

# Linked Conservation Data

Notes on discussing terminology

Athanasios Velios  ENGLISH  
HERITAGE

NKOS workshop – 9 Nov 2023



**Linked  
Conservation  
Data**



Portrait of Olive Boteler Porter by Anthony van Dyck, before and after conservation, The Bowes Museum

ual: university of the arts london | LIGATUS

Stanford LIBRARIES

Bodleian Libraries UNIVERSITY OF OXFORD

University of South Wales Prifysgol De Cymru

Medway The GUILDHALL Museum

The British Museum

IIC

LIBRARY LIBRARY OF CONGRESS

THE NATIONAL GALLERY

ICON THE INSTITUTE OF CONSERVATION

KENT STATE UNIVERSITY

SF MO MA

San Francisco Museum of Modern Art

BMIR

UNIVERSITÉ de Cergy-Pontoise

THE GETTY RESEARCH INSTITUTE THE FRANK GETTY TRUST

The Fitzwilliam Museum CAMBRIDGE

LIBRARY HSLIBRB

NATIONAL GALLERIES SCOTLAND

SAPIENZA UNIVERSITÀ DI ROMA

EAMENA

THE NATIONAL ARCHIVES

GallerySystems Collect. Manage. Share.

FORTH INSTITUTE OF COMPUTER SCIENCE



# Benefits of integration

- Understanding the technology of objects (materials and techniques)
  - better-informed decisions on how to care for and treat objects
- Assessing new (and old) materials and techniques for treatment
  - access observations of colleagues on new materials and techniques to decide on suitability
- Identifying patterns of damage
  - correlation between damage and other parameters such as treatment materials, geographical location and environmental conditions
- Environment control based on condition surveys and monitoring data
  - better-informed choices for managing the environment of objects
- Identifying trends in the work of conservators
  - study conservation work undertaken in different territories and how this relates to professional issues such as the type of training available, contractual status, age etc.
- Estimating the condition of collections at a broad/national level and supporting arguments for funding



## Linked Conservation Data

Linked Conservation Data is a Network of partners working on improving access to conservation documentation records. The aim of the Network is to discuss and report on ways that conservation documentation can be disseminated and re-used more effectively through Linked Data.

### Project objectives

We have identified three areas of development for the network's attention: Terminology, Modelling, and Dissemination.

- **Terminology:** In the Semantic Web, communicating by using a variety of terminology traditions is important for disambiguation. The Network will assess the suitability of existing vocabularies in conservation and identify the amount of work needed both in terms of coverage and in terms of formatting to improve them for use in Linked Data applications. The relevant Linked Data standard for vocabularies is SKOS.
- **Modelling:** In the Semantic Web, the type of each published record needs to be explicitly declared. For example, machines need to be able to handle records of type *condition assessment* and records of type *treatment proposal* differently. A standard which provides different types of records (classes) is the CIDOC-CRM. The Network will assess the suitability of the CRM and its extensions for conservation.
- **Dissemination:** The Network will share best practices for producing Linked Data from conservation documentation and report on the readiness and capacity of existing software to host and share Linked Data.

### Terminology guidelines



See the terminology guidelines on GitHub.

Following a review of terminology in conservation, the Linked Conservation Data project has established a GitHub repository to share encoded versions of vocabularies used by conservators. This repository will function as the basis of a terminology portal for conservation. A list of some of these vocabularies can be found on the controlled vocabularies page.

### Overview of current conservation and restoration models

This report is a review of published work describing models for conservation data based on the CIDOC-CRM and its extensions. The report is an analysis of the new classes and properties recommended, as well as the mapping decisions adopted for the published models. The report identifies common practices and recommends ways of harmonising the different models. It proposes areas where further modelling work is needed and attempts to organise this work in conceptually coherent groups from a conservation perspective.

[Overview of current conservation and restoration models \(PDF, 1.13MB\)](#) 11/19 MB

### Conservation data policy template and primer

Policy template and primer for conservation data in memory organisations.

<a href="#">Policy template and primer (BIBEX format)</a>	1/11/19 KB
<a href="#">Policy template and primer (GDF format)</a>	1/10/19 KB
<a href="#">Policy template and primer (PDF format)</a>	145/13 KB

### Linked Data Pilot



[View the LCD Linked Data Pilot.](#)

[Read the LCD Linked Data Pilot summary.](#)

The LCD pilot working group including the Bodleian Library, Oxford; the Library of Congress; the National Archives (UK); and Stanford Libraries undertook a Linked Data pilot to explore a book conservation case study; reattaching detached boards.

### Contact

#### Address:

Linked Conservation Data / Ligatus  
University of the Arts London  
Chelsea College of Arts  
16 John Islip Street  
London  
SW1P 4JU

#### Email:

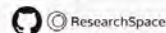
Athanasios Vellios

### Funded By



Arts and  
Humanities  
Research Council

### Repository and Portal



GitHub repository - ResearchSpace portal.

Username\*

Password\*

[Log in](#)

• Reset your password.

# Data sharing policy template



## Linked Conservation Data

Home Resources Meetings Consortium Introduction to Linked Data Updates

Home Resources Reports Conservation data policy template and primer (phase 2)

### Conservation data policy template and primer (phase 2)

Policy template and primer for conservation data in memory organisations. This is the version which incorporates changes as a result of discussions at the IIC Edinburgh Congress in 2020 (see: [Policy Document Workshop – IIC Congress 2020](#)). This version has been sent for ratification to the IIC, Icon and AIC.

#### Attachments:

- [Draft policy template and primer \(for ratification ODT format\)](#)
- [Draft policy template and primer \(for ratification DOCX format\)](#)
- [Draft policy template and primer \(for ratification PDF format\)](#)

#### Contact

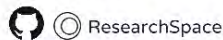
##### Address:

Linked Conservation Data / Ligatus  
University of the Arts London  
Chelsea College of Arts  
16 John Islip Street  
London  
SW1P 4JU

##### Email:

Brigitte Hart

#### Repository and Portal



GitHub repository - ResearchSpace portal

#### Funded by



Arts and  
Humanities  
Research Council

#### User login

Username \*

Password \*

- Request new password

Log in



## Linked Conservation Data

LCD Strategy Working Group

Data policy primer and template

<logos - IIC, Icon, AIC>

Contributions by: Sara Crofts, Anni Mantyniemi, Sarah Stannage, Kristen StJohn, Athanasios Vellos, Eryl Wentworth and others.

Linked Conservation Data is funded by:



Arts and  
Humanities  
Research Council



7

8

<https://www.ligatus.org.uk/lcd/output/216>

# CIDOC-CRM

Suitability for conservation

- Actively engaged in the community
- Event-based model
  - how objects change
  - development of condition
  - scientific measurements and sampling

# CIDOC-CRM

Suitability for conservation

- Identify missing classes and properties
  - recording risk
  - negative statements
  - treatment plans
  - exhibition instructions
- Decide how best to use the CRM
  - Discuss minimum conservation model
  - Consider extensions: CRMsci, CRMact etc.



# Linked Conservation Data Pilot

hosted by the British Museum using [ResearchSpace](#)

This site is a pilot implementation of Linked Data in conservation using a specific case study from the field of book conservation: board re-attachment. Books with detached boards are common in historic library collections and conservators have applied a variety of techniques and materials when re-attaching boards over the years.

[Board re-attachment research questions](#)

[Books](#)

[Example conservation record](#)

[Example conservation report](#)

The institutions which contributed to the pilot are:

- Stanford Libraries
- Bodleian Library
- Library of Congress
- The National Archives (UK)

In the near future we also expect contributions from the British Library

<https://lcd.researchspace.org/>

This project is funded by the



## What is Linked Conservation Data?

Linked Conservation Data is a Network of partners working on improving access of conservation documentation records. The aim of the Network is to discuss and report on ways that conservation documentation can be disseminated and re-used more effectively through Linked Data.

[LCD Project](#)

[LCD Partners and Contributors](#)

## What is Linked Data?

Linked Data refers to a set of technologies which are delivering the Semantic Web. Linked Data allows sharing of datasets across institutions by enabling joint query and cross-linking of individual records (not only document hyperlinks). We are exploring these technologies for the benefit of conservation.

[W3C introduction](#)

[LCD introduction to Linked Data](#)

# Linked Conservation Data Pilot

hosted by the British Museum using [ResearchSpace](#)

This site is a pilot implementation of Linked Data in conservation using a specific case study from the field of book conservation: board re-attachment. Books with detached boards are common in historic library collections and conservators have applied a variety of techniques and materials when re-attaching boards over the years.

[Board re-attachment research questions](#)

[Books](#)

[Example conservation record](#)

[Example conservation report](#)

The institutions which contributed to the pilot are:

- [Stanford Libraries](#)
- [Bodleian Library](#)
- [Library of Congress](#)
- [The National Archives \(UK\)](#)

In the near future we also expect contributions from the British Library

This project is funded by the



## What is Linked Conservation Data?

Linked Conservation Data is a Network of partners working on improving access of conservation documentation records. The aim of the Network is to discuss and report on ways that conservation documentation can be disseminated and re-used more effectively through Linked Data.

[LCD Project](#)

[LCD Partners and Contributors](#)

## What is Linked Data?

Linked Data refers to a set of technologies which are delivering the Semantic Web. Linked Data allows sharing of datasets across institutions by enabling joint query and cross-linking of individual records (not only document hyperlinks). We are exploring these technologies for the benefit of conservation.

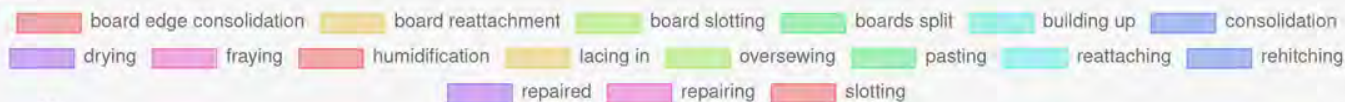
[W3C introduction](#)

[LCD introduction to Linked Data](#)

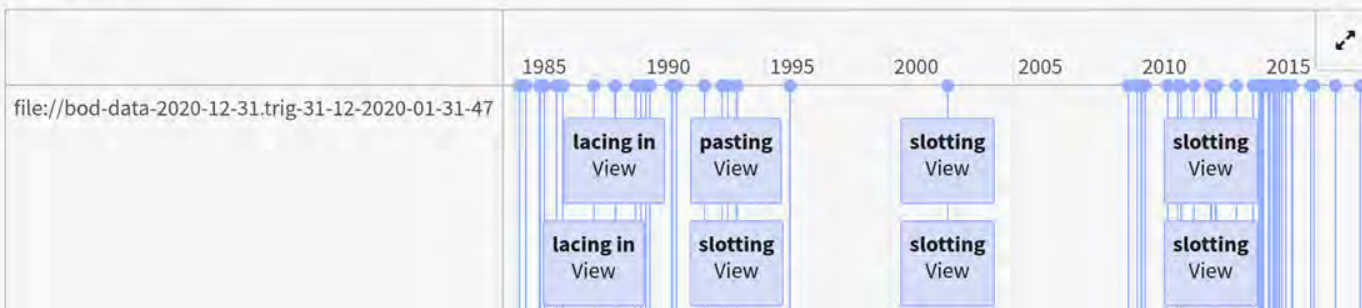
# LCD pilot research questions

## Board attachment techniques over time

### FREQUENCY OVER TIME



### FULL DATA



Clipboard

Details

### Board attachment materials over time

#### FREQUENCY OVER TIME

Material

animal glue

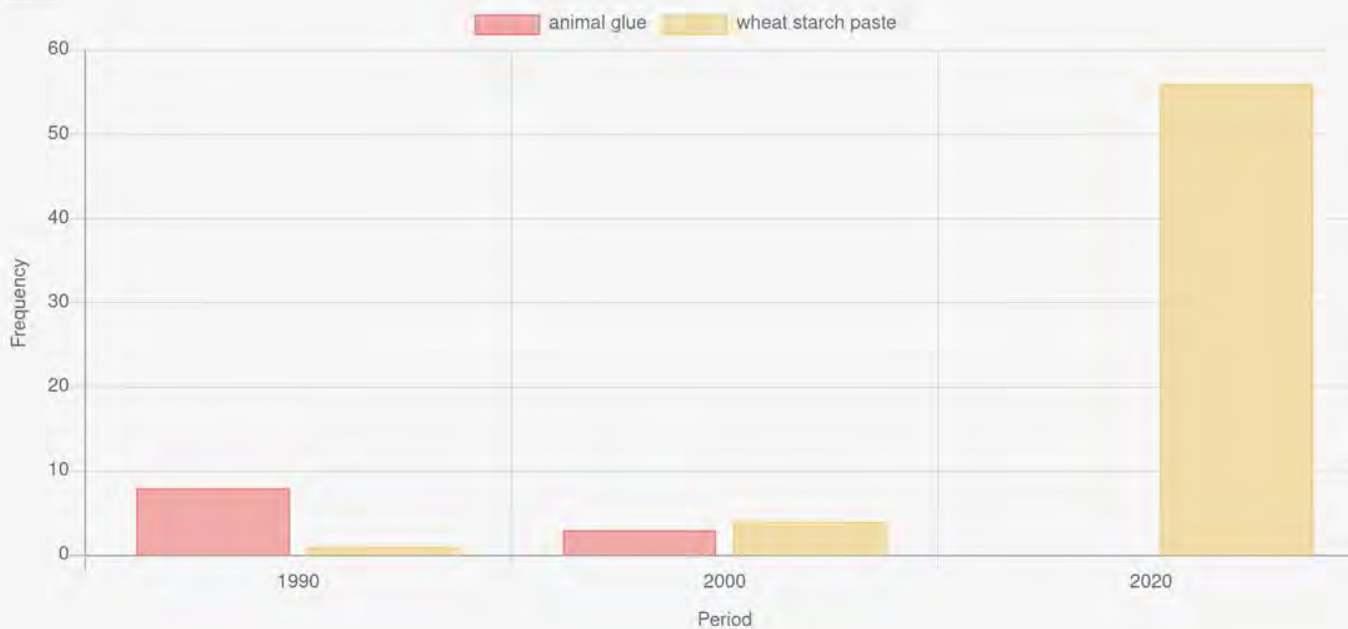
x ▼ x

wheat starch paste

x ▼ x

+ Add material

SEARCH



Clipboard

Details

# Terminology working group

- Anastasia Axaridou
- Ceri Binding
- Nicola Carboni
- Kirsten Dunne
- John Graybeal
- Ryan Lieu
- Jennifer Murray
- Joseph Padfield
- Eleni Tsouloucha
- Jon Ward
- Francesca Whymark
- Marcia Zeng
- Kristen St. John
- Athanasios Velios



Send



Questions

Responses

37

Settings

## Linked Conservation Data - Thesaurus and Glossary Usage Questionnaire

Email \*

Valid email address

This form is collecting email addresses. [Change settings](#)

Name \*

Short-answer text

Organization

Short-answer text

What is your current professional role? \*

Conservator

Conservation Scientist



# Questions

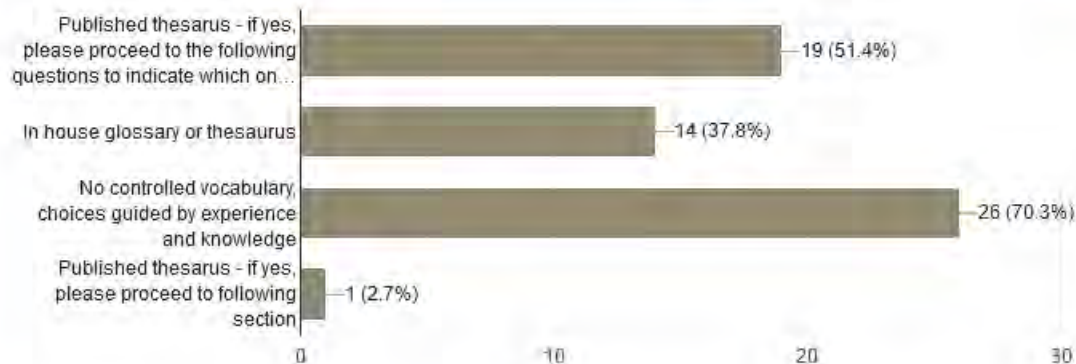
- Type of participant role
- Practical software/systems setup
- Thesauri/glossaries and vocabularies used



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



37 responses

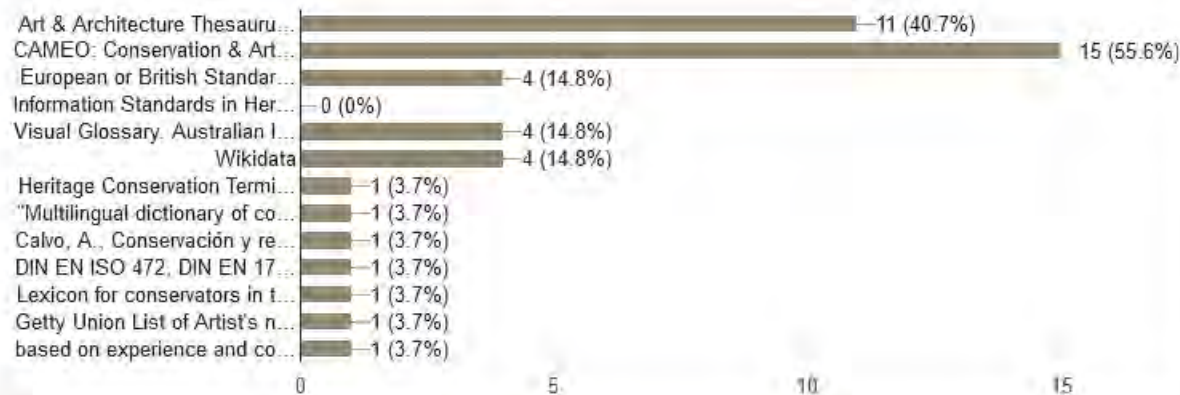


### Thesauri or Glossaries

#### General



27 responses







# Linked Conservation Data

[Home](#)[Resources ▾](#)[Meetings ▾](#)[Consortium](#)[Introduction to Linked Data ▾](#)[Updates](#)

## Conservation controlled vocabularies

### Index

[A Glossary of Stained Glass Terms](#)

[A library, media, and archival preservation glossary](#)

[ABC of bookbinding](#)

[AIC BPG Glossary of Terms](#)

[AIC Lexicon Terms](#)

[Art & Architecture Thesaurus \(AAT\)](#)

[AV Artifact Atlas](#)

[Binding Terms. Rare Books and Manuscripts Section: Controlled Vocabularies](#)

[BM Object names Thesaurus](#)

[Bookbinding and the Conservation of Books: A Dictionary of Descriptive Terminology. Roberts and Etherington.](#)

[BS EN 15898:2011 - Conservation of cultural property. Main general terms and definitions](#)

[BS EN 923:2015 Adhesives. Terms and definitions.](#)

[CAMEO: Conservation & Art Materials Encyclopedia Online](#)

[Compendium of Image Errors in Analogue Video](#)

[Condition Reporting – Paintings. Part III: Glossary – Canadian Conservation Institute \(CCI\) Notes 10/11](#)

[Conservation and Restoration of Stained Glass - an owner's guide](#)

[Conservation of the architectural surfaces in the tablinum of the House of the Bicentenary, Herculaneum : phase 1, examination, investigations, and condition assessment](#)

[Descriptive Terminology for Works of Art on Paper](#)

[Diccionario técnico Akal de conservación y restauración de bienes culturales](#)

[Dictionary of Book and Paper Conservation](#)

[Dictionary of Building Preservation](#)

[EwaGlos – European Illustrated Glossary for Conservation Terms of Wall Painting and Architectonic Surfaces](#)

[FIAF manual, D.16 Manifestation/Item Condition, Preservation and Restoration](#)

[FISH Archaeological Objects Thesaurus](#)

[FISH Archaeological Sciences Thesaurus](#)

[FISH Building Materials Thesaurus](#)

[FISH Components Thesaurus](#)

[FISH Event Types Thesaurus](#)

[FISH Evidence Thesaurus](#)

[FISH Thesaurus of Monument Types](#)

[FISH Building Materials Thesaurus](#)  
[FISH Components Thesaurus](#)  
[FISH Event Types Thesaurus](#)  
[FISH Evidence Thesaurus](#)  
[FISH Thesaurus of Monument Types](#)  
[Glassworking, Ancient and Medieval, Terminology, technology and typology](#)  
[Glosario de términos relacionados con el deterioro de las piedras de construcción](#)  
[Glossary of Arabic terms for the conservation of cultural heritage](#)  
[Glossary of Architectural Terms - Trust for Architectural Easements](#)  
[Glossary of Basic Archival and Library Conservation Terms](#)  
[Glossary of Conservation](#)  
[Graphics Atlas](#)  
[Guide for Conducting a Visual Inspection of Concrete in Service](#)  
[Herein: Terminology](#)  
[Heritage Conservation Terminology Definition of terms from various sources](#)  
[IASA Cataloging Rules: Appendix C. Terms for describing the physical condition of sound recordings](#)  
[Illustrated Dictionary of Architectural Preservation: Restoration, Renovation, Rehabilitation, Reuse](#)  
[Illustrated Glossary by the York Glaziers Trust](#)  
[Illustrated glossary on stone deterioration patterns \(English - French\)](#)  
[Illustrated glossary on stone deterioration patterns \(English - German\)](#)  
[Illustrated Glossary: Mosaics in Situ Project](#)  
[Illustrated Glossary: Technician Training for the Maintenance of In Situ Mosaics](#)  
[Language of Bindings Thesaurus](#)  
[Les altérations visibles du béton : définitions et aide au diagnostic](#)  
[Looking at Paintings: A Guide to Technical Terms](#)  
[Nomenclature for museum cataloguing](#)  
[Painting Conservation - Glossary of Terms - Smithsonian Museum Conservation Institute](#)  
[Petra archaeological park weathering and deterioration definitions](#)  
[Stained Glass Terminology](#)  
[Thesaurus for Graphic Materials](#)  
[Visual Glossary - Australian Institute for the Conservation of Cultural Material](#)

---

## A Glossary of Stained Glass Terms

glass, URI production, images included, encoding - scripting, concept matching, How something is made

---

## A library, media, and archival preservation glossary

DEPEW, J. N., & JONES, C. L. (1992). A library, media, and archival preservation glossary.

Alphabetical list, source cited at end of scope note, related terms in bold in scope note.

Preface: " The purpose of this glossary is to bring together many of the terms that relate to the conservation and preservation of archival, library and media center materials, many of which are scattered in articles, books, dictionaries, glossaries, and reports throughout the literature of preservation. The terms are drawn not only from the field of library and information science, but from the binding, paper, photographic and preservation reformatting literature, as well. The glossary also contains building, HVAC, insurance and statistical terms that are useful in

## Binding Terms. Rare Books and Manuscripts Section: Controlled Vocabularies

Controlled Vocabularies for Use in Rare Book and Special Collections Cataloging, developed and maintained by the Bibliographic Standards Committee of the Rare Books and Manuscripts Section (ACRL/ALA). These thesauri provide standardized vocabulary for retrieving special collections materials by form, genre, or by various physical characteristics that are typically of interest to researchers and special collections librarians, and for relating materials to individuals or corporate bodies.

Binding terms thesauri: option to search alphabetically or hierarchically. "The Alphabetical List contains unapproved terms and references. To determine whether a term is approved and to determine appropriate usage as described in the Scope Note the full term record must be consulted."

RBMS produce 6 thesauri: Binding Terms; Genre Terms; Paper Terms; Printing & Publishing Evidence; Provenance Evidence; Type Evidence. Examining here: Binding Terms; Genre Terms; Paper Terms

books, techniques, materials alternative labels included, no scope notes, hierarchies included, structured data, encoding - scripting, URI production, concept matching

## BM Object names Thesaurus

The thesaurus reflects the nature of the British Museum's internal reference tool. Some areas of terminology may be more specific than others, depending on the level of documentation available. books, techniques, materials, alternative labels included, structured data, encoding - scripting, URI production, concept matching

subject area within conservation

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

ROBERTS, M., ETHERINGTON, D., & BROWN, M. R. (1982). Bookbinding and the conservation of books: a dictionary of descriptive terminology.

Entries arranged alphabetically. References at the end of each entry. Small number of images at end. Search facility. References via link.

Preface: " Although this dictionary is intended first and foremost for those actively involved in one or more aspects of the overall field of bookbinding and book conservation, including bookbinders, conservators of library and archival materials, and the like, it is perhaps no less intended for those working in related fields, such as bibliography and librarianship, where the many terms and expressions relating to the overall field may be less familiar and even more confusing... The definitions themselves, .., were, whenever possible, drawn from the most authoritative sources available (as indicated by the number in parentheses at the end of a definition, which refers to the Sources and Bibliography section) and supplemented by the experience of the authors. Even though the bibliography and sources cited represent but a relatively small segment of the extensive corpus of literature in the field of conservation, we believe they provide a reasonably good sampling and may benefit the reader by offering an authoritative source for the terms and sometimes providing a source for further investigation. Definitions that do not cite a source are entirely the responsibility of the authors.

Where a term has more than one definition, each is numbered and arranged in its descending order of significance in relation to bookbinding."

books, techniques, materials, treatment, references included, images included, alternative labels included, URI production, concept matching, build hierarchies, encoding - scripting, encoding - transcribing, semi-structured data

## BS EN 15898:2011 - Conservation of cultural property. Main general terms and definitions

BSI BRITISH STANDARDS. (2011). Conservation of cultural property. Main general terms and definitions. London, BSI British Standards.

"This standard is part of a series of standards being developed by CEN/TC 346. It provides terms and their definitions for a set of general concepts which are widely used by those working in

## Binding Terms. Rare Books and Manuscripts Section: Controlled Vocabularies

Controlled Vocabularies for Use in Rare Book and Special Collections Cataloging, developed and maintained by the Bibliographic Standards Committee of the Rare Books and Manuscripts Section (ACRL/ALA). These thesauri provide standardized vocabulary for retrieving special collections materials by form, genre, or by various physical characteristics that are typically of interest to researchers and special collections librarians, and for relating materials to individuals or corporate bodies.

Binding terms thesauri: option to search alphabetically or hierarchically. "The Alphabetical List contains unapproved terms and references. To determine whether a term is approved and to determine appropriate usage as described in the Scope Note the full term record must be consulted."

RBMS produce 6 thesauri: Binding Terms; Genre Terms; Paper Terms; Printing & Publishing Evidence; Provenance Evidence; Type Evidence. Examining here: Binding Terms; Genre Terms; Paper Terms

books, techniques, materials, alternative labels included, no scope notes, hierarchies included, structured data, encoding - scripting, URI production, concept matching

## BM Object names Thesaurus

The thesaurus reflects the nature of the British Museum collections, and was originally set up as an internal reference, depending on the level of documentation available, or the size of particular collections.

objects, techniques, materials, alternative labels included, no scope notes, hierarchies included, structured data, encoding - scripting, URI production, concept matching

existing technical setup

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

ROBERTS, M., ETHERINGTON, D., & BROWN, M. R. (1982). Bookbinding and the conservation of books: a dictionary of descriptive terminology.

Entries arranged alphabetically. References at the end of each entry. Small number of images at end. Search facility. References via link.

Preface: " Although this dictionary is intended first and foremost for those actively involved in one or more aspects of the overall field of bookbinding and book conservation, including bookbinders, conservators of library and archival materials, and the like, it is perhaps no less intended for those working in related fields, such as bibliography and librarianship, where the many terms and expressions relating to the overall field may be less familiar and even more confusing... The definitions themselves, .., were, whenever possible, drawn from the most authoritative sources available (as indicated by the number in parentheses at the end of a definition, which refers to the Sources and Bibliography section) and supplemented by the experience of the authors. Even though the bibliography and sources cited represent but a relatively small segment of the extensive corpus of literature in the field of conservation, we believe they provide a reasonably good sampling and may benefit the reader by offering an authoritative source for the terms and sometimes providing a source for further investigation. Definitions that do not cite a source are entirely the responsibility of the authors.

Where a term has more than one definition, each is numbered and arranged in its descending order of significance in relation to bookbinding."

books, techniques, materials, treatment, references included, images included, alternative labels included, URI production, concept matching, build hierarchies, encoding - scripting, encoding - transcribing, semi-structured data

## BS EN 15898:2011 - Conservation of cultural property. Main general terms and definitions

BSI BRITISH STANDARDS. (2011). Conservation of cultural property. Main general terms and definitions. London, BSI British Standards.

"This standard is part of a series of standards being developed by CEN/TC 346. It provides terms and their definitions for a set of general concepts which are widely used by those working in

## Binding Terms. Rare Books and Manuscripts Section: Controlled Vocabularies

Controlled Vocabularies for Use in Rare Book and Special Collections Cataloging, developed and maintained by the Bibliographic Standards Committee of the Rare Books and Manuscripts Section (ACRL/ALA). These thesauri provide standardized vocabulary for retrieving special collections materials by form, genre, or by various physical characteristics that are typically of interest to researchers and special collections librarians, and for relating materials to individuals or corporate bodies.

Binding terms thesauri: option to search alphabetically or hierarchically. "The Alphabetical List contains unapproved terms and references. To determine whether a term is approved and to determine appropriate usage as described in the Scope Note the full term record must be consulted."

RBMS produce 6 thesauri: Binding Terms; Genre Terms; Paper Terms; Printing & Publishing Evidence; Provenance Evidence; Type Evidence. Examining here: Binding Terms; Genre Terms; Paper Terms

books, techniques, materials, alternative labels included, no scope notes, hierarchies included, structured data, encoding - scripting, URI production, concept matching

## BM Object names Thesaurus

The thesaurus reflects the nature of the British Museum collections, and was originally set up as an internal reference tool. Some areas of terminology may be more specific than others, depending on the level of documentation available, or the size of the collection.

objects, techniques, materials, alternative labels included, no scope notes, hierarchies included, structured data, encoding - scripting, URI production, concept matching

## Bookbinding and the Conservation of Books: Terminology. Roberts and Etherington.

ROBERTS, M., ETHERINGTON, D., & BROWN, M. R. (1982). Bookbinding and the Conservation of Books: Terminology. London: British Library.

Entries arranged alphabetically. References at the end of each entry. Small number of images at end. Search facility. References via link.

Preface: "Although this dictionary is intended first and foremost for those actively involved in one or more aspects of the overall field of bookbinding and book conservation, including bookbinders, conservators of library and archival materials, and the like, it is perhaps no less intended for those working in related fields, such as bibliography and librarianship, where the many terms and expressions relating to the overall field may be less familiar and even more confusing... The definitions themselves, .., were, whenever possible, drawn from the most authoritative sources available (as indicated by the number in parentheses at the end of a definition, which refers to the Sources and Bibliography section) and supplemented by the experience of the authors. Even though the bibliography and sources cited represent but a relatively small segment of the extensive corpus of literature in the field of conservation, we believe they provide a reasonably good sampling and may benefit the reader by offering an authoritative source for the terms and sometimes providing a source for further investigation. Definitions that do not cite a source are entirely the responsibility of the authors.

Where a term has more than one definition, each is numbered and arranged in its descending order of significance in relation to bookbinding."

books, techniques, materials, treatment, references included, images included, alternative labels included, URI production, concept matching, build hierarchies, encoding - scripting, encoding - transcribing, semi-structured data

## BS EN 15898:2011 - Conservation of cultural property. Main general terms and definitions

BSI BRITISH STANDARDS. (2011). Conservation of cultural property. Main general terms and definitions. London, BSI British Standards.

"This standard is part of a series of standards being developed by CEN/TC 346. It provides terms and their definitions for a set of general concepts which are widely used by those working in

essential work for  
vocabulary  
alignment

# Vocabulary questionnaire

- Help with:
  - Understand trends
  - Understand scope overlap
  - Prioritise/plan future work
  - Estimate costs and effort
  - Establish relationships with vocabulary maintainers

# Vocabulary alignment infrastructure

- Long-term sustainability
  - Low cost
  - Shared ownership
- Github repository and wiki



master

1 branch 0 tags

Go to file

Code

natuk Files for the Smithsonian glossary example of the guidelines. 503f0dd on Mar 25, 2021 27 commits

align	Files for the Smithsonian glossary example of the guidelines.	2 years ago
vocabs	Files for the Smithsonian glossary example of the guidelines.	2 years ago
.gitignore	Updated .gitignore file to include xlsx files	3 years ago
README.md	Files for the Smithsonian glossary example of the guidelines.	2 years ago

README.md

## Linked Conservation Data

<https://www.ligatus.org.uk/lcd>

## Conservation vocabularies repository

This repository hosts raw data representing conservation vocabularies: both published thesauri as well as local wordlists and vocabularies used in institutions or studios. The repository also hosts alignment data between different vocabularies.

## How to use

Best practices for the use of the repository are described in the [repository wiki pages](#).

### About

Datasets of contributed conservation vocabularies alongside vocabulary alignment data.

Readme

Activity

9 stars

4 watching

1 fork

Report repository

### Releases

No releases published

### Packages

No packages published

### Contributors 2

natuk Athanasios Velios

acampagnolo Alberto Campagnolo





linked-conservation-data / conservation-vocabularies Public

Notifications

Fork 1

Star 9

- Code
- Issues
- Pull requests
- Actions
- Projects
- Wiki
- Security
- Insights

Files

master

Go to file

- align
- vocabs
  - bod-vocab
  - loc-vocab
  - mibact-sapienza
  - spg
  - sul-vocab
  - tna-vocab
- .gitignore
- README.md

conservation-vocabularies / vocabs /



natuk Files for the Smithsonian glossary example of the guidelines.

503f0dd · 2 years ago

History

Name	Last commit message	Last commit date
..		
bod-vocab	Commit prior to merging from Alberto's branch	3 years ago
loc-vocab	Corrected https and URIs for bod and loc vocabs - delet...	3 years ago
mibact-sapienza	Added new version 20201132 of mibact vocabularies.	3 years ago
spg	Files for the Smithsonian glossary example of the guideli...	2 years ago
sul-vocab	Fixed http protocol in URIs. Made it plain (not https).	3 years ago
tna-vocab	Commit prior to merging from Alberto's branch	3 years ago



linked-conservation-data / conservation-vocabularies Public

Notifications Fork 1 Star 9

- Code Issues Pull requests Actions Projects Wiki Security Insights

Files

master

Go to file

- align
bod-vocab--aat
bod-vocab--lob
loc-vocab--aat
loc-vocab--cam
loc-vocab--lob
mibact-sapienza--aat
spg--aat
sul-vocab--aat
sul-vocab--bib
sul-vocab--bpg
sul-vocab--cam
sul-vocab--lob
tna-vocab--aat
tna-vocab--cam
tna-vocab--lob
vocabs
bod-vocab

conservation-vocabularies / align /

natuk Files for the Smithsonian glossary example of the guidelines. 503f0dd · 2 years ago History

Table with 3 columns: Name, Last commit message, Last commit date. Rows include folders like bod-vocab--aat, loc-vocab--aat, and files like spg--aat.



# Home

Athanasios Velios edited this page on Mar 25, 2021 · 4 revisions



## Linked Conservation Data: Vocabulary guidelines

### LCD Terminology Working Group

Edited by:

- Athanasios Velios
- Kristen St. John

Contributions by:

- Anastasia Axaridou
- Ceri Binding
- Nicola Carboni

Pages 10

### Browse

- [Introduction](#)
- [Types of vocabularies](#)
- [Aim](#)
- [Hosting vocabularies in LCD](#)
- [Aligning vocabularies for LCD](#)
- [How to upload to the LCD repository](#)
- [Quick start](#)
- [Example process](#)
- [Repository maintainers](#)

Clone this wiki locally

<https://github.com/linked-conser>



- Nicola Carboni
- Kirsten Dunne
- John Graybeal
- Ryan Lieu
- Joseph Padfield
- Eleni Tsouloucha
- Jon Ward
- Francesca Whymark
- Marcia Zeng

and others.



Linked Conservation Data is funded by the [Arts and Humanities Research Council](#)

## Table of contents [↗](#)

---

- [Introduction](#)
- [Types of vocabularies](#)
- [Aim](#)
- [Hosting vocabularies in LCD](#)
- [Aligning vocabularies for LCD](#)
- [How to upload to the LCD repository](#)
- [Quick start](#)
- [Example process](#)
- [Repository maintainers](#)





# Types of vocabularies

Athanasios Velios edited this page on Nov 15, 2020 · 2 revisions

## Types of vocabularies [↗](#)

This section outlines the types of vocabularies considered in LCD. Different processes are required for each type in order to be shared effectively through the LCD repository. In sections [Hosting vocabularies](#) in LCD and [Aligning vocabularies](#) for LCD, the described processes refer to one or more of these types. They are listed here in order of increasing complexity of structure. The structure of each type can be produced by building upon the previous one. More complexity of a vocabulary does not necessarily mean better quality, but in general it leads to easier use of vocabulary data. These types are illustrated in a figure at the end of this section (figure 1). Lists are used for looking up terms when filling in conservation records. Glossaries are used when audiences looking up the vocabularies have different contextual understandings and therefore a word on its own does not necessarily convey the required meaning. Thesauri are used for large vocabularies which are difficult to browse alphabetically especially when the sought term is not known. Thesauri provide extra ways of navigating terms in addition to alphabetical order, such as terms hierarchies and related terms ("see also"). For a thorough explanation please refer [here](#).

## Lists [↗](#)

This primarily includes plain lists of terms (word-lists), without definitions/descriptions, which are used as lookup lists or options in structured records, e.g. as database lookup fields or tick-boxes in survey forms. They do not always consist of terms, they could also include other sequences of symbols (for example, drawing patterns for marking condition on photographs of objects). These lists are often local in scope, i.e. they apply to institutions or conservation labs.

A word-list does not indicate whether some terms are more general than others. For example, the term 'oil' is more general than the term 'linseed oil' since linseed oil is a specific type of oil. In a word-list there is no way of indicating

Pages 10

### Browse [↗](#)

- [Introduction](#)
- [Types of vocabularies](#)
- [Aim](#)
- [Hosting vocabularies in LCD](#)
- [Aligning vocabularies for LCD](#)
- [How to upload to the LCD repository](#)
- [Quick start](#)
- [Example process](#)
- [Repository maintainers](#)

Clone this wiki locally

<https://github.com/linked-conser>

```
<related></related>  
</concept>
```

## Producing URIs [↗](#)

---

### Applies to [↗](#)

This process is required for all vocabularies which do not already provide URIs (see [Vocabularies expressed in SKOS](#)) for each of their concepts.

**Ideally URIs should be created and maintained long-term as described in this section by the vocabulary maintainer. If the technicalities of creating the URIs make the process too resource intensive, then contact the [LCD repository maintainers](#) for suggestions.**

### Purpose [↗](#)

To provide unique identifiers and unambiguous reference points for concepts at a global scope.

### How [↗](#)

LCD requires that a vocabulary concept has a single URI. Concepts that are updated in later versions of vocabularies should maintain the URIs from earlier versions. Updates to scope notes should not change the meaning of the concept but instead explain it in a better way. If the meaning does change, maintainers should consider creating new concepts while keeping the old ones. Using a different URI for a concept means that we are referring to a different concept. URIs used to refer to different versions of the **whole** vocabulary may change when the vocabulary is updated.

If the host organisation for the vocabulary has an existing practice for producing and maintaining URIs, then it is recommended to follow that practice. If there is no such practice then URIs should be produced as explained next.

### URIs for vocabularies [↗](#)

The following patterns for URI production can be applied to any namespace. Vocabulary maintainers or host organisations can use any namespace they are committed to manage. Users can be redirected from that namespace to a location presenting information about the vocabulary if this is at a different place. In practice this means that the URIs point to one server and each one of them is then passed to another server at request. This is seamless to the end user. This redirection is beneficial because it allows another host organisation to take over the management of the vocabulary without affecting the original URIs which remain the same (persistent). The cost for this flexibility is the requirement for managing the redirection server.

</related>></related>  
</concept>

## Producing URIs [↗](#)

### Applies to [↗](#)

This process is required for all vocabularies for each of their concepts.

**Ideally URIs should be created and maintained by a dedicated maintainer. If the technicalities of creating URIs are a barrier, see [LCD repository maintainers](#) for suggestions.**

### Purpose [↗](#)

To provide unique identifiers and unambiguous references to concepts.

### How [↗](#)

LCD requires that a vocabulary concept has a unique URI. It should maintain the URIs from earlier versions of the concept but instead explain it in a better way for new concepts while keeping the old ones. It should not have different URIs for different concepts. URIs used to refer to different concepts should be updated.

If the host organisation for the vocabulary has a preferred practice, it is recommended to follow that practice. If there is no preferred practice, the following is recommended.

### URIs for vocabularies [↗](#)

The following patterns for URI production are recommended. Organisations can use any namespace they prefer. It is recommended to use a namespace to a location presenting information about the vocabulary means that the URIs point to one server and are easy to use seamless to the end user. This redirection is a key part of the management of the vocabulary without affecting the user. This flexibility is the requirement for managing a large number of vocabularies.

A list of terms with unique labels and associated scope notes where necessary. This is broadly the case for the Painting Conservation Glossary from the Smithsonian Conservation Institute available [here](#).

## Encoding records [↗](#)

### Applies to [↗](#)

This process is required for vocabularies held in formats which cannot be processed easily by software to separate labels, scope notes and relationships, to find relevant terms, or to distinguish conflicting uses. Typically this includes vocabularies in print or typeset in PDF files. It may also include vocabularies in text which is partially or inconsistently tagged in wiki-type websites. It may also include vocabularies in text which rely on the textual narrative to communicate labels, scope notes and relationships. Another obstacle of formats which cannot be processed easily may be the fact that terms and concepts are dispersed across documents or resources and their grouping cannot be done automatically.

### Purpose [↗](#)

1. To separate vocabulary information into labels, scope notes and relationships
2. To produce a consistent list of relevant concepts with their associated labels, scope notes and related concepts

### How [↗](#)

Methods depend on the format. A simple but time-consuming method is transcribing text into a spreadsheet or database form by hand.

More complex methods may require [scraping websites](#) and automatically identifying tagged text of interest. The process involves writing a script to load webpages holding vocabulary information, extracting it and storing it in a structured document. An example of doing this on the Smithsonian Painting Conservation Glossary using a script can be found in section: [Encoding Python script](#). In other cases it may require transforming tagged text to a new structure, for example using [XSLT](#) to simplify an elaborate HTML page. Tools such as [Tabula](#) can help with extracting records from a PDF file.

It is likely that the process of encoding is simplified when identifiers are used for concepts and possibly for labels. These identifiers would offer unambiguous references to concepts and labels at local level. Maintainers should consider the next section ([Producing URIs](#)) before establishing local identifiers during encoding.

### Output [↗](#)

## Non-technical quick start

To share your vocabulary on the LCD repository, follow these steps:

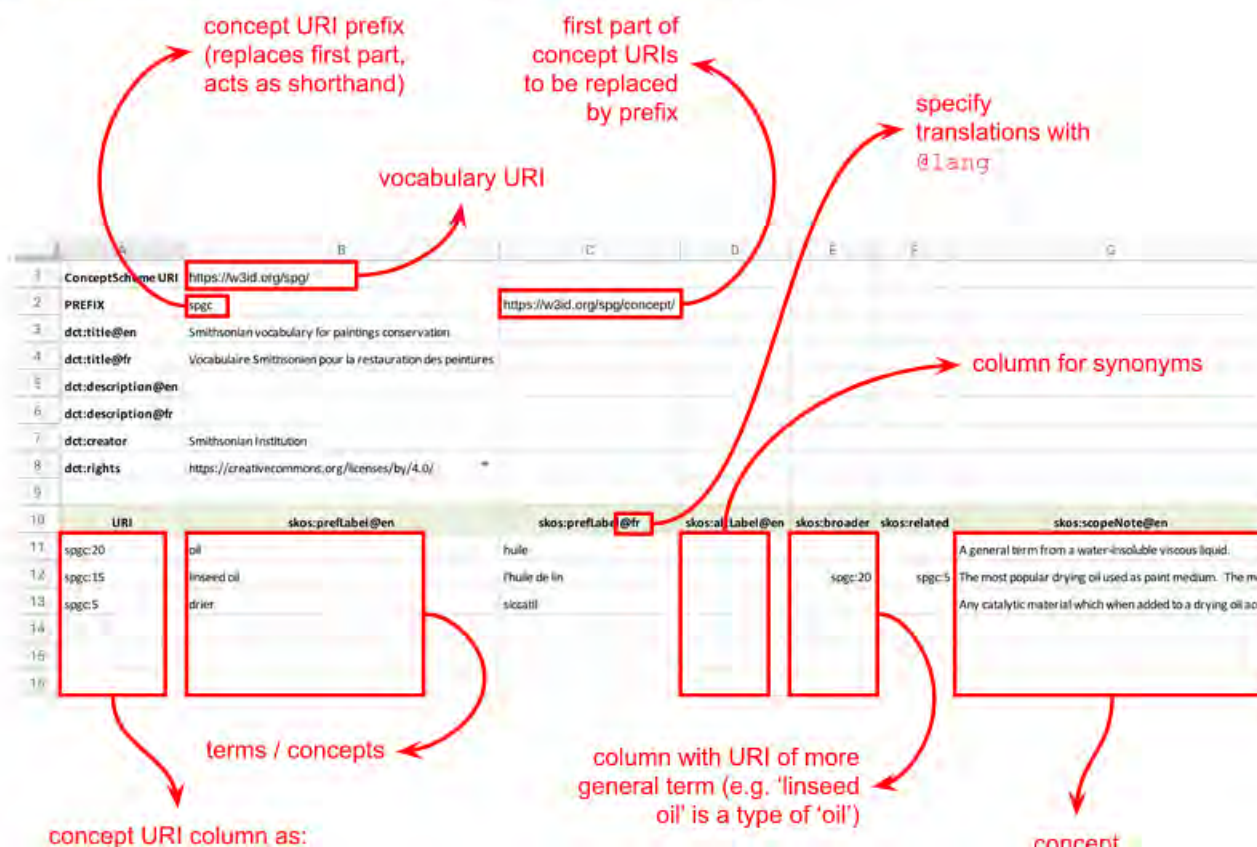
1. Ensure that you can share with an appropriate license as explained here.
2. Ensure you have noted the meaning of each of your terms and that there are no ambiguities (for example do not use the same term to mean two different things).
3. If you do not already have URIs for your terms, or if you are not sure what a URI is, contact the [repository maintainers](#) to help you produce them.
4. Use [this template](#) to enter your vocabulary data as explained here:

## Browse

- [Introduction](#)
- [Types of vocabularies](#)
- [Aim](#)
- [Hosting vocabularies in LCD](#)
- [Aligning vocabularies for LCD](#)
- [How to upload to the LCD repository](#)
- [Quick start](#)
- [Example process](#)
- [Repository maintainers](#)

## Clone this wiki locally

<https://github.com/linkedin-conser> 



The image shows a spreadsheet template for a vocabulary. Red arrows and boxes highlight key parts of the template:

- Concept URI prefix (replaces first part, acts as shorthand):** Points to the 'PREFIX' row (row 2) where the value is 'spgc'.
- first part of concept URIs to be replaced by prefix:** Points to the 'ConceptScheme URI' row (row 1) where the value is 'https://w3id.org/spgc/'.
- vocabulary URI:** Points to the 'PREFIX' row (row 2) where the value is 'spgc'.
- specify translations with @lang:** Points to the 'dct:title@en' and 'dct:title@fr' rows (rows 3 and 4).
- column for synonyms:** Points to the 'skos:prefLabel@fr' column header (row 10).
- terms / concepts:** Points to the 'URI' and 'skos:prefLabel@en' columns (rows 11-13).
- column with URI of more general term (e.g. 'linseed oil' is a type of 'oil'):** Points to the 'skos:broader' column header (row 10).
- concept:** Points to the 'skos:scopeNote@en' column header (row 10).

URI	skos:prefLabel@en	skos:prefLabel@fr	skos:alLabel@en	skos:broader	skos:related	skos:scopeNote@en
spgc:20	oil	huile				A general term from a water-insoluble viscous liquid.
spgc:15	linseed oil	/huile de lin		spgc:20	spgc:5	The most popular drying oil used as paint medium. The med
spgc:5	drier	siccant				Any catalytic material which when added to a drying oil acc



SKOS Play is a free application to render and visualise thesaurus, taxonomies or controlled vocabularies expressed in **SKOS**.

With SKOS Play you can print Knowledge Organization Systems that use the SKOS data model in HTML or PDF documents, and visualize them in graphical representations.

SKOS Play can be used :

- to *test and verify* a vocabulary during the conception phase
- to *exchange and communicate* the vocabulary when validating it with domain experts
- to *publish* it when it is shared on the web.

[Start Here](#)[Learn More](#)

SKOS Play also provides a way to convert Excel spreadsheets to **SKOS** files.

This makes it easy to produce SKOS files for taxonomies and authorities list.

The converter can even generate data in other RDF vocabularies.

[Convert Excel to SKOS](#)

🌟 Now you can also validate your SKOS data using the **SKOS Testing Tool**.

The SKOS Testing Tool is a frontend for **qSKOS** by Christian Mader.

[SKOS Testing Tool](#)

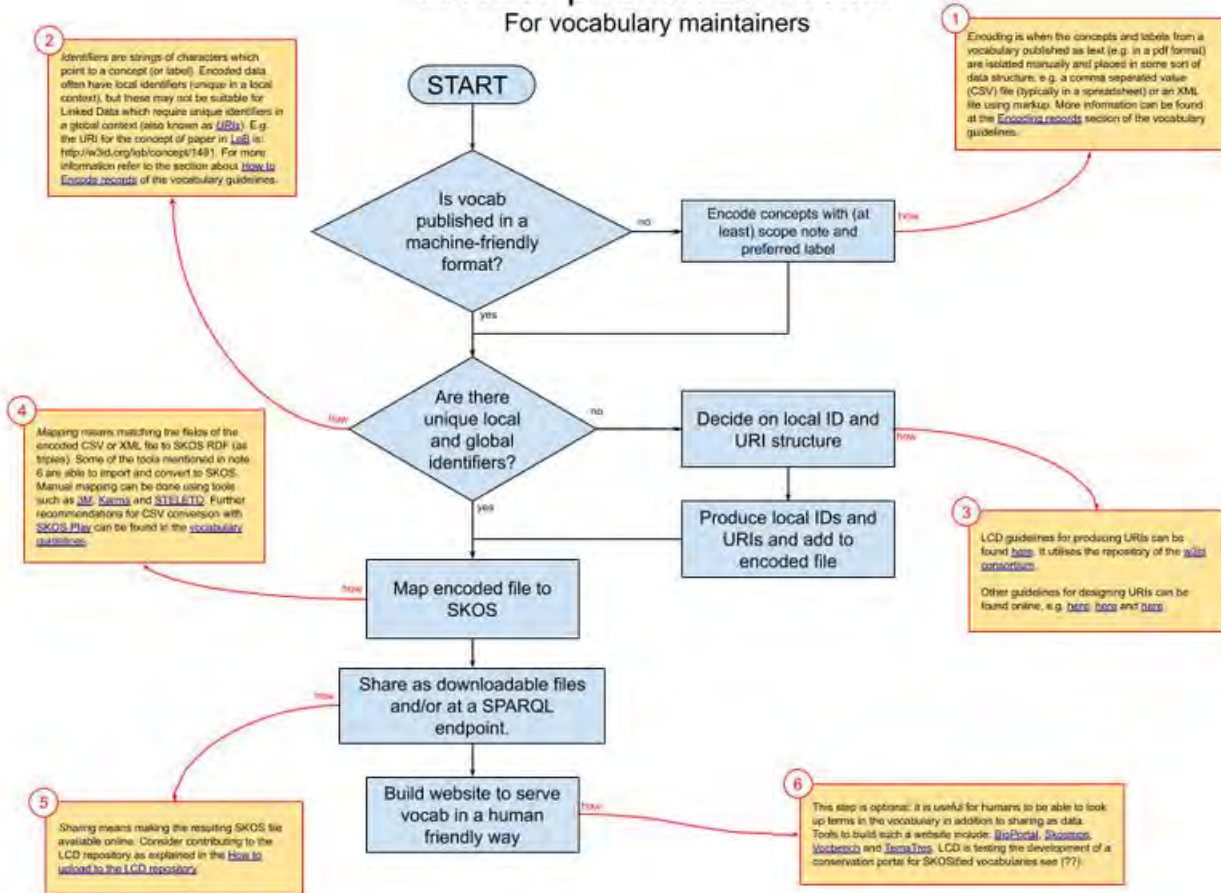
## Flowcharts

This document is based on work done by the LCD consortium during 2019. Please consult these flowcharts for easy reference.

## How to publish as SKOS

### How to publish as SKOS

For vocabulary maintainers

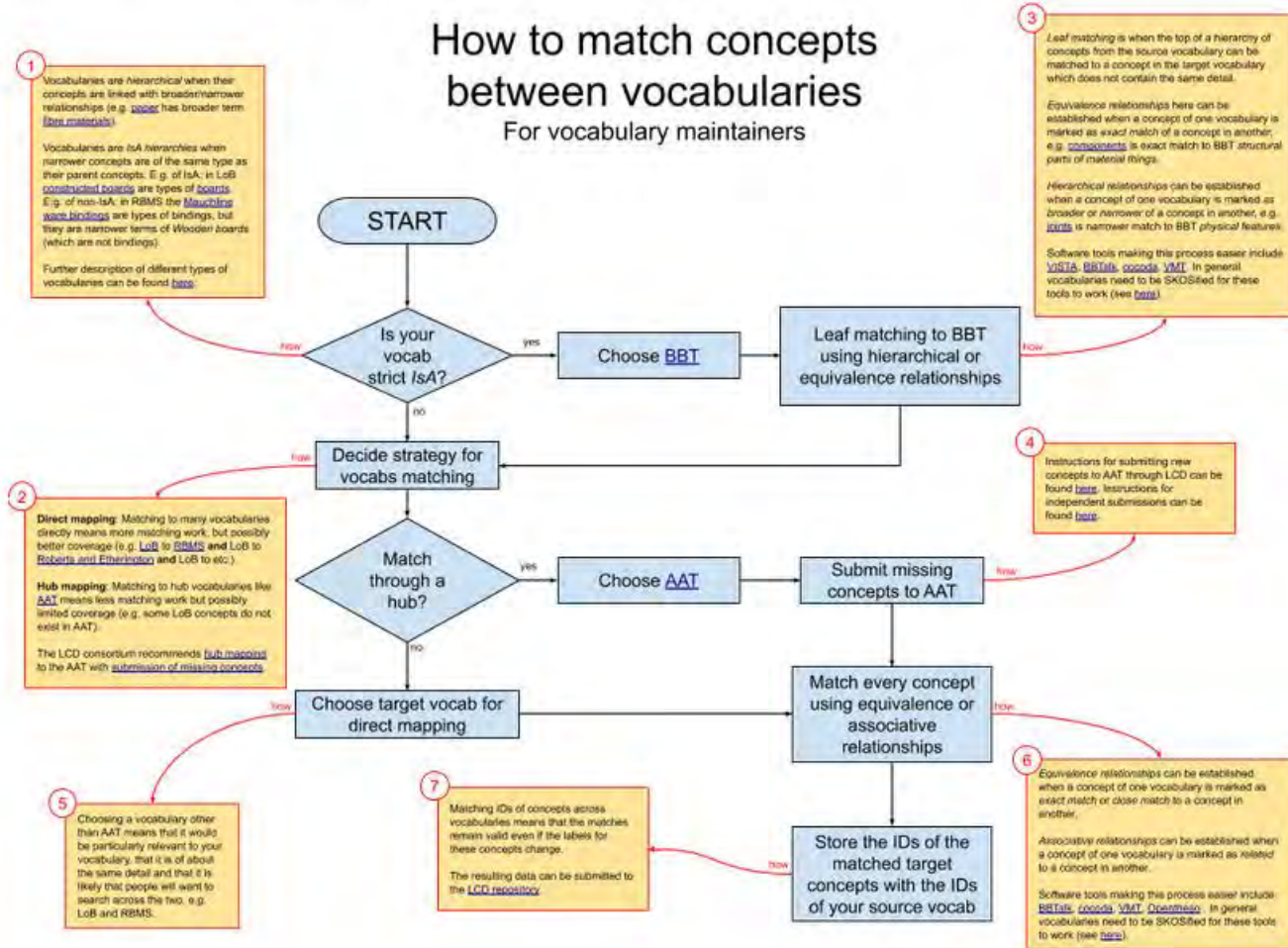


## How to match concepts

# How to match concepts

## How to match concepts between vocabularies

For vocabulary maintainers



paper

Term	Scope note	Vocabulary	Matched terms
<a href="#">paperboard</a>	Board made entirely of paper by whatever means (laminated, pulp, etc.). The term can be used where the boards of a book are completely covered and the means of their manufacture cannot be identified, but where it is clear that they are made of paper.	<a href="http://w3id.org/loc-vocab/">http://w3id.org/loc-vocab/</a>	<ul style="list-style-type: none"> <li>• <a href="http://vocab.getty.edu/aat/">paperboard</a> (<a href="http://vocab.getty.edu/aat/">http://vocab.getty.edu/aat/</a>)</li> <li>• <a href="http://w3id.org/lob/">paper board (materials)</a> (<a href="http://w3id.org/lob/">http://w3id.org/lob/</a>)</li> </ul>
<a href="#">paper</a>		<a href="http://w3id.org/bod-vocab/">http://w3id.org/bod-vocab/</a>	<ul style="list-style-type: none"> <li>• <a href="http://vocab.getty.edu/aat/">paper</a> (<a href="http://vocab.getty.edu/aat/">http://vocab.getty.edu/aat/</a>)</li> </ul>
<a href="#">Paper</a>	Refers generally to all types of thin matted or felted sheets or webs of fiber formed and dried on a fine screen from a pulpy water suspension. The fibers may be animal, such as hair, silk or wool, or mineral, such as asbestos, or synthetic. However most paper is made from cellulosic plant fiber, such as from wood pulp, grass, cotton, linen, and straw.	<a href="http://w3id.org/loc-vocab/">http://w3id.org/loc-vocab/</a>	<ul style="list-style-type: none"> <li>• <a href="http://vocab.getty.edu/aat/">paper</a> (<a href="http://vocab.getty.edu/aat/">http://vocab.getty.edu/aat/</a>)</li> </ul>
<a href="#">paper</a>	Refers generally to all types of thin matted or felted sheets or webs of fiber formed and dried on a fine screen from a pulpy water suspension. The fibers may be animal, such as hair, silk or wool, or mineral, such as asbestos, or synthetic. However most paper is made from cellulosic plant fiber, such as from wood pulp, grass, cotton, linen, and straw.	<a href="http://w3id.org/sul-vocab/">http://w3id.org/sul-vocab/</a>	<ul style="list-style-type: none"> <li>• <a href="http://vocab.getty.edu/aat/">paper</a> (<a href="http://vocab.getty.edu/aat/">http://vocab.getty.edu/aat/</a>)</li> </ul>

Clipboard

Details

## Some points to discuss:

Long term sustainability of resources for small communities  
Asking domain experts to engage with terminology at a technical level

# Thank you

Engagement with the NKOS community during and after the project was extremely helpful!

[athanasios.velios@english-heritage.org.uk](mailto:athanasios.velios@english-heritage.org.uk)