



What is a Controlled Vocabulary?



Definition of a Controlled Vocabulary

A vocabulary is a set of terms (words, codes, etc.) that are used in a specific community. [Vocabularies \(/glossary/29#term11058\)](#) provide a mechanism for communication- be it written, oral or electronic- because the meaning of the terms are known and agreed upon by the community members. When a vocabulary is formally managed, it becomes a controlled vocabulary. In this case, "managed" means the terms are stored and maintained using agreed-upon procedures. Procedures should exist for adding terms, modifying terms and, more rarely, deprecating terms from a controlled vocabulary.

A controlled vocabulary is a collection of terms that are:

- Accepted: The term must adhere to community practices.
- Defined: The terms are precisely characterized. Typically, this means the terms have rigorous definitions.
- Managed: In general, there will be a body of experts that create and maintain the controlled vocabulary. The controlled vocabulary maintenance will involve periodic review, addition of new terms, modification of terms, and occasionally deprecation of terms.

Notice, this definition does not specify a particular scope of usage. [Controlled vocabularies \(/glossary/29#term11024\)](#) could be developed for a local project (like the Scripps Institution of Oceanography Geological Data Center), a broader community (e.g. OOSTethys), or as a part of a widely used [standard \(/glossary/29#term11040\)](#) or tool ([ISO \(/glossary/56#term11604\)](#) 19115).

Controlled Vocabulary Categories and Types

To many people, the English language is a well-known vocabulary. We have many ways of representing the terms in the English language. For example, if we want to figure out what a specific word means we might consult a [glossary \(/glossary/29#term11575\)](#); if we want to know the origin of a term we might consult a [dictionary \(/glossary/29#term11026\)](#); and if we want to know how the term relates to other terms we might consult a [thesaurus \(/glossary/29#term11055\)](#). We also need to recognize that the meaning of terms may change through time. Generations use terms in different ways (cool in one generation means a low temperature, while cool in another is a positive adjective).

To enable formal management, a controlled vocabulary can be organized in several ways. There are three broad categories of controlled vocabularies: flat, multi-level and relational.

- Flat controlled vocabularies provide a set of used terms. Some flat controlled vocabularies will provide additional information about each term.
- Multi-level controlled vocabularies build upon a flat controlled vocabulary by assigning each term to a category.
- Relational controlled vocabularies provide a set of terms, and capture how they are associated with each other.

Within these three categories, there are additional controlled vocabulary types. The table below summarizes these categories and types. The table categorizes necessary conditions only. Some controlled vocabularies will appear as "hybrids" of one or more categories of controlled vocabularies. Please see the [Types of Controlled Vocabularies \(/guides/vocabs/vocetypes/voccat\)](#) guide for a more extensive explanation, or this article on [Knowledge Organization Systems \(http://www.clir.org/pubs/reports/pub91/1knowledge.html\)](#).

Broad Category	Controlled Vocabulary Types	Description
Flat Controlled Vocabulary	Authority File (/glossary/29#term11570)	List of terms
	Glossary	List of terms and definitions within a specific domain
	Dictionary	List of terms, definitions, and additional information
	Gazetteer (/glossary/29#term11574)	List of place names
	Code List (/glossary/29#term11571)	List of codes (e.g. abbreviations) and definitions

Multi-Level Controlled Vocabulary	Taxonomy (/glossary/29#term11578)	Terms classified into categories
	Subject Heading (/glossary/29#term11577)	Terms classified into categories, which may be broad classes (/glossary/29#term11023)
Relational Controlled Vocabulary	Thesaurus	Set of terms and relationships (/glossary/29#term11047) among individual values (/glossary/29#term11041)
	Semantic Network (/glossary/29#term11579)	Set of terms/concepts and directed relationships
	Ontology (/glossary/29#term11044)	Set of terms and relationships among terms, enhanced by additional information provided by rules and axioms.

The Purpose of a Controlled Vocabulary

Controlled vocabularies can serve several different purposes. For example, a controlled vocabulary might help users find data (also known as a "[discovery vocabulary \(/glossary/29#term11027\)](#)"), or assist in the interpretation of data (also known as a "[usage vocabulary \(/glossary/29#term11057\)](#)"). The controlled vocabulary might provide human-understandable meaning (also known as a "semantic vocabulary") or [machine-readable \(/glossary/29#term11032\)](#) format information (also known as a "syntactic vocabulary"). Controlled vocabularies provide these abilities by:

- establishing the permissible terms to be used;
- maintaining the proper and agreed-upon spelling of the terms;
- clarifying terms for those who are new to the community; and
- eliminating the use of arbitrary terms that can cause inconsistencies and confusion.

Posted November 7th, 2008

SUGGESTED CITATION

[Neiswender, C. \(/taxonomy/term/11567\)](#) 2009. "What is a Controlled Vocabulary?." In *The MMI Guides: Navigating the World of Marine Metadata*. <http://marinemetadata.org/guides/vocabs/vocdef>. Accessed October 21, 2014.

[Have suggestions? Please give us your Feedback \[+ \]](#)