Modelling museum critical cataloguing practices using CIDOC CRM and OntoLex

Erin Canning, DPhil student, University of Oxford
Department of Engineering Science,
Oxford e-Research Centre
Co-supervised by the Victoria and Albert Museum

61st joint meeting of the CIDOC CRM SIG, 16 Oct. 2025

Problematic terminology and critical cataloguing

Identifying problematic terminology

- Slurs, derogatory language
- Euphemisms, veiled slurs
- Objectification, dehumanisation
- Erasure
- Laudatory terms, evaluative language
- Colonial and incorrect names of peoples, places, and types of objects
- Framing and focus: what is prioritized vs. minimized or dismissed

Barriers to the work

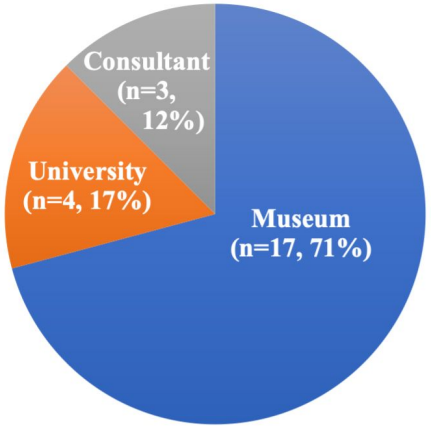
- Determining “where to start” - there is little domain-wide guidance on what to potentially look for
- Different institutions will require different levels of granularity on different issues
- Different institutions may be interested in different kinds of problematic language, or not agree on what terms are problematic
- Why do this? What comes next? How to proceed once instances have been found?

Proposing a linked data solution

- Address issues in critical cataloguing related to the need to develop domain-wide guidance about what to look for and possible remediative actions to take through the use of linked open data
- Share across institutions:
  - What is being looked for
  - How it is being identified
  - What is being recommended when instances are encountered
- Connect to additional lexicographical resources

Eliciting knowledge part 1: interviews

45-minute semi-structured interviews with 24 critical cataloguers working in or with museums conducted December 2023 - July 2024.



Representation of participants by institution type

4 areas of domain requirements covering 26 requirements:

- Reflexivity
- Complexity
- Accommodativeness
- Sustainability

In progress draft document: https://eecanning.github.io/caapt/img/dh2025.pdf

Eliciting knowledge part 2: competency questions

5 prompt questions

1. What questions come up for you when you are thinking about reviewing catalogue data for problematic terminology?
2. What questions come up for you when you encounter what you think might be an instance of problematic terminology in catalogue data?
3. What questions come up for you when you seek to engage with problematic terminology in cataloguing data?
4. What information do you look for when trying to make a decision about what to do when you encounter problematic terminology in catalogues?
5. Where do you go to try to find information or answers to the questions that you have?

43 competency questions

Examples:

CQ 7. What explanations of different meanings for this term over place and time are available for this term?

CQ 20. What are the suggestion details for the different possible uses of this term?

CQ 23. What are the ”preferred” terms suggested for this term?

CQ 37. Who has reviewed records that contained this term?

CQ 41. Which meaning of a term is most frequently used when the term is found in a record?

Eliciting knowledge part 3: terminology guidance documents

GLOSSARY OF TERMS

Aboriginal

HISTORY, USE & POSSIBLE SENSITIVITIES

The term "Aboriginal" stems from the Latin words *ab* and *origine*, meaning "from the beginning", and is a commonly used word, including in museum databases. "Aboriginal" describes the original inhabitants of a place and is primarily used to refer to Indigenous peoples in Australia and Canada (see also "Indigenous"). However, in Canada, there has been a recent preference for the use of Indigenous.

The term does not adequately describe the complexity and diversity of Indigenous peoples. Many Indigenous peoples in both countries do not like to be referred to as "Aboriginal", preferring to emphasize other markers of their identity such as language, land and clan relationships. In Canada the term came into general use in the 1980s, when a legal and overarching term was sought to collectively describe the existing—but controversial—categories "Indian," "Inuit" and "Métis" (see also "Indian" and "Eskimo").

SUGGESTIONS

- Adopt the terminology used and accepted as respectful by the people themselves.
- In cases where it is not possible to avoid generalization, for example, should it be impossible to find out the specific group from which a person/object comes, then use "Aboriginal" as a compound noun, always with capital "A": *Aboriginal people(s), Aboriginal Australians.*
- In Canada, certain First Nations groups in Ontario prefer "Indigenous," i.e. "Indigenous people(s)." "First Nation(s) people(s)" may also be acceptable.
- In Australia, "Aboriginal" and "Torres Strait Islander peoples" is in most situations appropriate.

Figure 1

Centering Indigenous Epistemologies

Excerpt from the finding aid of the Stephen Beauregard Weeks Papers (collection number 00762)

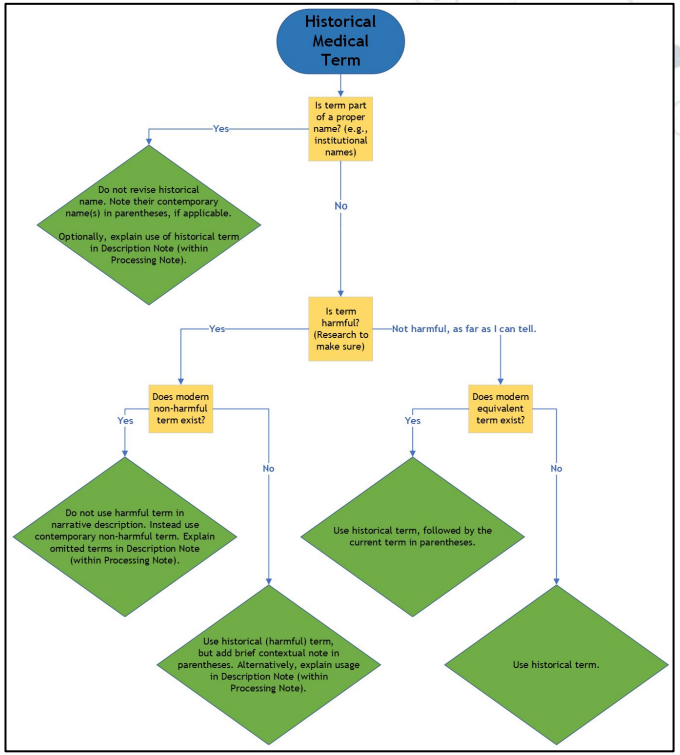
Legacy Description

Stephen Beauregard Weeks (1865-1918) was a North Carolina educator and historian. He was superintendent of an Indian school in Arizona. The collection contains papers and volumes related primarily to southern education and religion, compiled or created by Stephen B. Weeks, North Carolina educator and historian. Included are his correspondence about North Carolina historical matters, 1897-1913, and 75 items pertaining to a dispute at an Indian school in Arizona, 1903-1905, of which Weeks was superintendent.

Conscious Editing

Stephen Beauregard Weeks (1865-1918) was a white North Carolina educator, historian, and superintendent of San Carlos Boarding School, what was then called an "Indian school," for Apache Indians in San Carlos, Arizona. The collection consists of personal, family, and professional correspondence, papers, diaries, and other volumes. Topics include the history of education in southern states, religion, a dispute at the San Carlos Boarding School, North Carolina history and biography, the formation of the Southern Historical Association, southern Quakers, and slavery, and George Moses Horton, an African American poet who was enslaved in Chatham County, N.C., during the early 1800s.

The reductive use of the term "Indian School" is replaced with the full name of the school. The term "Indian School" is kept in the new description, but it is placed in quotation marks to indicate that this description is an outdated, Western-oriented historical term that does not accurately or respectfully describe Indigenous history and experience. The relevant Indigenous tribe associated with the school is also named in the remediated description.



National Museums of World Culture (2018)
Wilson Special Collections Library, University of North Carolina at Chapel Hill (2022)
Francis A. Countway Library of Medicine, Harvard University (2020)

Eliciting knowledge part 3: terminology guidance documents

GLOSSARY OF TERMS

Aboriginal

HISTORY, USE & POSSIBLE SENSITIVITIES

The term “Aboriginal” stems from the Latin words ab and origine, meaning “from the beginning”, and is a commonly used word, including in museum databases. “Aboriginal” describes the original inhabitants of a place and is primarily used to refer to Indigenous peoples in Australia and Canada (see also “Indigenous”). However, in Canada, there has been a recent preference for the use of Indigenous.
The term does not adequately describe the complexity and diversity of Indigenous peoples. Many Indigenous peoples in both countries do not like to be referred to as “Aboriginal”, preferring to emphasize other markers of their identity such as language, land and clan relationships.
In Canada the term came into general use in the 1980s, when a legal and overarching term was sought to collectively describe the existing—but controversial—categories “Indian,” “Inuit” and “Métis” (see also “Indian” and “Eskimo”).

SUGGESTIONS

- Adopt the terminology used and accepted as respectful by the people themselves.
- In cases where it is not possible to avoid generalization, for example, should it be impossible to find out the specific group from which a person/object comes, then use “Aboriginal” as a compound noun, always with capital “A”: Aboriginal people(s), Aboriginal Australians.
- In Canada, certain First Nations groups in Ontario prefer “Indigenous,” i.e. “Indigenous people(s).” “First Nation(s) people(s)” may also be acceptable.
- In Australia, “Aboriginal” and “Torres Strait Islander peoples” is in most situations appropriate.

1

Term	Sub-term	See also	Notes on the term	Guidance on the term
Aborigine(s) / Aboriginals		Indigenous	**Under review.** The word 'Aboriginal' describes the original inhabitants of a place. It is mostly used to refer to Indigenous Australians and Indigenous people in Canada (see also 'Indigenous'). However, the term does not adequately describe the complexity and diversity of indigenous peoples, and many indigenous people in both countries do not like to be referred to as 'Aboriginal', preferring instead to emphasise other markers of their identity such as language, land and clan relationships.	**Under review. Consult Interpretation Department.** Where possible, reword to be more specific. If being specific is not possible, for Australia change to 'Indigenous Australians'. For Canada, change to 'First Nation(s)' or 'First Nations people(s)' to refer to 'Indigenous people in Canada' in general, except for the Inuit and Métis. these communities are not First Nations and their individual names should be used. In cases where it is not possible to avoid or change the term, use 'Aboriginal' as a compound noun, always with capital 'A': e.g., Aboriginal people(s), Aboriginal Australians.

2

Term	Contextual note	Time period/ Region note	Reference
Aboriginal/s	The term "Aboriginal" (with a capital "A") is most often used to refer to Indigenous peoples in Australia and Canada. In Canada, the term Indigenous (with a capital "I") is preferred. Indigenous communities in both Australia and Canada often prefer to emphasise other markers of their identities, such as language, land, and/or clan. "First Nations" is sometimes a preferred term. The term "aboriginal" may be used to refer to flora/fauna that lived or existed in a place since the earliest known time. ***Reminder: capitalise "Aboriginal" when referring to Indigenous peoples in Australia and Canada***	Canada, Australia	Tropenmuseum - Words Matter

3

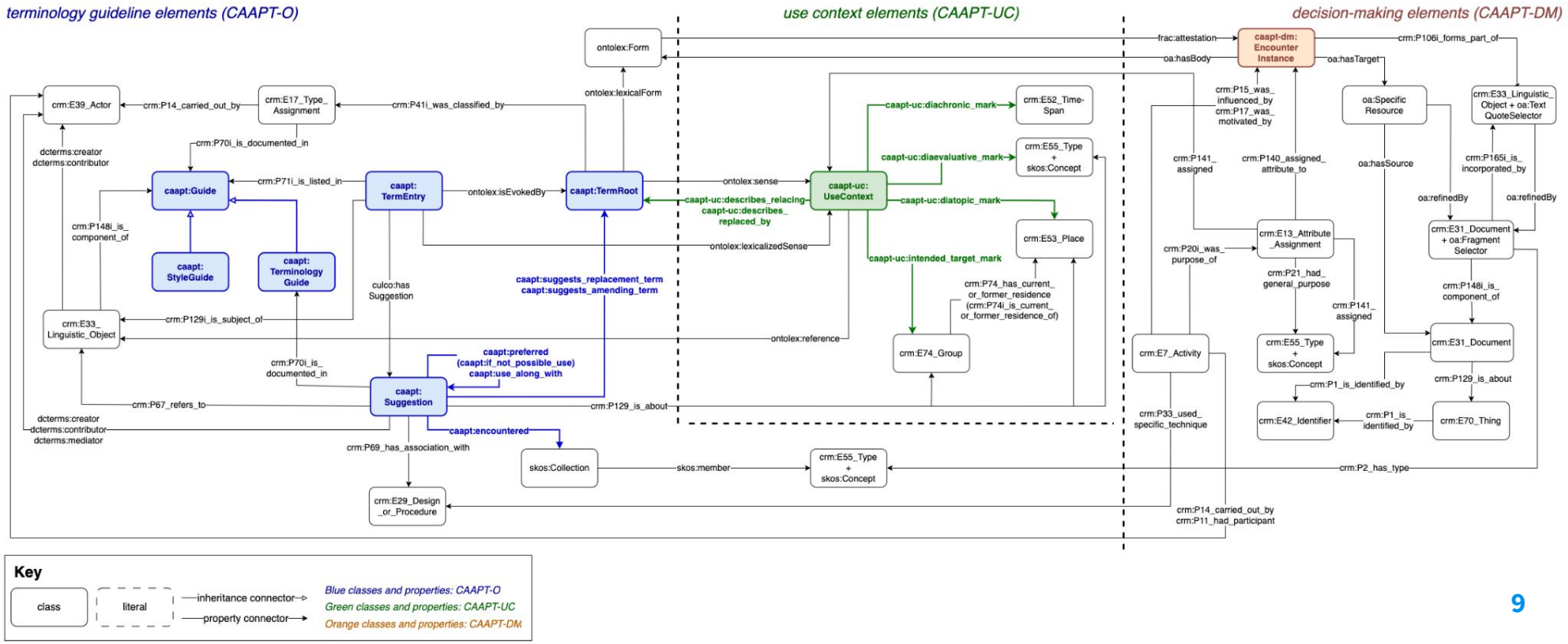
(1): V&A Museum Terminology Guidance Document; (2): Cultural Heritage Terminology Network Inclusive Terminology Glossary; (3): Words Matter

Critical cataloguing and linked open data

Ontology	Description	Use	Limitation
CIDOC CRM	A stable, widely used ontology for the representation of cultural heritage data	• Extensive coverage of cultural heritage data • Event-centric approach provides covers actions	• Insufficient coverage of linguistic elements
OntoLex	A stable, widely used ontology for the representation of linguistic data, namely lexica and dictionaries	• Extensive coverage of linguistic data elements • Extension for diachronic elements	• Introduces greater level of complexity in structuring of linguistic elements than is required for this use case • No coverage of actions
SKOS	A stable, widely used ontology for the representation of thesauri, terminology lists, and controlled vocabularies	• Covers basic thesaurus structuring • Used by related projects	• Only for relating terms and labels to each other, no coverage of diachronic and sense elements, actions, etc.
CULCO	A project ontology developed to describe the glossary section of the “Words Matter” publication	• Direct representation of domain • Used by DE-BIAS Project	• Insufficient granularity of representation • No coverage of actions

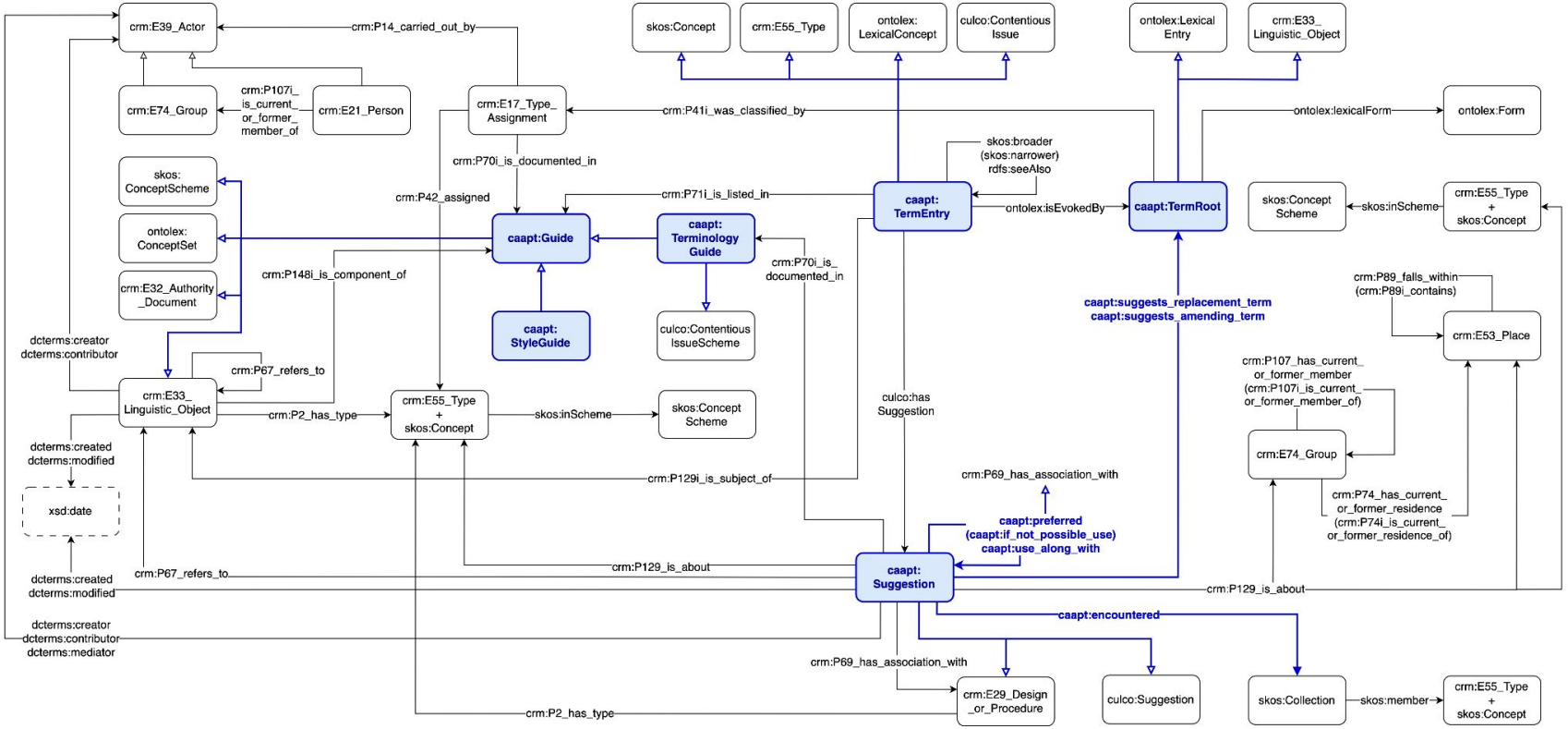
https://w3id.org/caapt

CAAPT ontologies overview



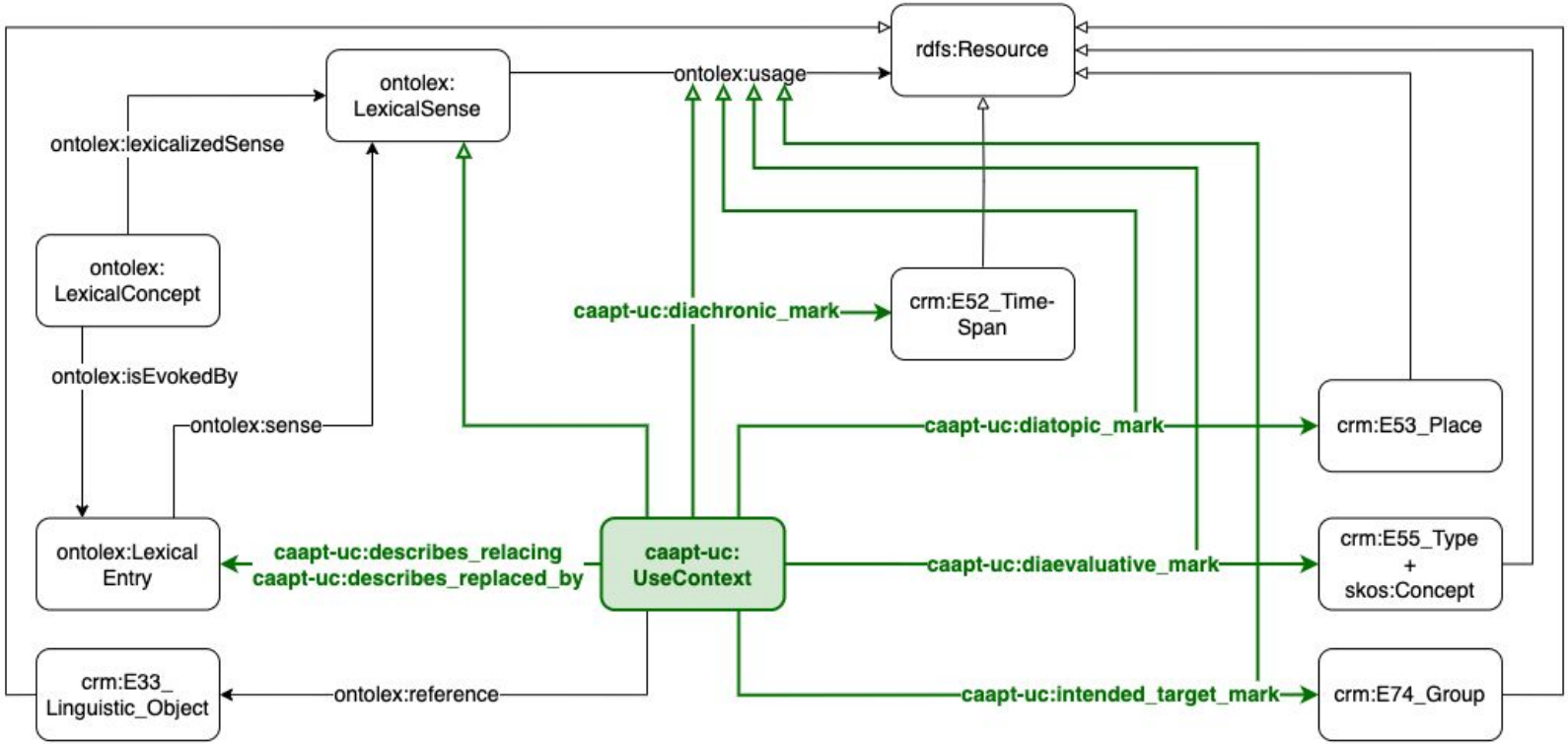
https://w3id.org/caapt

CAAPT-O: critical cataloguing guidelines



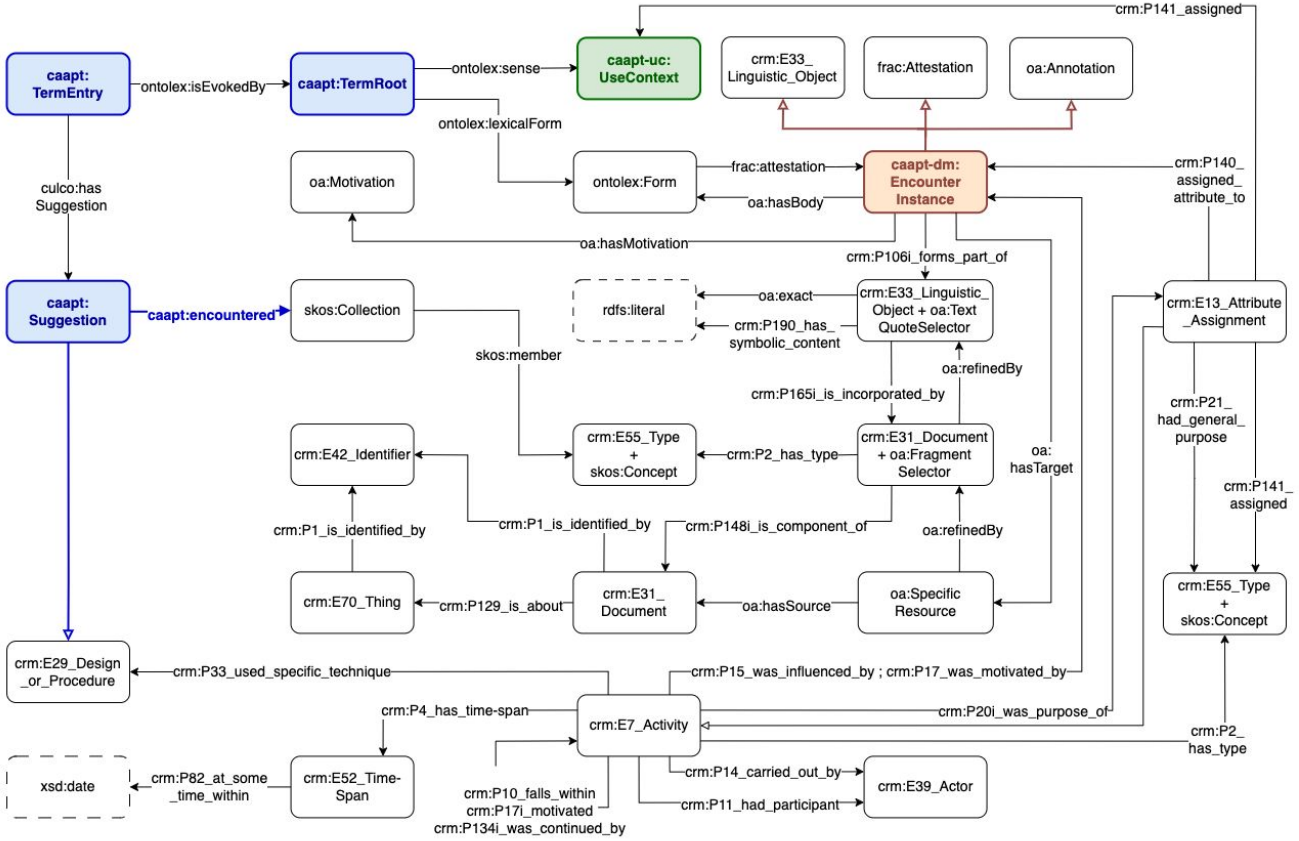
https://w3id.org/caapt

CAAPT-UC: terminology use contexts



https://w3id.org/caapt

CAAPT-DM: critical cataloguing decision-making



CAAPT knowledge graph (567,410 triples)

Section (number of…)	V&A Museum TDG	V&A LGBTQ+ TGD	Words Matter	Chew (CHTN) ITG	DE-BIAS Vocab.	Total
Entries listed in source documents	76	42	55	734	537	1,216
Terms listed or referenced in source documents	123	44	108	1,063	1,006	2,029
Terms classified as “to search for”	78	42	63	1,005	702	1,633
Terms classified as “suggested as preferred”	46	37	52	88	307	471
Use contexts	139	66	137	914	283	1,391
Suggestions	324	74	106	276	473	1,253

Overlapping contents:

No. of sources	Entries	Roots
4	24	31
3	29	36
2	98	150
1	1,065	1,812
(Total)	(1,216)	(2,029)

Annotated instances:

V&A annotations	11,246
V&A TDF decisions	257
DE-BIAS	662
Total	12,165

Leveraging this representation

CAAPT critical cataloguing decision-making support

The CAAPT critical cataloguing decision-making support tool helps you identify potentially problematic terms in your collection records and suggests next steps to take, using established sources and predictive models. You can upload records, identify the meaning of terms in context, and receive suggestions for potential next steps from a range of sources. At the end, you will be able to download your work as a CSV file for future use.

Upload records

Upload a CSV file with your records. It must contain one column named 'ID' with object or record identifiers; one column named 'Link' containing the URL of the online collection page of the object/record (contents optional but column required); and one or more columns to be analyzed, named by their Record Field. There should be one row per object/record ID.

Upload your CSV file

Drag and drop file here
Limit 200MB per file • CSV

Browse files

Select vocabularies

Select one or more options from below. You can select by source, subject area, or specific terms. If you select by source and/or subject area, you cannot also select specific terms. If nothing is selected, everything will be used in the analysis.

By source (guideline document)

- V&A Museum - Terminology Guidelines
- V&A Museum - LGTBQ+ Guidelines
- Words Matter
- Chew Inclusive Terminology Glossary
- DE-BIAS Vocabulary

By subject area (language related to...)

Select Subject Areas

By specific term

Click checkbox to access a list of specific terms (count = 1,159)

Select Terms

Search and select terms

Search

- aakāne, aikane, moe aikāne, moeaikane
- abbie*, abbie, abe, abie
- Abbo, Abbos, Abo, Abos
- abeed*, abeed, abid
- able, able-bodied, able bodied, abled, differently able, specially able
- abnormal, abnormality, not normal, subnormal
- Abolition*, abolitionist, abolitionists
- abominable pollutions
- Aborigin*, Aboriginal, Aboriginals, Aborigine, Aborigines, aborigine, aborrigine, The Aboriginal people, The Aboriginals, The Aborigines

Select Subject Areas

By subject area (language related to...)

Choose a subject area

Sexuality and gender

- Sexuality and gender
- Ability and disability
- Indigenous histories and peoples
- Class and socioeconomic status
- Empire, imperialism, and colonialism
- Antisemitism
- Other

- Women's history

Confirm selection

Set preferences

Set your preferences for how the analysis will be conducted. Whole word matching ensures that only complete words are matched, not parts of words. Predictive models attempt to disambiguate term meanings and identify problematic uses of terms based on patterns in existing data.

Search for whole words only?

- Search for whole words only

Use predictive models?

- Yes

Analyse records

Press the Submit button to start the analysis. This may take a few minutes depending on the size of your dataset and the number of terms you have chosen to look for.

Submit

Leveraging this representation

Filter options

Filter by Object ID

Choose options

Filter by Field

Choose options

Filter by Term

Choose options

Sort options

Sort by

None

CAAPT critical cataloguing decision-making support

The CAAPT critical cataloguing decision-making support tool helps you identify potentially problematic terms in your collection records and suggests next steps to take, using established sources and predictive models. You can upload records, identify the meaning of terms in context, and receive suggestions for potential next steps from a range of sources. At the end, you will be able to download your work as a CSV file for future use.

1. Records, terms, and meanings

The first step is to identify the meaning of the term as it is being used in the record. This step is not required, but it will help the app provide more accurate suggestions. Additionally, you will be able to filter by identified meaning on the following page. At this time, you can also indicate whether or not you are able to make direct edits to the field of the record in question.

#	Object ID	Field	Term	Instance	Highlighted Text	Term Meaning	Direct edit?
1	Q640758	Physical description (EtC)	Queer	Queerinthian	...an, Doric, Ionic, Corinthian, and Composite) include comical titles such as 'Old Peerian of Alder-Manic' and '**Queerinthian** or queue de renard'. Expand / Collapse	Select meaning	☐
2	Q1632617	Public access description value	Queer	queerness	... the photographic studio. His pictures often take the form of portraiture, a medium through which he explores **queerness**, homoeroticism, and blackness. In 2004, Sepuya received a BFA from New York University's Tisch School of the ... Expand / Collapse	Select meaning	☐

3. Generate suggestions

Press the Submit button to generate suggestions for your records. This may take a few minutes depending on the size of your dataset and the number of instances you are reviewing.

Submit

Back to Top

Select the meaning for the word "Queer" that is taking place in this record. Clicking a row will show more details below. Scroll down after selecting a row.

Value(s)	Place(s)	Time(s)
☐ Particularly since the 1980s, 'queer' has served as an umbrella term for sexual interests and identities that challenge social norms for sexual behavior. The term is not only shorthand for 'LGBT' but also the full range of human sexuality, for example people with sexual fetishes or who practice polyamory. An umbrella term which embraces a spectrum of sexual preferences and identities, which can include lesbians, gay men, bisexual people, transgender people, intersex people, the radical sex communities, and many other people. It allows you to not strictly define yourself.		since the 1980s
☑ In the late 20th century, the word 'queer' became a political reclamation. 'Queer' has been reclaimed as a proud political and sexual identity from earlier etymological usages that meant strange, aslant or curious.		late 20th century
☐ It has historically been used as a slur or term of abuse. The term, however, is still used as a slur against people perceived to be sexually deviant.		historical

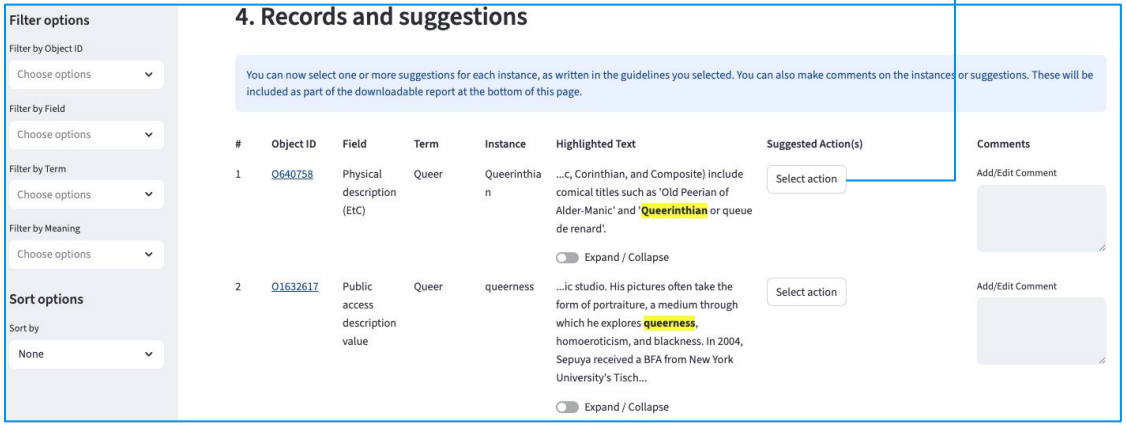
In the late 20th century, the word 'queer' became a political reclamation. 'Queer' has been reclaimed as a proud political and sexual identity from earlier etymological usages that meant strange, aslant or curious.

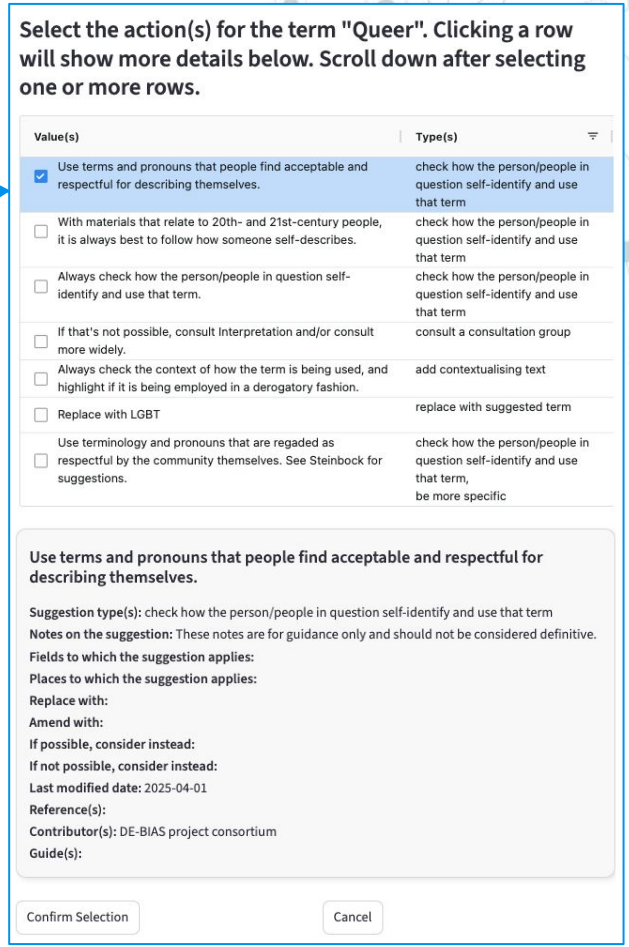
Note(s): These notes are for guidance only and should not be considered definitive.
Place(s):
Time(s): late 20th century
Intent(s): Harmful term reclaimed by community
Describes replacing:
Describes being replaced by:
Reference(s): Tropen Museum et al., eds., "Words Matter: An Unfinished Guide to Word Choices in the Cultural Sector," 2018, 134. https://www.materialculture.nl/sites/default/files/2018-08/words_matter.pdf.pdf.
Andrei Nesterov, Laura Hollink, Marieke van Erp, and Jacco van Ossenbruggen. (2023). cultural-ai/wordsmatter: Words Matter: a knowledge graph of contentious terms (v1.0.2) [Data set]. European Semantic Web Conference (ESWC), Hersonissos, Greece. Zenodo. https://doi.org/10.5281/zenodo.7713157 ; https://w3id.org/culco/wordsmatter/134.
Contributor(s): Victoria and Albert Museum National Museum for World Cultures
Guide(s): Victoria and Albert Museum LGBTQ+ Terminology Guide Words Matter: An Unfinished Guide to Word Choices in the Cultural Sector

Confirm Selection

Cancel

Leveraging this representation





Thank you! Questions?

Erin Canning | erin.canning@eng.ox.ac.uk

DPhil student, Oxford e-Research Centre, Dept. of Engineering Science, University of Oxford & Victoria and Albert Museum

Supervisors:
Prof. David De Roure[1], Prof. Sarah Ogilvie[2], Dr. Kevin Page[1]
Dr. Richard Espley[3], Pamela Young[3]
[1] University of Oxford, Department of Engineering Science
[2] University of Oxford, Department of Linguistics, Philology and Phonetics
[3] Victoria and Albert Museum

This research is funded by the UKRI-AHRC CDP scheme grant AH/X004775/1, “Applications of computational approaches in addressing problematic terminology in V&A Museum catalogues”