

Stand 20.01.2014

Allgemeine Hinweise:

- im Idealfall sollten Inhalte immer auch in Englisch angegeben werden, Abweichung von dieser Regel sind evt. sinnvoll bei Namen (Personen, Organisationen etc.)
- explizite Angabe der Sprache über xml:lang für nicht-englischen Inhalt
- möglichst vorhandene und weit verbreitete (community-spezifische) Vokabulare nutzen (Vorschläge für einzelne Facetten siehe unten)
- bei Verwendung existierender Vokabulare immer 1:1 Schreibweise aus dem Vokabular übernehmen
- bei Namen (Personen, Organisation etc.) grundsätzlich immer einheitliche und möglichst standardisierte Schreibweise nutzen
- möglichst immer Einzelwerte statt mehrere Werte in ein Element (z.B. statt einem Language-Element mit Inhalt "deu, eng", lieber zwei Elemente)
- verschiedene Elemente sollten zueinander orthogonal sein, d.h. sie sollten sich inhaltlich möglichst nicht überschneiden
- unspezifische Angaben wie "multilingual" oder "multimodal" sollten vermieden werden, stattdessen sollten alle zutreffenden Werte benannt werden
- Werte die hauptsächlich für die facetiierte Suche gedacht sind, sollten im Idealfall möglichst kurz sein (einzelne Worte oder kurze Wortgruppen)

Vorschläge für nützliche Vokabulare:

Facette	Vokabulare	Empfohlene CMDI-Komponente
Collection	-	-
Language	Language Codes nach ISO 639-3 (Sprachen) bzw. ISO 639-5 (Sprachfamilien) bzw. entsprechender englischer Name	cmdi-language ¹ (clarin.eu:cr1:c_1271859438111), iso-language-639-3 ¹ (clarin.eu:cr1:c_1271859438110), iso-languagefamily-639-5 ¹³ (clarin.eu:cr1:c_1271859438108)
ResourceClass	Stark abhängig von der Facetten-Definition: Dublin Core Type (DCMI Type Vocabulary), DC-3806, OLAC Linguistic Data Type	-
Modality	NaLiDa Modalities ² (DC-2490 ²)	ModalityInfo (clarin.eu:cr1:c_1290431694513)
Continent	Codes entsprechend Geonames Konvention (Geonames Webservices) bzw. entsprechender englischer Name (DC-2531 ¹²)	iso-continent ¹ (clarin.eu:cr1:c_1271859438105)
Genre	-	-
Country	Codes nach ISO 3166-1 bzw. englischer Name entsprechend Standard	iso-country ¹ (clarin.eu:cr1:c_1271859438104)
Subject	-	-

Facette	Vokabulare	Empfohlene CMDI-Komponente
Organisation	-	-
Format	Beliebiger MIME-Type	cmdi-mimetype (clarin.eu:cr1:c_1271859438106)
Keywords	-	-
Year (bzw. beliebige Zeitangaben)	W3C DateTime	-

Vorschläge für Vokabulare zu evt. zukünftigen Facetten:

Facette	Vokabulare	Empfohlene CMDI-Komponente
LiveCycleStatus	DC-3818	-
Time Coverage	Open Date Range Format	-

¹ Definition verwendeter ISOcat-Kategorie(n) problematisch

² Definition nur über CMDI Komponente (=kein festes Vokabular in ISOcat)

³ Unvollständige Definition der Komponente (fehlende Conceptlinks)