

# Defining an Ontology for Museum Critical Cataloguing Terminology Guidelines

**Erin Canning**

University of Oxford  
Oxford, UK

erin.canning@eng.ox.ac.uk

## Abstract

This paper presents the proposed ontology for the project “Computational Approaches for Addressing Problematic Terminology” (CAAPT). This schema seeks to represent contents and structure of language guideline documents produced by cultural heritage institutions seeking to engage with critical cataloguing or reparative description work, known as terminology guidance documents. It takes the Victoria and Albert Museum Terminology Guidance Document as a source for the initial modelling work. Ultimately, CAAPT seeks to expand the knowledge graph beyond the context of the Victoria and Albert Museum to incorporate additional terminology guidance documents and linked open data vocabularies. The ontology seeks to bring together scholarly communities in areas relevant to this project, most notably those in cultural heritage and linguistics linked open data, by leveraging existing linked data resources in these areas: as such, OntoLex, CIDOC CRM, and SKOS are used as a foundation for this work, along with a proposed schema from a related project, CULCO. As the CAAPT project is in early stages, this paper presents the preliminary results of work undertaken thus far in order to seek feedback from the linguistics linked open data community.

**Keywords:** cultural heritage, problematic terminology, linked open data, ontology

## 1. Introduction

Cultural heritage institutions are increasingly aware of the presence of bias and problematic and offensive language in texts of their catalogue records, as evidenced by the growing interest in critical cataloguing (Watson, 2023). There has been effort in the field to define what is meant by “problematic terminology” and therefore what institutions could, or should, examine in catalogue reviews (Chew, 2022; Cress, 2021; Dalal-Clayton & Rutherford, n.d.; Lawther, 2021; Muñoz, 2021; Museums Association, 2021; Ortolja-Baird & Nyhan, 2022; Rutherford, 2021a, 2021b, 2022). For example, the above authors identify “problematic terminology” as encompassing explicit slurs, euphemisms, and derogatory, objectifying, and dehumanizing language, as well as colonial and incorrect names of peoples, places, and types of objects.<sup>1</sup> However, there is little sector-wide guidance on what all this heading could include and the need to share information between institutions in pursuit of the development of best practices is well known (Chew, 2023; Dalal-Clayton & Rutherford, n.d.; Museums Association, 2020, 2021). At the level of individual institutions, museums are developing—and implementing—terminology guidance documents: these are glossary-like documents that list terms that the institution is interested in looking for in their cataloguing, often accompanied by a description of the term and a history of use that may give context to why the term was used when authoring catalogue records, paired with suggestions for actions to take when the term is found in the record. These suggestions are highly dependent on context, and include options such as to replace the term, to format it in a particular way that indicates its historical nature, or to add specific or general explanatory text, to give

three examples. These documents themselves are objects of potential scholarly interest: in addition to being a way for museums to communicate internally about emerging best practices, they show what terms museums are interested in addressing in their catalogue records and how they are thinking about defining such language. As such, terminology guidance documents may hold interest for linguistics as well as cultural heritage scholarly communities.

The project “Computational Approaches for Addressing Problematic Terminology” (CAAPT) seeks to make the contents of these terminology guidance documents available to institutions looking to engage in critical cataloguing as well as to relevant scholarly communities through the use of linked open data (LOD). The first step in this is to define the structure required to represent this information. This paper introduces the proposed ontology for CAAPT, based off of the Victoria and Albert Museum Terminology Guidance Document.

## 2. The Victoria and Albert Museum Terminology Guidance Document

The Victoria and Albert Museum (V&A) contains close to 1.7 million works of art and design objects acquired over more than 170 years of collecting activity, and which is still ongoing (Victoria and Albert Museum, 2023). The museum’s catalogue records represent objects from vast reaches of time and place, and as such contain a wide variety of problematics. The V&A holds regular cross-department meetings to discuss terminology questions and concerns raised by staff. This working group, in collaboration with the Interpretation Department and additional staff-led internal advisory groups, has produced and maintains

<sup>1</sup> For practical purposes within context of this paper, the author considers “problematic terminology” to be the terms listed in terminology guidance documents. For conceptual framing, the author will propose that

“problematic terminology” be understood as language which enables a catalogue record to perform or play into Haraway’s (1988) concept of “the god trick”; discussion of this falls outside of the present scope.



resources. Therefore, a small number of classes are proposed, with the focus instead on properties that bring together classes from these four ontologies. Furthermore, all classes and almost all properties—labelled here as “Computational Approaches for Addressing Problematic Terminology” (caapt)—are declared as subclasses and subproperties of elements from one or more of these four ontologies.

## 4.2 CAAPT Proposed Classes

The six classes proposed for CAAPT are:

1. *TermRoot*: root form of a term or phrase.  
rdfs:subClassOf: skos:Concept,  
crm:E55\_Type, ontolex:LexicalEntry,  
culco:ContentiousIssue
2. *Guide*: written guidance on language use.  
rdfs:subClassOf: skos:ConceptScheme,  
crm:E32\_Authority\_Document,  
ontolex:ConceptSet
3. *TerminologyGuide*: written guidance for addressing problematic terminology created with the intention of assisting the work of or education around reparative description.  
rdfs:subClassOf: caapt:Guide,  
culco:ContentiousIssueScheme
4. *StyleGuide*: written guidance for language style and use.  
rdfs:subClassOf: caapt:Guide
5. *UseContext*: context bounding the meaning intended by the use of a term.  
rdfs:subClassOf: ontolex:LexicalSense
6. *Suggestion*: action to be considered or taken when a term is encountered.  
rdfs:subClassOf: culco:Suggestion,  
crm:E29\_Design\_or\_Procedure

These classes represent core concepts for this model: the term being considered (*TermRoot*), the ways it has been used (*UseContext*), the suggestions written in the guidelines (*Suggestion*), and the guidelines documents themselves (*Guide*, *TerminologyGuide*, and *StyleGuide*). The decision to propose these classes as subclasses of multiple ontologies works to build a bridge between these communities. This is reminiscent of the approach taken by Khan & Salgado (2021) in their work to forge connections between OntoLex, FRBRoo, and CIDOC CRM through the creation of two new classes that inherit from each of these ontologies.

The requirement for new properties for three of these classes also justifies the need to create these classes, as opposed to proposing the use of multiple instantiation in the representation of instances of these classes. For example, in the case of the proposed class *UseContext*, which inherits only from OntoLex (*LexicalSense*), the need to connect to different classes in specific ways—discussed below as subproperties for *ontolex:usage*—drove the need to declare a new class.

Inheriting from multiple classes can also introduce new nuances of meaning, such as in the case of the proposed class *Suggestion*: while CULCO’s *Suggestion* class is defined as “a suggestion gives recommendations on how to use a contentious term” (Nesterov et al., 2022), inheritance from CIDOC CRM’s class *E29\_Design\_or\_Procedure* introduces the specification that suggestions as they are understood in this context are, in fact, “documented plans for the execution of actions in order to achieve a result of a specific quality, form or contents” (Bekiari et al., 2022). Therefore, the *Suggestion* class here is a type of documented plan for how to address specific problematic terminology in a museum’s catalogue records.

Finally, the scope of these new classes is narrower in definition than the combination of meanings introduced by the classes from which they inherit. In the case of *Guide* and its subclasses *TerminologyGuide* and *StyleGuide*, the scope is defined more narrowly than for each of the classes from which *Guide* inherits, and terminology guides are differentiated from other forms of language guides, including writing style guides, in the source materials. This distinction is also reflected by *TerminologyGuide* inheriting from CULCO’s *ContentiousIssueScheme* as well as the *Guide* class, as it is only this specific kind of document that meets the additional criteria.

## 4.3 CAAPT Proposed Properties

The properties that are proposed, listed below in Table 2, are also connected to existing ontologies where possible: four are declared as subproperties of OntoLex’s *usage* predicate, and three as subproperties of CIDOC CRM’s *P69\_has\_association\_with* predicate.

Property	Draft scope note	Domain	Range	rdfs:subPropertyOf
caapt:used_where	Geographic location in which the use context existed or was/is relevant	caapt:UseContext	crm:E53_Place	ontolex:usage
caapt:used_when	Time period in which the use context existed or was/is relevant	caapt:UseContext	rdfs:literal	
caapt:used_why	Intended purpose of use of the term in a use context	caapt:UseContext	crm:E55_Type	
caapt:about_who	Group of persons intended to be described by a term in a use context	caapt:UseContext	crm:E74_Group	
caapt:preferred	Suggestion that is preferred	caapt:Suggestion	caapt:Suggestion	crm:P69_has_association_with
caapt:if_not_possible_use	Suggestion to be considered if not possible to use preferred suggestion	caapt:Suggestion	caapt:Suggestion	

caapt:use_along_with	Suggestion to be used concurrently	caapt:Suggestion	caapt:Suggestion	
caapt:suggests_replacement	Suggested replacement term used in the suggestion	caapt:Suggestion	caapt:TermRoot	---
caapt:suggests_amendment	Suggested amending term used in the suggestion	caapt:Suggestion	caapt:TermRoot	
caapt:encountered	Suggestion encountered in type of catalogue field	caapt:Suggestion	skos:Collection	

Table 2: CAAPT proposed properties

The four properties proposed as subproperties of usage all refine the notion of the original predicate—to “indicate[] usage conditions or pragmatic implications when using the lexical entry to refer to the given ontological meaning” (Cimiano et al., 2016)—to the specific kinds of use cases discussed in the source materials, where an analysis of the descriptions revealed four main considerations: where the use of the term took place, when the use of the term took place, who the term was intended to describe when it was being used, and the intended purpose or use of the term. These four properties therefore introduce these meanings to the relationships they represent, and narrow the scope of the range from *rdfs:Resource* to specific classes according to the needs of those relationships.

The three properties proposed as subproperties of *P69\_has\_association\_with* similarly refine the generic relationship between different instances of the *E29\_Design\_or\_Procedure* class, of which *Suggestion* is proposed as a subclass, in order to specify three ways in which *Suggestions* are related to each other in the source documentation: two of these are hierarchical, representing a preference order in the listed suggestions, and the third indicates when two suggestions should be used at the same time or as two parts of a larger remediative cataloguing actions. For example, adding contextualising text and adding a content warning are often recommended together, as explaining the use of the problematic term does not negate the need for a warning, and adding a warning to a record does not negate the need to add text explaining what the term means in the context of the record or why it was retained.

Three properties are not proposed as subproperties to existing LOD predicates: *suggests\_replacement*, *suggests\_amendment*, and *encountered*. The first two connect a *Suggestion* to a *TermRoot* and specify whether a term is suggested to be used to replace a term, or to be included alongside the existing term as an amendment of the text. The final property connects a *Suggestion* with the kind of field in which the term is located in the catalogue record. Initial values for instances of this class are “historical context” (e.g. a Title field) and “contemporary context” (e.g. the current display label text for the object’s online collection page) as this is the language used in the

source materials when suggestions are made according to the location of the term in the record. This is an important element to consider as different suggestions are made depending on what kind of field the term appears in.

## 5. Conclusions

A knowledge graph structured according to the ontology defined here has been populated with the contents of the V&A Terminology Guidance Document, resulting in an initial graph describing 328 Suggestions for 73 potentially problematic terms. The schema and contents have been reviewed by key stakeholders at the V&A. These validation meetings have been successful and the schema in both theory and practice has been well received. The primary suggestion to come out of the knowledge graph review meeting was to include additional *use\_along\_with* properties between a greater number of *Suggestions*: only relationships that had been made explicit in the source document had been included in the knowledge graph population, and this review meeting revealed that this kind of relationship was often implicit in the museum’s documentation.

Next steps will be to consider two additional terminology guidance documents for inclusion: the Cultural Heritage Terminology Network Glossary and the glossary section of the *Words Matter* publication (Chew, n.d.; Tropenmuseum, 2018). Integrating these sources will validate the schema as being generalizable beyond the sole context of the V&A, as well as produce a knowledge graph that will begin to allow for inter-institutional comparisons of terms and suggestions. Following this, reconciliation with LOD vocabularies—namely the Getty Art & Architecture Thesaurus and the Homosaurus vocabulary—will take place, as connecting to these two resources will demonstrate integration with a vocabulary that is commonly used in the cultural heritage domain (Getty Art & Architecture Thesaurus) and a community-developed vocabulary that is already working in the space of critical cataloguing (Homosaurus).<sup>2</sup>

The steps taken thus far have built a solid foundation for this work to proceed. Initial validation of the schema and knowledge graph have been successful, and further feedback alongside the integration and reconciliation work will inform future developments.

<sup>2</sup> Getty Art & Architecture Thesaurus: <https://vocab.getty.edu/aat/>; Homosaurus: <https://homosaurus.org>

## 6. Acknowledgements

This research is taking place as part of the AHRC-funded Collaborative Doctoral Partnership scheme grant AH/X004775/1.

## 7. Bibliographic References

- Bekiari, C., Bruseker, G., Canning, E., Doerr, M., Ore, C.-E., Stead, S., & Velios, A. (2022). *CIDOC Conceptual Reference Model, Version 7.1.2*. <https://cidoc-crm.org/Version/version-7.1.2>
- Chew, C. (n.d.). Welcome to CHTNUK. *Cultural Heritage Terminology Network*. Retrieved April 28, 2023, from <https://culturalheritageterminology.co.uk/>
- Chew, C. (2022, May 27). Inclusive Terminology for the Heritage Sector. *ARLIS UK & Ireland*. [https://www.youtube.com/watch?v=bb2hMcKz\\_aY](https://www.youtube.com/watch?v=bb2hMcKz_aY)
- Chew, C. (2023). Decolonising Description: Addressing Discriminatory Language in Scottish Public Heritage and Beyond. *Journal of Irish and Scottish Studies*, 11(1), Article 1. <https://doi.org/10.57132/jiss.213>
- Cimiano, P., McCrae, J. P., & Buitelaar, P. (Eds.). (2016). *Lexicon Model for Ontologies: Community Report*. <https://www.w3.org/2016/05/ontolex/>
- Cress, L. (2021, March 25). An Intern's Investigation on Decolonizing Archival Descriptions and Legacy Metadata. *Bitstreams: The Duke University Libraries Digital Collections Blog*. <https://blogs.library.duke.edu/bitstreams/2021/03/25/an-interns-investigation-on-decolonizing-archival-descriptions-and-legacy-metadata/>
- Dalal-Clayton, A., & Rutherford, A. (n.d.). Against a New Orthodoxy: Decolonised "Objectivity" in the Cataloguing and Description of Artworks. *Paul Mellon Centre Photographic Archive*. <https://photoarchive.paul-mellon-centre.ac.uk/groups/against-a-new-orthodoxy>
- Doerr, M., Light, R., & Hiebel, G. (2020). *Implementing the CIDOC Conceptual Reference Model in RDF*. <https://www.cidoc-crm.org/sites/default/files/issue%20443%20-%20implementing%20CIDOC%20CRM%20in%20RDF%20v1.1.pdf>
- Haraway, D. (1988). Situated Knowledges: The Science Question in Feminism and the Privilege of Partial Perspective. *Feminist Studies*, 14(3), 575–599.
- Khan, F., & et al. (2021). When Linguistics Meets Web Technologies. Recent advances in Modelling Linguistic Linked Open Data. *Semantic Web*, 0(0), 1–62.
- Khan, F., & Salgado, A. (2021). Modelling Lexicographic Resources using CIDOC-CRM, FRBRoo and Ontolex-Lemon. *SWODCH: Semantic Web and Ontology Design for Cultural Heritage*.
- Lawther, K. (2021, July 7). Decolonising the Database (Part 1). *Acid Free Blog*. <http://acidfreeblog.com/documentation/decolonising-the-database/>
- Miles, A., & Bechhofer, S. (2009). *SKOS Simple Knowledge Organization System*. <https://www.w3.org/2009/08/skos-reference/skos.html>
- Muñoz, G. (2021, April 22). Reframing Reparative Description Initiatives through Critical Race Theory and Black Feminism. *Descriptive Notes*. <https://saadescription.wordpress.com/2021/04/22/reframing-reparative-description-initiatives-through-critical-race-theory-and-black-feminism/>
- Museums Association. (2020). Our statement on Decolonisation. *Museums Association*. <https://www.museumsassociation.org/campaigns/decolonising-museums/our-statement-on-decolonisation/>
- Museums Association. (2021). Supporting Decolonisation in Museums. *Museums Association*. <https://www.museumsassociation.org/campaigns/decolonising-museums/supporting-decolonisation-in-museums/>
- Nesterov, A., Hollink, L., Van Erp, M., & Van Ossenbruggen, J. (2022). *CULCO: Concept Scheme for contentious terminology*. <https://cultural-ai.github.io/wordsmatter/>
- Nesterov, A., Hollink, L., van Erp, M., & van Ossenbruggen, J. (2023). A Knowledge Graph of Contentious Terminology for Inclusive Representation of Cultural Heritage. In C. Pesquita, E. Jimenez-Ruiz, J. McCusker, D. Faria, M. Dragoni, A. Dimou, R. Troncy, & S. Hertling (Eds.), *The Semantic Web* (pp. 502–519). Springer Nature Switzerland. [https://doi.org/10.1007/978-3-031-33455-9\\_30](https://doi.org/10.1007/978-3-031-33455-9_30)
- Ortolja-Baird, A., & Nyhan, J. (2022). 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. *Digital Scholarship in the Humanities*, 37(3), 844–867. <https://doi.org/10.1093/lc/fqab065>
- Rutherford, A. (2021a, March 24). Documentation and Decolonisation. *Museum - Data - Laundry*. <https://museumdatalaundry.com/2021/03/24/documentation-and-decolonisation/>
- Rutherford, A. (2021b, July 7). Decolonising the Database event, Centre for Design History, University of Brighton, July 5th 2021. *Museum - Data - Laundry*. <https://museumdatalaundry.com/2021/07/07/decolonising-the-database-event-centre-for-design-history-university-of-brighton-july-5th-2021/>
- Rutherford, A. (2022, June 10). Provisional semantics, projects and positionality. *ARLIS Cataloguing and Classification Ethics Series 2022*. [https://www.youtube.com/watch?v=tjTyi\\_WViSI](https://www.youtube.com/watch?v=tjTyi_WViSI)
- Tropenmuseum. (2018). *Words Matter: An Unfinished Guide to Word Choices in the Cultural Sector*. <http://www.tropenmuseum.nl/en/about-tropenmuseum/words-matter-publication>
- Victoria and Albert Museum. (2023). Size of the V&A Collections on 31 March 2023. *Victoria and Albert Museum*. <https://vanda-production-assets.s3.amazonaws.com/2023/05/18/12/15/18/adca7f35-1299-48cb-be74-4aa07151525e/V&A%20Size%20of%20Collections%20Statement%20March%202023.pdf>
- Watson, B. M. (2023). Critiquing the Machine: The Critical Cataloging Database. *TCB: Technical Services in Religion & Theology*, 31(1), Article 1. <https://doi.org/10.31046/tcb.v31i1.3216>