

Controlled Vocabularies Decision Tree Working Group

Samvera Partners
October 14-15, 2020

Julie Hardesty
Metadata Analyst, Indiana University

CV Decision Tree Working Group Members

Dawn Childress (UCLA)

Gretchen Gueguen (PALCI)

Julie Hardesty (Indiana University)

Rebecca Pattillo (University of Louisville)

Erik Radio (University of Colorado)

Lynette Rayle (Cornell University)

Scope and Objectives

Provide guidance for **selecting and using controlled vocabularies** behind descriptive metadata fields. Show how this guidance can be used with **Hyrax** and other software incorporating metadata fields where CVs can be helpful.

Deliverables and Timeframe

Define use cases for needing CVs behind metadata fields (in Hyrax / in general)

Identify sources/steps for locating controlled vocabularies

Identify decisions/steps necessary to evaluate the use of a controlled vocabulary

Hyrax-specific: Identify steps to take to add/change a controlled vocabulary behind a descriptive metadata field in Hyrax

Timeframe: Samvera Connect 2020

Decision Tree

Uses **levels of CV review, replacement, and creation** to discuss ways to evaluate and look for CV options

Hyrax examples include understanding CVs that come with Hyrax, editing those, and adding new CVs to metadata fields in Hyrax

Connects to **Controlled Vocabularies and Criteria spreadsheet** of CV options

Shared to samvera-metadata for review in September 2020

Controlled Vocabularies (CV) Decision Tree

September 2020

Dawn Childress, Gretchen Gueguen, Julie Hardesty, Rebecca Pattillo, Erik Radio, Lynette Rayle

Introduction

Discovery and retrieval in various digital platforms and repositories can be greatly enhanced by the use of controlled vocabularies (CV). Controlled vocabularies are sets of terms that can be used to ensure greater consistency among metadata records, as opposed to uncurated metadata values. This consistency allows a user to collocate, filter, and compare resources and facilitates their ability to find ones that they need. While commonly used for subjects, genres, and other categorical fields, CVs can also be used for rights statements, names, places, and other metadata fields for which more consistency can enable better system functionings and discovery.

While there are many well-established CVs in use (e.g. Library of Congress Subject Headings; Getty Art and Architecture Thesaurus), it is important to remember that CVs are works in progress, being continuously updated to better reflect the areas and disciplines they help describe. Similarly, CVs can also be a tool for marginalized communities who historically have not had the power to describe or shape their representation in the world. As such, the use of a particular CV and its impact on the description of resources in a given collection warrant evaluation to determine if the CV is a good match for a set of resources. See the accompanying [Controlled Vocabularies and Criteria](#) for CV options and consider adding any CVs you know about that are not listed.

Scope

This decision tree is meant to provide guidance for evaluating, selecting, and using controlled vocabularies for descriptive metadata fields. Using a tiered approach, the document explores three scenarios for using CVs. First, it discusses how to evaluate a CV already built into a given platform. Second, it looks at replacing a CV in a platform with another existing one. Third, it addresses when it is necessary to customize, modify, or build a local CV due to gaps or deficiencies in existing CVs.

Audience

Controlled Vocabularies and Criteria

List of CV options

Hoping for collaboration to add to list

Identifying domain (subject of CV),
date last update/maintenance
indicator, and who is responsible for
CV

Would like to include CVs created by
marginalized communities that
represent that community's interests
(LGBTQ, First Nations, Disabilities,
Women)

Controlled Vocabulary (CV)	Stable identifiers / URIs	Maintained by	Date last updated	Domain
Getty AAT	<input checked="" type="checkbox"/>	Getty	2020-09-14	Generic terms for art, architecture, and other material culture.
LCGFT	<input checked="" type="checkbox"/>	Library of Congress	2020-08-17	Thesaurus that combines both genres and forms, describing what a work is versus what it is about.
ISO639-2 Codes for Languages	<input checked="" type="checkbox"/>	Library of Congress	2017-12-21	Provides representation of names of languages, equivalent to MARC Language Codes.
creativecommons.org	<input checked="" type="checkbox"/>	Creative Commons	2020-09-23	Six different license types that offer a standardized way to answer the question, "What can I do with this work?"
Geonames	<input checked="" type="checkbox"/>	Multiple contributors	2020-09-25	Contains over 25 million geographical names and consists of over 11 million unique features whereof 4.8 million populated places and 13 million alternate names.
DCMI Type Vocabulary	<input checked="" type="checkbox"/>	DCMI Usage Board	2020-01-20	Included are the fifteen terms of the Dublin Core™ Metadata Element Set (also known as "the Dublin Core") plus several dozen properties, classes, datatypes, and vocabulary encoding schemes.
rightsstatements.org	<input checked="" type="checkbox"/>	Europeana, DPLA, National Digital Library of India, Library and Archives Canada, and Trove at the National Digital Library of Australia	2020-05-27	3 categories with 12 standardized rights statements to summarize the copyright status of objects in their collection, as well as how those objects may be used.
LCSH	<input checked="" type="checkbox"/>	Library of Congress	2020-08-17	Subject terms used by the Library of Congress and other libraries around the United States.
OCLC FAST	<input checked="" type="checkbox"/>			
MeSH	<input checked="" type="checkbox"/>			
Homosaurus	<input checked="" type="checkbox"/>	Homosaurus Editorial Board	2019-05	International linked data vocabulary of LGBTQ terms that supports improved access to LGBTQ resources within cultural institutions.

CVs and Criteria

Meta-Lists of Controlled Vocabularies

Lists of vocabularies	URL	Notes
List of LC vocabs	LC Linked Data Service	
List of Getty vocabs	Getty Vocabularies	
Rare Book and Mss vocabs	RBMS Vocabularies	
Temporal subject headings	PeriodO	for temporal subject headings - aggregates many time period CVs - can be a source itself, or a good place to find an appropriate CV
BARTOC registry	http://bartoc.org	
JISC Directory of Metadata Vocabularies	https://www.webarchive.org.uk/wayback/archive/20160101151732/http://www.jiscdigitalmedia.ac.uk/guide/controlling-your-language-links-to-metadata-vocabularies	
Metadata Map	http://jennriley.com/metadatamap	
Link to community developed resource for CVs created by underrepresented groups		

Thanks!

[CV Decision Tree Working Group](#)
Under the SMIG on the Samvera Wiki