Discovering Ontologies And Instance Data

RareMat Meeting | 2018.01.11 | Columbia University
Why is reuse important?

Without reuse of shared models and reuse of instance data from those shared models, there is no “linked data”.
Not to mention... ontology engineering is expensive. :)
Reuse: Identify your collaborators and their data models.

- What data do you want to consume or integrate with?
  - Is the modeling for those datasets good enough to meet your prioritized use cases?
  - Or do you need to extend it?
  - Or engage to influence changes for the better?
Reuse: SKOS is Ubiquitous in GLAM’s

- Central is the SKOS is the skos:Concept
  - conceptualizations in a particular scheme to build thesauri (similar semantics as Authority Records)
- Strategically limited semantics
  - broader/narrower
  - pref/alt labels
  - alignment properties, like skos:closeMatch
  - skos:inScheme
Tools

● Example of an Instance Data Discovery Tool: Lotus from the LOD Laundromat

● Examples of Ontology Discovery Tools:
  ○ Linked Open Vocabularies (LOV, cross domain)
  ○ Bioportal (began with a Biomedical Focus, but expanding)
Definitions shared on the web

http://lov.okfn.org/dataset/lov/
Navigating LOV: Searching for Vocabularies

VOCABS

family

plink - PersonLink Ontology
http://vocabs.cnam.fr/ontologies/PersonLink.owl
A Multilingual and Multicultural Ontology Representing Family Relationships. @en

prov - W3C PROVance Interchange
http://www.w3.org/ns/prov#
The namespace name http://www.w3.org/ns/prov# is intended for use with the PROV family of documents that support the interchange of provenance on the web. @en

rdac - RDA Classes
http://rdaregistry.info/Elements/c/
The Classes element set consists of classes representing the RDA entities, including Work, Expression, Manifestation, Item, Person, Family, Corporate Body, and Agent. @en

xhv - XHTML Vocabulary
http://www.w3.org/1999/xhtml/vocab
This is a vocabulary definition utilized by XHTML. Family modules and document types using XHTML. Metadata, including HTML, Role and XML. RDFa can

Type
- vocabulary (7)
  - property/class >
  - agent >

Tag
- FRBR (2)
- People (2)
- API (1)
Navigating LOV: Searching for Specific Terms
Navigating LOV: Searching for Ontologies by Agents

Library of Congress

has role in vocab bf, mrel, mads, premis
Navigating LOV: Faceting
Navigating LOV: Context for Decision Making
Navigating LOV: Context for Decision Making
As an Ontology Owner: Registering New Ontologies

- [http://lov.okfn.org/dataset/lov/suggest](http://lov.okfn.org/dataset/lov/suggest)
  - Supply an ontology URI that LOV can then index

  - Guidelines for metadata *about the ontology*
    - Title
    - Version URIs, version information, modifications
    - Creators, rights, licenses

  - New initiative coming out of LD4P
    - Based on same infrastructure as BioPortal
Readings on LOV & Reuse

Discovery/Evaluation Ontologies for Reuse

- non-Academic description of Linked Open Vocabularies:
- Academic Article on LOV:
- Very Academic Article on Ontology Reuse (available through Hollis):
  - Survey on Common Strategies of Vocabulary Reuse in Linked Open Data Modeling:
    - [http://link.springer.com/chapter/10.1007/978-3-319-07443-6_31](http://link.springer.com/chapter/10.1007/978-3-319-07443-6_31)

Discovering Linked Data (Instance Data) for Reuse

- From DLib’s In Brief section: Identifier Hubs: OCLC Launches New Person Entity Lookup Pilot: [http://www.dlib.org/dlib/november15/11inbrief.html](http://www.dlib.org/dlib/november15/11inbrief.html)
Reuse isn’t easy

Reasons one might not reuse from an existing model

- A definition isn’t what is needed, often a judgement call
  - Terms might have a looser definition than is desired… possible candidate for reuse with an application profile.
  - Properties may have unwanted entailments… some are more critical than others
- Lack of trust in the sustainability of the ontology maintainers
- Lack of trust in the versioning practices by ontology maintainers

Discovering Linked Data (Instance Data) for Reuse

- From DLib’s In Brief section: Identifier Hubs: OCLC Launches New Person Entity Lookup Pilot: http://www.dlib.org/dlib/november15/11inbrief.html
Ontology Design Patterns (ODPs)

Plan A: Reuse

Plan B: Follow established Ontology Design Patterns

- ODP’s- effort to establish modularized *patterns* for reuse to enable easier alignment with other models and combat anti-patterns
- Lack of trust in the sustainability of the ontology maintainers
- Lack of trust in the versioning practices by ontology maintainers

Related Links

- Website: [http://ontologydesignpatterns.org/wiki/Main_Page](http://ontologydesignpatterns.org/wiki/Main_Page)
- ODP google group, [https://groups.google.com/forum/#!forum/ontology-design-patterns](https://groups.google.com/forum/#!forum/ontology-design-patterns)
Ontology Design Patterns (ODPs)


  These are lists for available ODP catalogues.

**Submissions**

This area aims at collecting Ontology Design Pattern proposals from ODP users.

After the author has finished the submission and asked for a review, the proposals are assigned to at least two members of the ODP Quality Committee, who are expected to provide a review.

Positive reviews can be accompanied with guidelines for fixing possible problems of the proposed Content ODP.

Once such problems have been addressed, the proposed Content ODP can be certified and published in the official catalogue.

See the submissions list:

- Content ODPs
- Reengineering ODPs
- Alignment ODPs
- Logical ODPs
- Architectural ODPs
- Lexico-Syntactic ODPs

<table>
<thead>
<tr>
<th></th>
<th>Catalogue</th>
<th>Submissions</th>
<th>All</th>
</tr>
</thead>
<tbody>
<tr>
<td>Content ODPs</td>
<td>0</td>
<td>147</td>
<td>147</td>
</tr>
<tr>
<td>Reengineering ODPs</td>
<td>0</td>
<td>12</td>
<td>12</td>
</tr>
<tr>
<td>Alignment ODPs</td>
<td>0</td>
<td>14</td>
<td>14</td>
</tr>
<tr>
<td>Logical ODPs</td>
<td>0</td>
<td>18</td>
<td>18</td>
</tr>
<tr>
<td>Architectural ODPs</td>
<td>0</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Lexico-Syntactic ODPs</td>
<td>0</td>
<td>20</td>
<td>20</td>
</tr>
</tbody>
</table>
Content ODPs


These are lists for available ODP catalogues.

### Submissions

This area aims at collecting Ontology Design Pattern proposals from ODP users.

After the author has finished the submission and asked for a review, the proposals are assigned to at least two members of the ODP Quality Committee, who are expected to provide a review.

Positive reviews can be accompanied with guidelines for fixing possible problems of the proposed Content ODP.

Once such problems have been addressed, the proposed Content ODP can be certified and published in the official catalogue.

See the submissions list:

- Content ODPs
- Reengineering ODPs
- Alignment ODPs
- Logical ODPs
- Architectural ODPs
- Lexico-Syntactic ODPs

<table>
<thead>
<tr>
<th>ODP Type</th>
<th>Submissions</th>
<th>All</th>
</tr>
</thead>
<tbody>
<tr>
<td>Content ODPs</td>
<td>0</td>
<td>147</td>
</tr>
<tr>
<td>Reengineering ODPs</td>
<td>0</td>
<td>12</td>
</tr>
<tr>
<td>Alignment ODPs</td>
<td>0</td>
<td>14</td>
</tr>
<tr>
<td>Logical ODPs</td>
<td>0</td>
<td>18</td>
</tr>
<tr>
<td>Architectural ODPs</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Lexico-Syntactic ODPs</td>
<td>0</td>
<td>20</td>
</tr>
</tbody>
</table>
ODP Example: Parts

Graphical representation

Diagram

General description

Name: part of
Submitted by: Valentina Presutti
Also Known As: part whole
Intent: To represent entities and their parts.
Domains: Parts and Collections

Competency Questions:
- What is this entity part of?
- What are the parts of this entity?

Solution description:

Reusable OWL Building Block:

Consequences:

Scenarios:

Brain and heart are parts of the human body, substantia nigra is part of brain.

Known Uses:

Web References:

Other References:

Examples (OWL files):

- http://www.ontologydesignpatterns.org/ocp/owl/partOf.owl (483)
- http://www.ontologydesignpatterns.org/ocp/examples/partOfHumanBodyParts.owl (6)
Questions and Answers?

Tour of LOV and ODP’s?

Both?