doi.org/10.3233/SSW200034>

14. Giuliano, J., Heitman, C.: Difficult heritage and the complexities of indigenous data. *J. Cult. Anal.* **4**(1) (2019). <https://doi.org/10.22148/16.044>
15. Hardesty, J., Nolan, A.: Mitigating bias in metadata. *Inf. Technol. Libr.* **40**(3) (2021). <https://doi.org/10.6017/ital.v40i3.13053>
16. Hawkins, A.: Archives, linked data and the digital humanities: increasing access to digitised and born-digital archives via the semantic web. *Arch. Sci.* **22**(3), 319–344 (2022). <https://doi.org/10.1007/s10502-021-09381-0>
17. Holterhoff, K.: From disclaimer to critique: race and the digital image archivist. *Digit. Humanit. Q.* **011**(3) (2017). <http://www.digitalhumanities.org/dhq/vol/11/3/000324/000324.html>
18. Kahn, R., Simon, R.: Feast and famine: the problem of sources for linked data creation. In: Andrews, T., Diehr, F., Efer, T., Kuczera, A., Zundert, J.V. (eds.) *Graph Technologies in the Humanities - Proceedings 2020*. CEUR Workshop Proceedings, vol. 3110, p. 86. CEUR (2020). <http://ceur-ws.org/Vol-3110/#paper5>
19. Lazić, D., Mihaljević, A.: Stereotypes and taboo words in dictionaries from a diachronic and a synchronic perspective - the case study of Croatian and Croatian Church Slavonic. In: *Proceedings of XIX EURALEX Congress: Lexicography for Inclusion*, vol. II, no. 2, pp. 643–653 (2021). https://www.euralex.org/elx_proceedings/Euralex2020-2021/EURALEX2020-2021_Vol2-p643-653.pdf
20. Lewandowska-Tomaszczyk, B., Źitnik, S., Bączkowska, A., Liebeskind, C., Mitrović, J., Oleskeviciene, G.V.: LOD-connected offensive language ontology and tagset enrichment. In: Carvalho, S., et al. (eds.) *Proceedings of the Workshops and Tutorials held at LDK 2021*. CEUR Workshop Proceedings, vol. 3064, pp. 135–150. CEUR (2021). <https://ceur-ws.org/Vol-3064/#salld-03>
21. McCrae, J.P., Wood, I., Hicks, A.: The colloquial wordnet: extending princeton wordnet with neologisms. In: Gracia, J., Bond, F., McCrae, J.P., Buitelaar, P., Chiarcos, C., Hellmann, S. (eds.) *LDK 2017. LNCS (LNAI)*, vol. 10318, pp. 194–202. Springer, Cham (2017). https://doi.org/10.1007/978-3-319-59888-8_17
22. Modest, W., Lelijveld, R.: Words matter: an unfinished guide to word choices in the cultural sector (2018). <https://www.tropenmuseum.nl/en/about-tropenmuseum/words-matter-publication>
23. Noy, N., McGuinness, D.: Ontology development 101: a guide to creating your first ontology. *Knowl. Syst. Lab.* **32** (2001). https://protege.stanford.edu/publications/ontology_development/ontology101.pdf
24. Ortolja-Baird, A., Nyhan, J.: Encoding the haunting of an object catalogue: on the potential of digital technologies to perpetuate or subvert the silence and bias of the early-modern archive. *Digit. Scholarsh. Humanit.* **37**(3), 844–867 (2021). <https://doi.org/10.1093/llc/fqab065>
25. Ras, G., Gootjes, E., Pennock, H., Vermaat, S.: Handreiking ‘Onderzoek naar sporen van slavernij en het koloniale verleden in de collectieregistratie’ (2021). <https://www.cultureelerfgoed.nl/publicaties/publicaties/2021/01/01/handreiking-onderzoek-naar-sporen-van-slavernij-en-het-koloniale-verleden-in-de-collectieregistratie>
26. Recasens, M., Danescu-Niculescu-Mizil, C., Jurafsky, D.: Linguistic models for analyzing and detecting biased language. In: *Proceedings of the 51st Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pp. 1650–1659. Association for Computational Linguistics, Sofia (2013). <https://aclanthology.org/P13-1162>
27. Rossenova, L., Duchesne, P., Blümel, I.: Wikidata and Wikibase as complementary research data management services for cultural heritage data. In: Kaffee, L.A.,

- Razniewski, S., Amaral, G., Alghamdi, K.S. (eds.) Proceedings of the 3rd Wikidata Workshop 2022. CEUR Workshop Proceedings, vol. 3262. CEUR (2022). <https://ceur-ws.org/Vol-3262/#paper15>
- 28. Sérasset, G.: DBnary: wiktionary as a Lemon-based multilingual lexical resource in RDF. *Semant. Web* **6**(4), 355–361 (2015). <https://doi.org/10.3233/SW-140147>
 - 29. Simou, N., Chortaras, A., Stamou, G., Kollias, S.: Enriching and publishing cultural heritage as linked open data. In: Ioannides, M., Magnenat-Thalmann, N., Papagiannakis, G. (eds.) Mixed Reality and Gamification for Cultural Heritage, pp. 201–223. Springer, Cham (2017). https://doi.org/10.1007/978-3-319-49607-8_7
 - 30. Sinn, D.H.: Archival description and records from historically marginalized cultures: a view from a postmodern window. *J. Korean Soc. Libr. Inf. Sci.* **44**(4), 115–130 (2010). <https://doi.org/10.4275/KSLIS.2010.44.4.115>
 - 31. Vásquez-Bravo, D.-M., Sánchez-Segura, M.-I., Medina-Domínguez, F., Amescua, A.: Guideline to select knowledge elicitation techniques. In: Lytras, M.D., Ruan, D., Tennyson, R.D., Ordóñez De Pablos, P., García Peñalvo, F.J., Rusu, L. (eds.) WSKS 2011. CCIS, vol. 278, pp. 374–384. Springer, Heidelberg (2013). https://doi.org/10.1007/978-3-642-35879-1_45
 - 32. Wright, K.: Archival interventions and the language we use. *Arch. Sci.* **19**(4), 331–348 (2019). <https://doi.org/10.1007/s10502-019-09306-y>