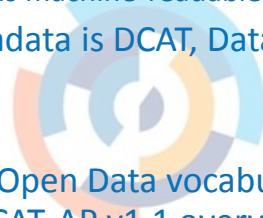


Important aspects while using controlled vocabularies

- A controlled vocabulary is a restricted list of words or terms used for labeling, indexing or categorizing. It is controlled because only terms from the list may be used for the subject area covered by the controlled vocabulary. Most controlled vocabularies also have some form of cross-references pointing from one or more “non-preferred” terms to the designated “preferred” term.
- Specific types of controlled vocabularies:
 - **Thesaurus:** A thesaurus is a more structured kind of controlled vocabulary.
 - **Taxonomy:** Has become a popular term now for any hierarchical classification or categorization system
 - **Ontology:** Set of concepts with attributes and relationships between the various concepts that contain various meanings, all to define a domain of knowledge, and is expressed in a format that is machine-readable.
- The controlled vocabulary used to describe metadata is DCAT, Data Catalogue Vocabulary
 - <http://www.w3c.org/TR/vocab-dcat/>
- The European Commission has created a Linked Open Data vocabulary specification called DCAT-AP. More information about DCAT-AP can be found in the dedicated DCAT-AP v1.1 overview.
- For more general information and training about metadata, please look at the following links:
 - <http://w3c.github.io/dwbp/bp.html#metadata>
 - <http://europeandataportal.eu/en/content/training-library/library>
 - <https://theodi.org/guides/marking-up-your-dataset-with-dcat>
- Examples of important controlled vocabularies are available on the next slide.



Most important vocabularies to look into

Licences:

 <http://publications.europa.eu/mdr/resource/authority/licence/html/licences-eng.html>

File types:

 <http://publications.europa.eu/mdr/resource/authority/file-type/html/filetypes-eng.html>

Data themes:

 <http://publications.europa.eu/mdr/resource/authority/data-theme/html/data-theme-eng.html>



Benefits of controlled vocabularies:

1. Clarification and specification of the language and terms being searched
2. Standardization of the terms being searched
3. Discoverability of data is enhanced
4. Comparability and consistency over datasets and across organisations