CRMEX Position Paper: mapping patterns for CIDOC CRM

Douglas Tudhope, Ceri Binding

Different mappings and different implementation primitives can potentially pose significant problems for semantic interoperability, as discussed in our workshop paper. It would be useful if we could provide more examples and guidelines. Examples of some possible mapping issues from our experience include

- Should an E57 Material (e.g. *gold)* be mapped as a property of an E11 Modification event or as a property of an E22 Man-Made Object?
- Should a method of manufacture (e.g. *hammered*) be mapped as an E55 Type of an E12 Production event or as an Appellation of an E29 Design or Procedure?
- Should E22 Man-Made Objects be directly identified by an E42 Identifier or should the connection be made via a record that has an Identifier?
- All CRM classes can be assigned types (used for domain terminology) – any guidelines for default choices?
- When is it appropriate to create an assignment event when assigning an attribute to an object? Essentially this depends whether the decision to assign an attribute is considered worthy to record. Again this can result in different mapping expressions depending on the judgement.
- How to deal with short cuts in a consistent way and link a shortcut with its underlying path?

The potential to employ reasoning over the CRM graph is one of the reasons for semantic integration. Nonetheless in our view, a multiplicity of approaches for similar data will pose unnecessary problems for implementation in the medium term. Specific rules will probably be required, which raises difficulties for generalising and introducing a new alternative mapping. A pragmatic approach is to combine developments in reasoning with efforts at consensus on patterns for CRM mappings and guidelines. This could involve patterns for particular domains and also general patterns for common situations.

Issues

If different implementations of the CRM follow different low level implementation specifications or employ different mappings for the same underlying semantics then this raises barriers for semantic interoperability.

Working from established RDF patterns guarantees the semantic interoperability of the resultant data and also that the syntactical implementation details are handled consistently. It is also more friendly to non-specialists and can make it easier to express datasets via CIDOC CRM.

- Agreement on implementation details (e.g. primitives, namespace, definitive URIs)?
- Agreement on mapping patterns and guidelines?
- Desirability of expressing the end-purpose or use cases of a mapping exercise?
- Provision of appropriate registries of mapping patterns?
- Provision of core metadata for mapping patterns?