POZNAN
SUPERCOMPUTING AND NETWORKING CENTER
Simple Visualization of Structures of Interrelated Concepts in the FRBRoo Ontology

Krzysztof Sielski, Marcin Werla

TPDL 2013
Foreword: the Knowledge Base

Part of the SYNAT project financed by the National Center for Research and Development in Poland*

Aim:
Integrate data from heterogeneous sources such as:
• digital libraries and museums,
• bibliographic union catalogues,
• inventory databases,
• archives

* (grant no SP/I/1/77065/10, funding period: 2010 – 2013)
Foreword: the Knowledge Base

Source data:
- Union Catalog MARC XML
- Digital Libraries DC
- National Museum in Warsaw MONA
- National Museum in Krakow CDWA LITE

Aggregator (Clepsydra):
- Semantic cleaning and normalization (e.g. dates)

Mapping to ontology FRBRoo (jMet2Ont)

Relation Detection

Knowledge Base

Enrichment

Auxiliary Data Sources:
- Subject Headings
- Geonames
- VIAF
- (LOD)
Knowledge Base statistics

- Over 3,100,000 cultural heritage objects
- 535,602,864 RDF triples (including 234,530,568 explicit)
- 60,644,249 instances of FRBRoo concepts (URIs + bnodes):
  - 3,133,629 works,
  - 1,717,455 persons,
  - 507,907 legal bodies,
  - 890,252 subject headings
- RDF database engine: Owlim SE
Entailment rules

Very reduced rule set, e.g. no full type materialisation because of complex type hierarchy in FRBRoo

- e.g. efrbroo:F24_Publication_Expression is a subclass of 12 other FRBRoo classes + owl:Thing + rdfs:Resource + 13 instances of owl:Restriction (in implementation by Erlangen) = 27 classes
Entailment rules

Very reduced rule set, e.g. no full type materialisation because of complex type hierarchy in FRBRoo

- e.g. efrbroo:F24_Publication_Expression is a subclass of 12 other FRBRoo classes + owl:Thing + rdfs:Resource + 13 instances of owl:Restriction (in implementation by Erlangen) = 27 classes

- We can query for abstract classes in SPARQL simply by using a/rdfs:subClassOf

```sparql
PREFIX rdfs:<http://www.w3.org/2000/01/rdf-schema#>
PREFIX ecrm:<http://erlangen-crm.org/current/>

SELECT * {
  ?actors a/rdfs:subClassOf ecrm:E39_Actor
}
```
P14.1_in_the_role_of problem

- property of a property instance in RDF/OWL?
- No implementation of such property in ECRM
P14.1_in_the_role_of problem

- property of a property instance in RDF/OWL?
- No implementation of such property in ECRM
- Our workaround by introducing a new class: \texttt{F28a\_Contribution} – a subclass of \texttt{F28\_Expression\_Creation}
Explore the Knowledge Base contents (1)

- a raw SPARQL endpoint, which is aimed at expert users who know the ontology very well and have precisely defined goals
Explore the Knowledge Base contents (2)

- a geographical search application, which allows user to select an area on a map to find all objects connected with places contained in that area (e.g. all publications whose subject is a particular city)
Explore the Knowledge Base contents (3)

- a full text search application, which searches for keywords provided by user in RDF literals from the triplestore and uses the Query Processing Module (QPM) which maps on-the-fly information represented in the FRBRoo ontology to a simplified model, consisting of a small number of concepts: works, items, persons, places, legal bodies, and subjects
Explore the Knowledge Base contents (4)

- an application to explore semantic database with dynamically fetched portions of data describing particular object (RDF Units), which are presented as FRBRoo concepts in a legible way understandable by non-experts.

- RDF Units are graphs which consist of several ontology objects of different classes that are needed to provide all the essential information about a certain resource. For example, an RDF Unit for a particular instance of Publication Expression from the Knowledge Base would include objects representing its Title, Publication Event and Place of Publishing, but not geographical coordinates of that place.
Explore the Knowledge Base contents
RDF Unit simplification rules

[E21_Person]  P100_i_died_in  [E69_Death]
  P4_has_time_span  [E52_Time-Span]
  P1_is_identified_by  [?]
  \rightarrow date of death

[F18_Serial_Work]  P148_has_component  [F14_Individual_Work]
  \rightarrow series element

[?]  P9_consists_of  [F28a_Contribution]
  P14_carried_out_by  [?]
  \rightarrow contributor

[?]  P9_consists_of  [F28a_Contribution]
  P2_has_type  [?]
  \rightarrow in the role of

______

[?]  \text{―}  \text{stands for any class}