

CIDOC CRM mapping

Seta Štuhec, Ksenia Zaysteva, Edeltraud Aspöck

Ontology

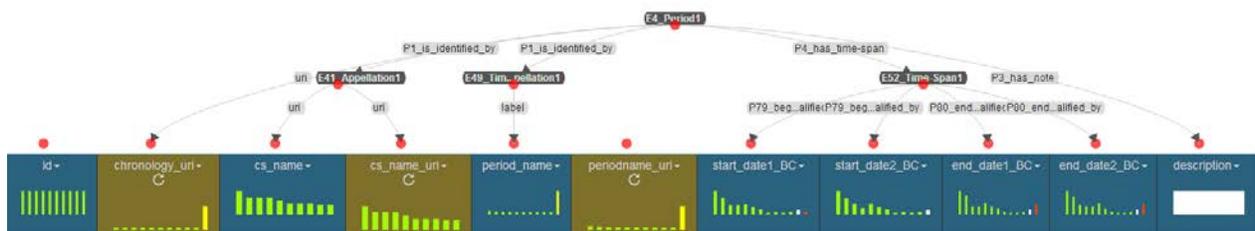
The word “ontology” is used in many communities and therefore has several different meanings. Most commonly it is used in philosophy, where it refers to the discipline dealing with the *nature* and *structure* of “reality”, independently of any further consideration (even actual existence). In computer science ontologies are means to formally model the structure of a system using concepts and relations between them (Guarino, Oberle, & Staab, 2009). [CIDOC CRM](#) is such a reference work for concepts and relations built to describe the world of cultural heritage. Dataset that has been mapped to CIDOC CRM framework is semantically enriched and furthermore it can be (at least theoretically) integrated with other datasets also mapped to the same ontology. This way the problem of interoperability and integration between various complementary information from different systems can be efficiently addressed (Doerr, 2009).

Mapping

Mapping of DEFC app – or any other relational database for that matter – to CIDOC CRM is no straightforward task. The current version (6.2.2.) has 102 classes and 193 properties excluding CIDOC CRM extensions such as CRMarchaeo and CRMsci, which have to be understood well, before one can start with the mapping.

To map a dataset to CIDOC CRM ontology basically means that each database field has to be assigned to a CIDOC CRM class and relationships between them have to be expressed with CIDOC CRM properties. The definition of a class is described in scope notes. Furthermore, each class has a certain domain and range of properties, which means not all properties can be used with all classes. However, since classes are hierarchically ordered, each subclass inherits all properties (relationships as well as the semantic meaning) of its upper class.

For example, if we have a table describing different properties of archaeological periods, the main concept would be mapped as E4_Period, which for example P1_is_identified_by E49TimeAppellation (period name), P7_took_place_at E44_PlaceAppellation (region where this dating system is in use), P4_has_time_span E52_Time_Span (beginning and the end dates)...



Sometimes the path from one class to the other is not so straightforward and leads through many empty nodes before we “arrive to the destination” class. For example let’s say there is field in the database describing the type of artefact that was found in a settlement this table is describing. Then we can map the settlement to the E27_Site and create a path of empty nodes through S4_Observation to E24_Physical_Man_Made_Thing that finally has a E55_Type which describes the type of the artefact.

Rather often it is difficult to say, whether the mapping has been done “correctly” or not, because it depends on what we would like to express, what is relevant, and what detail we need – are we going to use a shortcut (e. g. say that an objects has dimension x) or do we need a longer path to describe the whole process that lead to that specific result (e. g. say that an object has been observed, measurements taken – when, by who etc.). However, as already discussed by the CIDOC CRM community, some sort of quality control will be needed in the future.

For more detailed information about DEFC app mapping check the [mapping documentation](#).

Data transformation

Once the dataset has been mapped to the ontology it is ready to be transformed into RDF (Resource Description Framework) triples. An RDF triple comprises three components: subject, predicate and object (from the above example: an artefact is a subject, has_dimension is a predicate and the dimension is the object). Data stored in this way can be queried using SPARQL Protocol.

References

- Doerr, M. (2009). Ontologies for Cultural Heritage. In S. Staab & R. Studer (Eds.), *Handbook on Ontologies* (pp. 463-485). Berlin; Heidelberg: Springer.
- Guarino, N., Oberle, D., & Staab, S. (2009). What Is an Ontology? In S. Staab & R. Studer (Eds.), *Handbook on Ontologies* (pp. 1-17). Berlin; Heidelberg: Springer.