



Title	Strategy Report on Controlled Vocabularies (D4.2)
Work Package	WP4
Authors	Taina Jääskeläinen
Source	
Date	19 August 2009
Dissemination Level	PU (Public)

Summary/abstract

WP4 Recommendation for the cessda-ERIC regarding the use of controlled vocabularies in specified metadata elements.

1. WP4 policy recommendations regarding the use of controlled vocabularies

1.1 Obligations of cessda-ERIC members

Resource providers will translate the source controlled vocabularies, produced by WP4 and implemented by the cessda-ERIC, into their local language(s) within six months of entering Full or Designate Membership of the ERIC.

1.2 Recommendations for best practices

WP4 proposes that resource providers are strongly recommended to use controlled vocabularies for certain DDI elements, specified by the cessda-ERIC, for data published in the common portal, particularly for all new data published after entering Full or Designate Membership.

The Authentication and Authorisation requirements of the cessda-ERIC may require that the use of certain DDI elements and related controlled vocabularies becomes mandatory at the start of single-sign-on system.

Using all cessda-ERIC controlled vocabularies for data documentation may become mandatory after the cessda-ERIC metadata model has been decided.

WP4 also recommends using controlled vocabularies for other DDI elements as much as possible to increase harmonisation and comparability of data.

2.1 Reasons behind the WP4 recommendation:

Thesauri and other controlled vocabularies support consistent metadata by specifying terms that can be used for a particular metadata element. This consistent use of terms within metadata aids discovery of resources across organisations. The use of controlled vocabularies in indexing, documenting, or searching for resources by use of their assigned metadata will enhance:

- precision;
- consistency;
- temporal, spatial and topical comparability;
- semantic and technical interoperability;
- multilingual access and documentation;
- efficiency;
- harmonization.

Controlled vocabularies can be used to partly automate multilingual documentation.

2.3 Source controlled vocabularies

In close co-operation with the DDI Alliance Controlled Vocabularies group, WP4 has prepared draft source vocabularies for the following DDI elements:

Life Cycle Event Type
Commonality Type
Data Collection: Intended Frequency
Data Collection: Time Method
Data Collection: Mode of Collection
Response Unit
Aggregation method
Storage Type
Statistical Software Type
Character Set Type
Category Statistics
Summary Statistics
Date/Time Representation
Kind of Data Type

Currently under review are:

Variable Group Types
Question Group Types (element to be implemented in the next version of DDI3)
Analysis Unit
Contributor Type Roles
Variable Representation Roles

The decision which controlled vocabularies will be adopted for use within the cessda-ERIC will depend on the requirements of the authentication and authorisation procedures, and on metadata requirements for comparison, harmonisation and the Question Bank.