

General Guidelines:

- Content **MUST** be provided in English, exceptions **MAY** be acceptable in case of names (persons, organisations etc.)
- Attribute *xml:lang* **MUST** be used for all non-English content
- Existing and widely-used vocabularies **SHOULD** be used
- for used vocabularies the exact spelling specified in the vocabulary **MUST** be used
- For names (persons, organisations etc.) standardised variants **SHOULD** always be used
- Atomic values **MUST** be used (instead of a single 'deu, eng' → 'deu', 'eng')
- Different values **SHOULD** be orthogonal, they **SHOULD** not have semantic overlap
- Unspecified values **SHOULD** be avoided (e.g. instead of 'multilingual' give all valid values)
- Values supposed to be used for faceted search **SHOULD** be relatively short (i.e. single words, short phrases)
- for some facets (like genre, subject, keywords) different vocabularies **MAY** be used, which **SHOULD** be used at least consistently in the same collection

Useful vocabularies:

Facet	Vokabulary	Useful CMDI Component
Collection	-	-
Language	ISO 639-3 (language) or ISO 639-5 (language families) code or corresponding English 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	DC-3806	-
Modality	NaLiDa Modalities ²	ModalityInfo (clarin.eu:cr1:c_1290431694513)
Continent	Geonames convention (Geonames Webservices) or corresponding English name (DC-2531 ¹²)	iso-continent ¹ (clarin.eu:cr1:c_1271859438105)
Genre	-	-
Country	ISO 3166-1 code or their corresponding English names Jörg Knappen	iso-country ¹ (clarin.eu:cr1:c_1271859438104)
Subject	-	-
Organisation	-	-
Format	all MIME-Types	cmdi-mimetype (clarin.eu:cr1:c_1271859438106)
Keywords	-	-
Year (bzw. beliebige)	W3C DateTime	-

Facet	Vocabulary	Useful CMDI Component
Zeitangaben)		
LiveCycleStatus	DC-3818	-
Time Coverage	Open Date Range Format	-

¹ Definition of used ISOcat category problematic

² Definition only in CMDI component (=not a complex/closed datcat in ISOcat)

³ Incomplete definition in component (missing conceptlinks)