Title:

Linked Conservation Data: the adoption and use of vocabularies in the field of heritage conservation for publishing conservation records as Linked Data

Authors:

Athanasios Velios, University of the Arts London, Chelsea College of Arts, 16 John Islip Street, London SW1P 4JU, a.velios@arts.ac.uk

Kristen St.John, Stanford Libraries, 415 Broadway, Redwood City, CA 94063. kstjohn@stanford.edu

Linked Conservation Data: the adoption and use of vocabularies in the field of heritage conservation for publishing conservation records as Linked Data

Abstract

One of the fundamental roles of memory organisations is to safe-keep collections and this includes activities around their preservation and conservation. Conservators produce documentation records of their work to assist future interpretation of objects and to explain decision making for conservation. This documentation may exist as structured data or free text and in both cases they require vocabularies that can be understood widely in the domain. This paper describes a survey of conservation professionals which allowed us to compile the vocabularies used in the domain. It includes an analysis of the vocabularies with key findings: a) the overlapping terms with multiple definitions, b) the partial coverage of the domain which is lacking controlled vocabularies for condition types and treatment techniques and c) the limited formats in which vocabularies are published, making them difficult to use within Linked Data implementations. The paper then describes an approach to improve the vocabulary landscape in conservation by providing guidelines for encoding and aligning vocabularies as well as considering third party platforms for sharing vocabularies in a sustainable way. The paper concludes with a summary of our findings and recommendations.

Introduction

Heritage conservation is one of the main activities of memory organisations such as museums, libraries, galleries, archives but also archaeological sites and organisations which are responsible for built heritage. Conservators work to stabilize, repair, house, assess and describe collections from built heritage and historic sites to fine art to books in libraries. The overall

purpose of the domain is to ensure that cultural heritage is maintained through the safe-keeping of evidence that is required for our understanding of the context and history of heritage. Heritage conservation comprises practice, research and education on methods, materials and equipment employed when working on collections. Conservators are organised in several national and international bodies (e.g. the International Institute for Conservation of Historic and Artistic Works). These organizations have developed and maintain charters and codes of ethics to guide members in their practices and procedures. Examples are: “Our Code of Ethics” (https://www.culturalheritage.org/about-conservation/code-of-ethics) by the American Institute for Conservation and “Icon professional Standards and Judgment & Ethics” (https://icon.org.uk/about-us/standards-ethics/icon-professional-standards) by the Institute of Conservation in the UK.

Conservators are required by those codes of ethics to document treatments so that their actions for preserving collections are understood by other collection custodians and available to future conservators who may have to revisit a treatment or continue working on a project.

While documentation is required, the form it takes and how it is preserved are largely not prescribed. Specific practice is determined by conservators, their departments, or their institutional mandate and/or resources. Documentation is retained permanently, but may be stored locally either in file cabinets, servers or stand alone databases. In larger institutions, documentation might be tied in with collection management systems.

Common types of documentation generated by conservators and conservations scientists include:

- Treatment proposals and reports. These generally include identifying information of the item, a physical description, the current condition, testing results, and proposal(s) to stabilize or remediate damage to the item. A report would include actions undertaken to carry out the treatment. This type of documentation generally applies to a single item or

a small set of similar items. Sometimes more generalized proposals/reports may apply to larger groups of materials.

- Surveys and assessments. This type of documentation contains many of the initial components of treatment reports: identifying information, description, condition, and perhaps proposals for needed treatment. They tend to apply to larger numbers of items that may be surveyed individually or through a representative sample. Rationale for carrying out surveys or assessments can vary from wanting a broad overview of collection condition to specific assessment prior to exhibit or loan.
- Scientific data: Scientific data gathered due to or through Conservation treatment can be included in other types of documentation and/or may result in stand-alone documentation as well.
- Environmental or collection-wide documentation: Conservators and other collection professionals gather information about overall conditions of collection storage or exhibit. This information may be tied to a physical space and/or a group of materials.

Within these broad classes of documentation type, the representation of information as structured data can vary. Free text is a common feature of most types of documentation. In many cases instead of employing free-text, conservators encode their records in forms with checkboxes and pre-selected options that are then stored in databases. Condition surveys and routine repair work are often recorded using structured forms. Often free-text conservation reports are findable through a set of metadata which point to the full text document, i.e. a set of structured data acting as an index to the free-text document.

In both free-text and structured records the issue of ambiguous terminology is evident. The terminology used in conservation practice can stem from different craft or art practice with variation across regional and national traditions. It is not unusual for the same word to be used

differently within or between specializations (for example, the word ‘textblock’ is used to mean both a) the area of a printed page where the letters are located as a block, and b) the volume of all leaves of the book) or for different words to be used for the same concept (the ‘endbands’ of a book are also called ‘headbands’ by different conservators). Synonyms introduce ambiguity in records and context-specific use of words limits understanding to the audience within the same context. Structured records may be in a disadvantage to free-text where the syntax of a sentence could be useful in understanding the correct context of a word. It is the use of terminology specifically within structured records that we are considering in this article.

Linked Conservation Data

Integration of conservation documentation records can help the development of the domain and profession. The benefit of sharing conservation records has been discussed in a recent article (Velios 2021). This is mainly the availability of larger and more representative samples of observations which can lead to better conclusions on a) historic materials and techniques, b) conservation materials and analysis, c) correlation of damage types and environment and d) trends in conservation work (e.g. speed of work or type of conservation work executed).

Addressing this need for greater sharing and integration of conservation records, Linked Conservation Data (LCD) is a project funded by the Arts and Humanities Research Council in the United Kingdom to explore the potential of Linked Data technologies for sharing and integrating conservation documentation records. The project has established a consortium of partners primarily from the UK and the US with outreach activities introducing Linked Data to conservators as an opportunity to think about the role and structure of conservation documentation records.

Project Objectives

Two main areas of research are underway: a) modelling conservation records: where the consortium is exploring the capacity of the CIDOC-CRM ontology (http://www.cidoc-crm.org/) for expressing the relationships between concepts referred to in conservation documentation, and b) conservation terminology: developing a strategy to utilize and cope with a diverse array of vocabularies has been critical which is what we are reporting on here. Our requirements when working with vocabularies are summarised here:

1. That conservators are not required to switch to a different vocabulary from what they are used to. Conservators, like other professionals, have established methods and naming conventions when undertaking their work for clarity within their working environments. It is not desirable to alter these naming conventions as this causes lack of communication and disruption within conservation studios. This is why we require that conservators are able to maintain their current vocabulary.

2. That, while recognising the centrality of the English language in the Linked Conservation Data project, any recommendations would also be applicable to other languages. This aligns with requirement 1 in that conservators who produce documentation records in languages other than English, wish to continue their practice while receiving the same benefits as English speaking colleagues.

3. That searching records using one vocabulary would return results from records using other vocabularies. Following requirements 1 and 2 which are primarily for creating new records, conservators wish to maintain the use of their preferred vocabularies, by preference or by necessity (for example if they do not speak a foreign to them language), also for querying existing records, even though those records have been produced using a different vocabulary or language.

We began our investigation by mapping the landscape of vocabularies in conservation.

Conservation vocabularies

In order to understand the way that conservators use vocabularies we undertook a survey exercise. A questionnaire was sent out to conservation professionals to identify which vocabularies, glossaries, thesauri or other sources of terms they used in their work, if they use them at all. The set of questions included in the survey are mentioned in the Appendix. We received 27 responses mainly from conservators, but also from professionals in relevant fields. We recognise that the sample may be small and biased towards conservators who are interested in issues around conservation terminology but there are no studies with a larger sample in the field. The main findings of the survey are:

1. That conservators generate three main types of documentation which depends on terminology: a) documentation describing treatment undertaken on an object, b) documentation from surveys & assessments, typical of the condition of collections, and c) conservation science data and analysis.
2. Almost ⅔ of the respondents choose terms primarily guided by previous experience and current knowledge.
3. About 40% of the respondents refer to published vocabulary resources, such as glossaries and thesauri, when choosing their terminology.
4. For about 40% of the respondents, a large part of the documentation records produced utilize in-house glossaries or word lists.

The survey ended by asking respondents to identify which published vocabularies they use. Based on this information we compiled a list of vocabularies used in conservation which formed

the basis of our discussion in a follow up workshop. The list is available on the project website: https://www.ligatus.org.uk/lcd/controlled-vocabularies . This includes a description of the structure of each vocabulary as well as its readiness and additional work needed for publication as Linked Data. For example a frequently used vocabulary among conservators is the Conservation & Art Materials Encyclopedia Online (CAMEO – http://cameo.mfa.org). As part of the analysis of the vocabularies identified by the survey, CAMEO includes terminology around materials, techniques and condition types. It also includes references to concepts described in the vocabulary, illustrative images, alternative labels and unique identifiers which may act as URIs in a Linked Data context although they are dependent on labels which are not considered stable enough. Further work is needed for transcribing and encoding the vocabulary before it can be shared as Linked Data. Similar assessments took place for each of the vocabularies and the results are summarised next.

Challenges with conservation vocabularies

We examined the reported vocabularies and considered how easily they can be used within Linked Data implementations. We also examined them in relation to coverage of the conservation domain as well as their internal structure, i.e. on whether they employ hierarchies or facets. We discovered that:

1. Vocabularies tend to focus on a conservation sub-domain rather than attempt to cover the whole domain. Book and paper appeared to be the subject of a significant number of vocabularies followed by architecture and built heritage with many other sub-domains also represented. This creates a patchwork of coverage for the conservation domain which is not consistent.
2. About a quarter of vocabularies feature hierarchical arrangement of terms. These vocabularies are often broad, covering areas also outside conservation. Less than 10%

of the vocabularies included terms which would allow the development of a hierarchical arrangement of terms using broader/narrower relationships. A possible solution for overcoming this limitation is through hierarchical alignment to external thesauri as explained in the section about Application. This indicates that currently there are limited applications of reasoning based on broader/narrower relationships in the domain of conservation. About a fifth of the vocabularies include associative relationships.

3 The majority of vocabularies include materials that objects are made of and about half of them included the historic techniques used. This reflects the fact that many of these vocabularies were produced by professionals in allied fields with particular interest in cataloguing and not for use by conservators. Conservation treatment and condition types including damage were represented in less than a quarter of the vocabularies with limited coverage. This indicates a need for further development of vocabularies for treatment and condition.

4 About a quarter of vocabularies provide URIs suitable for Linked Data applications. About a quarter of the vocabularies are available as structured data in a format which can be machine-processed. Some of them are semi-structured and many would require transcription from text documents. This indicates that an exercise for processing and hosting conservation vocabularies as Linked Data is necessary. An associated issue is that vocabularies do not provide identifiers and their use in documentation systems is based on labels.

The review and analysis of the vocabularies identified during the survey was done in advance and during a workshop where the main challenges were discussed alongside proposals on future work.

Vocabularies workshop

The first workshop of the Linked Conservation Data project was held over a day and a half in June 2019 at Stanford University. 22 participants, both academics and professionals, with a diverse set of backgrounds including conservation, conservation science, information science and computer science engaged with various aspects of Linked Data and vocabularies. A large part of the workshop was devoted in the evaluation of the current state of vocabularies with the main findings described in the previous section. The participants also looked at technical challenges of alignment and searching across disparate vocabularies which was one of our requirements. Under the guidance of Prof. Marcia Zeng we explored the role that thesauri can play and began to develop a possible course of action.

We considered the following problems as identified during the workshop and survey: a) overlapping definitions of terms from different resources, b) lack of vocabulary readiness for publication as Linked Data and c) lack of coverage.

Overlapping definitions

Terms and concepts used in conservation are often defined in different vocabularies. These include synonyms. Requiring the use of one vocabulary instead of another would have resolved this problem but it would have contradicted one of our requirements which was to allow conservators to use their preferred vocabularies. To resolve this problem we agreed to align vocabularies by identifying matching concepts. Communities maintaining specialist vocabularies may choose which of them are useful to align directly so that there is maximum benefit when searching datasets described with these vocabularies. For example the Language of Bindings (LoB) Thesaurus (we can consider it as a source vocabulary) and the Controlled Vocabulary with binding terms from the Rare Books and Manuscripts Section of the American Library

Association (target vocabulary) are both popular resources used to describe historic bookbinding structures. Establishing a direct alignment between them would perhaps satisfy the requirements for search of a wide community of bookbinding experts. However, the introduction of a third bookbinding vocabulary would multiply the amount of effort needed for alignment and scaling this to more vocabularies would require resources which are not available.

In order to make the alignment process efficient we reviewed the conservation vocabularies and identified those which could act as hub vocabularies so that we identify one target vocabulary for all other to link to. The Getty Arts & Architecture Thesaurus (AAT) was the most appropriate candidate for a hub vocabulary. While the AAT has limited coverage for conservation, it is the most extensive vocabulary in our list with capacity for expansion. One of our principles is that all source vocabularies, including local lists where applicable should be aligned with the AAT first and then with other vocabularies of interest if there are resources available.

Vocabularies as Linked Data

As mentioned in the section about Characteristics of Conservation Vocabularies, many of them are not published as structured data and do not provide identifiers for concepts and labels therefore making them unsuitable for Linked Data applications. This is primarily because vocabularies are maintained by conservation professionals who are not necessarily aware of best practices in information science and knowledge organisation and do not have the necessary resources to process and publish their vocabularies as Linked Data. To address this problem we established a repository where conservation vocabularies can be hosted long term. This provides a basic infrastructure for conservation vocabularies which may not be available to the teams that maintain them. The repository is hosted on GitHub and therefore none of the consortium members are burdened with maintaining infrastructure long term (https://github.com/linked-conservation-data/conservation-vocabularies). Contributions by

members or third parties can be done using well-established procedures within GitHub. Moderation and approval of such contributions require time from repository maintainers and we believe this is the minimum effort required for the repository to remain functional. This effort is shared among consortium members regardless of availability of project-based funding. The idea being that among a group of consortium members there will always be some resource available to approve new contributions to the repository.

The repository stores vocabulary data and does not feature any user-friendly tools for searching scope notes and discovering alignments of terms. Maintaining such a system long term requires resources which are difficult for an institution to provide without continuous funding. We focused on the development of the data repository which will enable funding applications to establish a user-friendly system for interacting with the contributed vocabularies.

Through an iterative design process among consortium members two diagrammatic workflows were developed which assist vocabulary maintainers with publishing vocabularies as Linked Data and aligning these to other vocabularies. These are shown in figures 1 and 2. These workflows are enriched with explanation notes and implementation suggestions which point vocabulary maintainers with no experience on Linked Data to more information. For example, note 5 of figure 1 corresponds to information about submitting a vocabulary encoded as SKOS to the LCD vocabularies repository on GitHub. We consider this as a good starting point for vocabulary maintainers to explore Linked Data providing both didactical material and practical implementation advice.

Coverage

As mentioned in the section about Characteristics of Conservation Vocabularies, there are still areas of conservation documentation, namely damage types and treatments, for which there is limited or no availability of terms in vocabularies that can be referenced from datasets. One

solution to this problem would be to embark on a new project to develop such a vocabulary. While this may be possible it would require a large exercise with contributions from all conservation sub-domains and some central infrastructure to manage it long term. Another way of approaching this problem is by gradually enriching the AAT by adding missing terms. Organisations which produce conservation records adopt internal lists of terms to describe things like types of damage and conservation treatments. When these lists are published as SKOS and aligned with the AAT, the missing AAT terms are easy to flag. We recommend that at that stage vocabulary maintainers submit the missing terms to the AAT so that they can be used by others. Information about submitting new terms to the AAT is included in our workflows (note 4 of figure 2). We are currently exploring the role of LCD as a group of experts which can act as an editorial group for conservation related submissions to the AAT.

As a result of the discussions at the workshop and the process of developing the vocabulary workflows we have also provided more detailed guidelines for processing, aligning and submitting conservation vocabularies as described next.

LCD vocabulary guidelines

The LCD vocabulary guidelines (https://github.com/linked-conservation-data/conservation-vocabularies/wiki and https://www.ligatus.org.uk/lcd/output/193) expand on the diagrammatic workflows by providing more detailed information for each step of the process as well as examples and references with further didactic material. We note here some recommendations made in the guidelines alongside the associated rationale.

1 Production of URIs: an important decision when publishing Linked Data is the choice of the URIs to act as global identifiers for concepts (note 3 of figure 1). The decision is important because the vocabulary maintainer also maintains these URIs in the long term

to ensure that a (browser) request for a URI resolves to meaningful content returned by some infrastructure. As the resources available for publishing datasets to conservators are limited, the guidelines recommend the option of w3id.org for providing URIs (http://w3id.org). This allows issuing new URIs as subdirectories of w3id.org relatively easily and without the concern of an institutional domain name becoming obsolete. For example, the Language of Bindings Thesaurus is using w3id.org/lob for concept URIs which currently redirect to ligatus.org.uk/lob. Changing an institutional domain name means simply changing the redirection rules on w3id.org. Rules issued by w3id.org indicate that a valid response to a request for the main subdirectory is necessary, therefore at least a holding page for each thesaurus is necessary, but this is a small commitment in comparison to providing such access for each concept.

2 Versioning: the guidelines recommend that managing vocabularies versions should be done at dataset level as opposed to concept level. Version control through breaking down a vocabulary into individual components has been proposed in the past (Halilaj et al. 2016). This approach may be more practical when it comes to tracking changes in specific concepts and also tracking the provenance of changes. However it proves too complex when aligning vocabularies as most of the available software for vocabulary alignment is designed to work at dataset level.

3 Vocabulary data versus alignment data: the guidelines recommend that vocabulary data are kept in separate files to alignment data. This serves two purposes: a) it allows contributions to the alignment tasks by persons/teams other than the main vocabulary maintainer/s, b) it allows for different versions of alignment data without requiring new versions of labels or scope notes.

The guidelines also recommend SKOS validation rules for vocabularies based on qSKOS tool (Suominen and Mader 2014) as well as the file formats CSV and TRIG for depositing records.

The former to be used by vocabulary maintainers who require support in publishing their vocabularies as Linked Data and the latter for direct use in thesaurus software.

Application

As part of the Linked Conservation Data project a Linked Data pilot (https://lcd.researchspace.org/) was implemented on the subject of a specific conservation treatment for books: board re-attachment. As part of this exercise we integrated records from four different institutions which had adopted different vocabularies. We applied the guidelines mentioned here to these vocabularies. The production of URIs was a straight-forward process by using a UUID for each concept. The vocabularies were encoded as versioned SKOS datasets both to organise the development of the datasets during the pilot but also to enable future use without ambiguous referencing. The vocabularies where then aligned with the AAT and the LoB and the alignment datasets kept separate to the main vocabulary SKOS encodings.

Vocabularies in the pilot were flat word lists and lacked broader/narrower relationships. This meant that it was not possible to produce querying tools based on hierarchical retrieval or faceted searching, but we were able to interrogate the datasets using label search. As an alternative solution to the lack of broader/narrower relationships typically in conservation vocabularies one can adopted hierarchical matching relationships such SKOS broad match and take advantage of the existing hierarchies in target vocabularies. This was also necessary in our case due to the large number of composite terms in the dataset vocabularies which typically do not have direct matches in target vocabularies but can be connected with broad match relationships. For example, a local concept ‘goat tanned skin’ has broader match in a target vocabulary ‘goat skin’ and also ‘tanned skin’. This would make it appear under both of these hierarchies.

The vocabularies were integrated in the pilot successfully based on these principles and a number of querying pages were structured around them. An example is a webpage highlighting the alignment of vocabularies with the AAT and LoB: https://lcd.researchspace.org/resource/rsp:Vocabularies .

Conclusions

Our survey showed there is a wide range of vocabularies that are at very different levels of readiness for use in Linked Data. We use those already published as Linked Data as examples of good practice and we provide resources through the workflows and the guidelines to support vocabulary maintainers to process and publish their dataset.

The coverage of terms needed to represent conservation documentation is mixed. Many good sources to describe collection items exist. These are often compiled by people in allied fields, for example art historians and rare books catalogers. When it comes to coverage of terms specific to conservation like damage condition or specific treatment techniques further work is required. We encourage vocabulary maintainers to consider enriching thesauri like the AAT by submitting missing terms which they otherwise use in their daily practice.

Establishing and maintaining required infrastructure for publishing Linked Data as well as securing staff time to undertake data processing for vocabularies is a challenging task in conservation. The LCD consortium brings together partners from many organisations and aims to build enough expertise to crowd-source the maintenance and processing of vocabularies. We recommend shared infrastructure provided by third parties, such as GitHub and w3id.org to host essential tools for sharing data. We intend to work collaboratively with professional bodies to identify resources to allow the development of a user friendly vocabulary discovery tool.

We recognise that the main focus of the project is primarily conservation practice in the UK, US and Europe and that this does not provide a representative view of practices in other parts of the world. We also recognise that the technologies and systems used to manage vocabularies are centered in the knowledge systems and traditions that may not reflect different epistemologies. We try to approach conservators worldwide through engagement with professional bodies to improve our understanding of the use of vocabularies in conservation.

Acknowledgements

The authors thank Prof. Marcia Zeng for her guidance during the project, Dr Jennifer Murray for organising and reviewing the conservation vocabularies included, the participants of the terminology workshop and the terminology working group and the members of the LCD consortium. This project is funded by the Arts and Humanities Research Council in the UK.

Appendix

Linked Conservation Data - Thesaurus and Glossary Usage Questionnaire

Distributed via professional listservs Spring 2019.

Questions:

Email

Name

Organization

What is your current professional role?

- Conservator

Conservation Scientist

Thesaurus editor

Data manager

Other

What types of documentation do you generate?

Conservation Treatment Documentation

Scientific Data Collection and Analysis

Condition Surveys and Assessments

Not applicable in current role

Is your documentation integrated with a larger institutional system (museum management system, integrated library system, digital repository, etc)

Yes

No

Unsure

How do you choose your terminology? (check all that apply)

Published thesaurus - if, yes please proceed to the following questions to indicate which ones.

In house glossary or thesaurus

No controlled vocabulary, choices guided by experience and knowledge

If you do use thesauri or glossaries in your documentation, please let us know which ones. If they are not given in the choices below, we would appreciate your listing them in the "other" category. Add as many as you use separated by commas.

General

Art & Architecture Thesaurus. Getty Research Institute

CAMEO: Conservation & Art Materials Encyclopedia Online. Museum of Fine Arts Boston.

European or British Standards such as: EN 15898, EN 16095, etc

Information Standards in Heritage (FISH) Thesauri

Visual Glossary. Australian Institute for the Conservation of Cultural Material

Wikidata

Other

Architecture/Building Sites/Immovable Heritage

EwaGlos – European Illustrated Glossary for Conservation Terms of Wall Painting and Architectonic Surfaces

Glossary of Architectural Terms. Trust for Architectural Easements

HEREIN System: Terminology. Council of Europe, European Heritage Network.

Illustrated Glossary: Mosaics in Situ Project. Getty Conservation Institute and Israeli Antiquities Authority;

Illustrated glossary on stone deterioration patterns. ICOMOS-ISCS

Other

Book, Paper, and Photographs

Binding Terms. Rare Books and Manuscripts Section: Controlled Vocabularies

Bookbinding and the Conservation of Books: A Dictionary of Descriptive Terminology. Roberts and Etherington.

Descriptive Terminology for Works of Art on Paper. Philadelphia Museum.

Glossary of Terms: American Institute for Conservation, Book and Paper Group

Graphics Atlas. Image Permanence Institute

Language of Bindings. Ligatus

Thesaurus for Graphic Materials. Library of Congress

Other

Paintings, Objects, and Textiles

Object Names Thesaurus. British Museum

Painting Conservation. Glossary of Terms. Smithsonian Museum Conservation Institute.

Condition Reporting – Paintings. Part III: Glossary. Canadian Conservation Institute (CCI)

Other

Figures

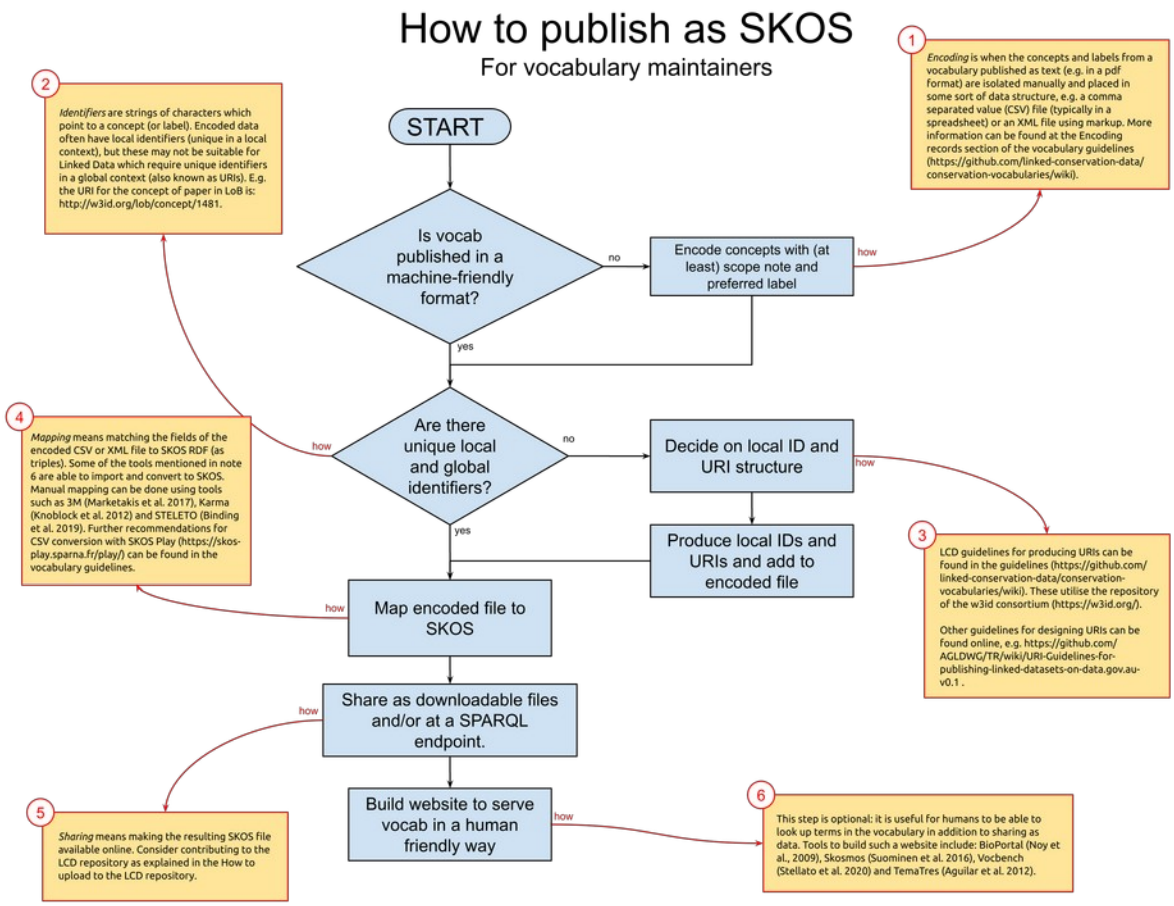

Figure 1: Schematic workflow for publishing vocabularies using SKOS and Linked Data. Yellow boxes provide additional information on how each decision or step should be completed.

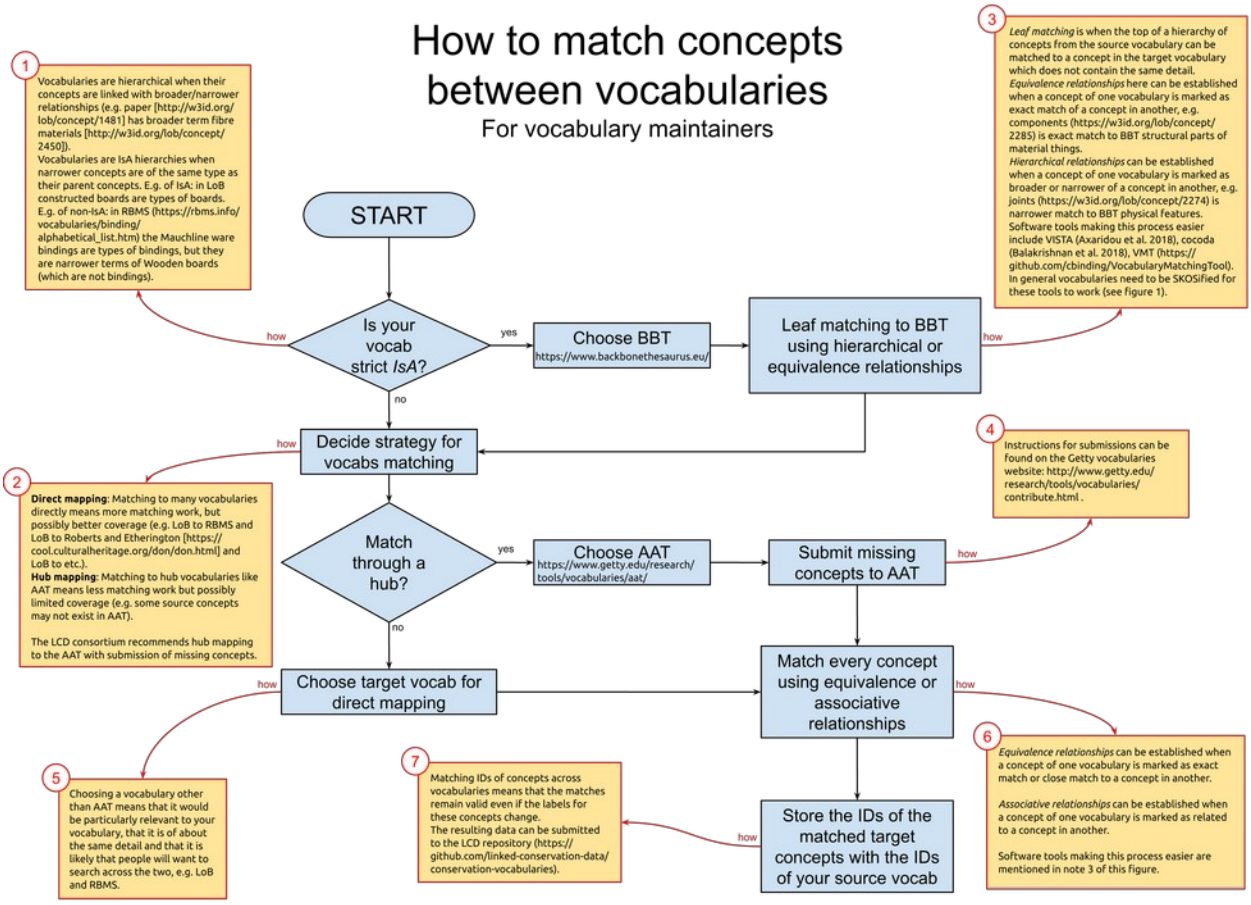

Figure 2: Schematic workflow for aligning vocabularies. Yellow boxes provide additional information on how each decision or step should be completed.

References

Axaridou, Anastasia, Konstantina Konsolaki, Artem Kozlov, Peter Haase, and Martin Doerr. 2018. 'VisTA: Visual Terminology Alignment Tool for Factual Knowledge Aggregation'. *Proceedings of the Third International Workshop on Semantic Web for Cultural Heritage Co-Located with the 15th Extended Semantic Web Conference (ESWC 2018)* 2094: 10.

Balakrishnan, Uma. 2018. 'Towards Integrated Systems for KOS Management, Mapping, and Access: Coli-Conc and Its Collaborative Computer-Assisted KOS Mapping Tool Cocoda'. In *Challenges and Opportunities for Knowledge Organization in the Digital Age*, 693–701. Ergon-Verlag. https://doi.org/10.5771/9783956504211-693.

Binding, Ceri, Douglas Tudhope, and Andreas Vlachidis. 2019. ‘A Study of Semantic Integration across Archaeological Data and Reports in Different Languages’. *Journal of Information Science* 45 (3): 364–86. https://doi.org/10.1177/0165551518789874.

Gonzales Aguilar, Audilio, María Ramírez-Posada, and Diego Ferreyra. 2012. ‘TemaTres’. *El Profesional de La Información* 21 (3): 319–25.

Halilaj, Lavdim, Irlán Grangel-González, Gökhan Coskun, and Sören Auer. 2016. ‘Git4Voc: Git-Based Versioning for Collaborative Vocabulary Development’. *ArXiv:1601.02433 [Cs]*, January. http://arxiv.org/abs/1601.02433.

Knoblock, Craig A., Pedro Szekely, José Luis Ambite, Aman Goel, Shubham Gupta, Kristina Lerman, Maria Muslea, Mohsen Taheriyan, and Parag Mallick. 2012. ‘Semi-Automatically Mapping Structured Sources into the Semantic Web’. In *The Semantic Web: Research and Applications*, edited by Elena Simperl, Philipp Cimiano, Axel Polleres, Oscar Corcho, and Valentina Presutti, 7295:375–90. Lecture Notes in Computer Science. Berlin, Heidelberg: Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-642-30284-8_32.

Mader, Christian. n.d. ‘QSKOS - Integrated Quality Management for SKOS Thesauri’, 14.

Mader, Christian, and Christian Wartena. 2014. ‘Supporting Web Vocabulary Development by Automated Quality Assessment: Results of a Case Study in a Teaching Context’. In . Crete, Greece. http://eprints.cs.univie.ac.at/4045/.

Marketakis, Yannis, Nikos Minadakis, Haridimos Kondylakis, Konstantina Konsolaki, Georgios Samaritakis, Maria Theodoridou, Giorgos Flouris, and Martin Doerr. 2017. ‘X3ML Mapping Framework for Information Integration in Cultural Heritage and Beyond’. *International Journal on Digital Libraries* 18 (4): 301–19. https://doi.org/10.1007/s00799-016-0179-1.

Noy, Natalya F., Nigam H. Shah, Patricia L. Whetzel, Benjamin Dai, Michael Dorf, Nicholas Griffith, Clement Jonquet, et al. 2009. ‘BioPortal: Ontologies and Integrated Data Resources at the Click of a Mouse’. *Nucleic Acids Research* 37 (suppl_2): W170–73. https://doi.org/10.1093/nar/gkp440.

Stellato, Armando, Manuel Fiorelli, Andrea Turbati, Tiziano Lorenzetti, Willem van Gemert, Denis Dechandon, Christine Laaboudi-Spoiden, et al. 2020. ‘VocBench 3: A Collaborative Semantic Web Editor for Ontologies, Thesauri and Lexicons’. *Semantic Web* 11 (5): 855–81. https://doi.org/10.3233/SW-200370.

Suominen, Osma, and Christian Mader. 2014. ‘Assessing and Improving the Quality of SKOS Vocabularies’. *Journal on Data Semantics* 3 (1): 47–73. https://doi.org/10.1007/s13740-013-0026-0.

Suominen, Osma, Henri Ylikotila, Sini Pessala, Mikko Lappalainen, Matias Frosterus, Jouni Tuominen, Thomas Baker, Caterina Caracciolo, and Armin Retterath. 2016. ‘Publishing SKOS Vocabularies with Skosmos’. https://seco.cs.aalto.fi/publications/2016/suominen-et-al-skosmos.pdf.

Velios, Athanasios. 2021. ‘Towards an Open Conservation Documentation Service’. *Journal of the Institute of Conservation* 44 (1): 66–78. https://doi.org/10.1080/19455224.2020.1865176.
‘W3id.Org - Permanent Identifiers for the Web’. n.d. Accessed 14 January 2021. https://w3id.org/.