A constraint language for Description Set Profiles

Thomas Baker
Dublin Core Metadata Initiative

2013-09-04: Session on Profiles and Ontologies, DC-2013
2013-09-11: RDF Validation Workshop, Cambridge, MA
“Dublin Core” as a Record Format

• **1995** workshop goal “simple metadata record”
• **2001** XML format for Open Archives Initiative
• **Today** many still see DC as an exchange format
“Dublin Core” as an RDF Vocabulary

- **1997** RDF organizers at DC Canberra workshop
- **1999** RDF Rec addresses DC requirements
  - DCMI Metadata Terms published as RDF properties
- **2006** Top-10 vocabulary in “Linked Data cloud”
## From Record Elements to alignment with RDF

<table>
<thead>
<tr>
<th>1995</th>
<th>1997</th>
<th>2001</th>
<th>2007</th>
<th>RDF</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Element</strong></td>
<td><strong>Element</strong></td>
<td><strong>Property</strong></td>
<td><strong>rdf:Property</strong></td>
<td></td>
</tr>
<tr>
<td><strong>Qualifier</strong></td>
<td><strong>Element Refinement</strong></td>
<td><strong>Property</strong></td>
<td>(rdfs:sub PropertyOf)</td>
<td></td>
</tr>
<tr>
<td><strong>Encoding Scheme</strong></td>
<td><strong>Syntax Encoding Scheme</strong></td>
<td><strong>rdf:Datatype</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Vocabulary Encoding Scheme</strong></td>
<td><strong>skos:Concept Scheme</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

- **Eleamon Element**
- **Element Qualifier**
- **Element Refinement**
- **Property**
- **rdf:Property**
- **(rdfs:sub PropertyOf)**
- **rdf:Datatype**
- **skos:Concept Scheme**

---

Wednesday, September 4, 2013
Application Profiles

• **2000**: Customize Dublin Core for specific uses.
  – Mix-and-match terms from different standards

• **But problems:**
  – Implemented incompatibly in HTML, XML, RDF...
Abstract Model (DCAM)

• **2003-2007** Development of DCAM
  – Common target for mapping from HTML, XML...
  – Mapped to RDF to the extent possible
Bridging two mindsets

• Orientation to *Record Formats*
  – Bounded sets of fields to be “filled in” with information

• Orientation to *Graphs*
  – Unbounded webs of information connected by statements
Viewing a graph in terms of a “record”
Seeing components that can be validated.
Abstracting those components to a generalized model.
Grouping constructs with no equivalent in RDF.
Heuschrecken brauchen ökologische Ausgleichsflächen

Subject | Predicate | Object
--- | --- | ---
agris:CD2001000179 | dct:subject | agrovoc:c_4416k
agris:CD2001000179 | dct:title | "Heuschrecken..."@de
agris:CD2001000179 | dct:creator | :PB
:PBS | foaf:name | "Peter, B."
Heuschrecken brauchen ökologische Ausgleichsflächen

Peter, B.

Described Resource URI

Value URI

Value String

Expressed as triples

Subject | Predicate | Object
--------|----------|-------
agris:CD2001000179 | dct:subject | agrovoc:c_4416k
agris:CD2001000179 | dct:title | "Heuschrecken..."@de
agris:CD2001000179 | dct:creator | :PB
:PBS | foaf:name | "Peter, B."
2008: Description Set Profile
Constraint Language (DSP)

Description Set [template]
Description [template]
Statement [template]
  Property [constraint]
    <http://purl.org/dc/terms/subject>
  VocabularyEncodingSchemeURI [constraint]
    <http://aims.fao.org/aos/agrovoc>

Statement [template]
  Property [constraint]
    <http://purl.org/dc/terms/title>
  MinOccurs [constraint]
    1
  MaxOccurs [constraint]
    1

Statement [template]
  Property [constraint]
    <http://xmlns.com/foaf/0.1/creator>

Description [template]
  Resource Class [constraint]
    <http://xmlns.com/foaf/0.1/Person>
  Statement [template]
    Property [constraint]
      <http://xmlns.com/foaf/0.1/name>

- Data using this Description Set Profile describes:
  Resource with *exactly one title*,
  Value of *subject* is an *AGROVOC URI*.
  Authors are members of the class *Person* and have *names*. 

Wednesday, September 4, 2013
Templates and Constraints

• **Description Template constraints**: Identifier, Standalone, Minimum occurrence, Maximum occurrence, Resource class membership

• **Statement Template constraints**: Minimum occurrence, Maximum occurrence, Type
  – Property constraints: Property list, Sub-property
  – Literal Statement Template constraints:
    • Literal value constraints: Literal list, Literal language, Literal language list, Syntax Encoding Scheme, Syntax Encoding Scheme list
  – Non-literal Statement Template constraints:
    • Non-literal value constraints: Description template reference, Class membership
      – Value URI constraints: Value URI occurrence, Value URI list
      – Vocabulary encoding scheme (VES) constraints: VES occurrence, VES list
      – Value string constraints: Minimum occurrence, Maximum occurrence (plus all other constraints that apply to literal values - see above)
Broad humor and bitter irony collide in this fictional autobiography of Rabo Karabekian, who, at age seventy-one, wants to be left alone on his Long Island estate with the secret he has locked inside his potato barn. But then a voluptuous young widow badgers Rabo into telling his life story—and Vonnegut in turn tells us the plain, heart-hammering truth about man’s careless fancy to create or destroy what he loves.

"hasInstance": "http://bibfra.me/instance/u2-1"
Broad humor and bitter irony collide in this fictional autobiography of Rabo Karabekian, who, at age seventy-one, wants to be left alone on his Long Island estate with the secret he has locked inside his potato barn. But then a voluptuous young widow badgers Rabo into telling his life story—and Vonnegut in turn tells us the plain, heart-hammering truth about man's careless fancy to create or destroy what he loves.

**Work**

<table>
<thead>
<tr>
<th>Field</th>
<th>Description</th>
<th>Example</th>
</tr>
</thead>
<tbody>
<tr>
<td>id</td>
<td>URI that must be <a href="http://bibfra.me/">http://bibfra.me/</a>...</td>
<td></td>
</tr>
<tr>
<td>title</td>
<td>String that must be one</td>
<td></td>
</tr>
<tr>
<td>titleRemainder</td>
<td>String</td>
<td></td>
</tr>
<tr>
<td>type</td>
<td>String from controlled vocabulary</td>
<td></td>
</tr>
<tr>
<td>classification</td>
<td>String</td>
<td></td>
</tr>
<tr>
<td>author</td>
<td>URI</td>
<td></td>
</tr>
<tr>
<td>subject</td>
<td>URI</td>
<td></td>
</tr>
<tr>
<td>language</td>
<td>String</td>
<td></td>
</tr>
<tr>
<td>summary</td>
<td>String</td>
<td></td>
</tr>
<tr>
<td>hasInstance</td>
<td>URI that must be one and only one</td>
<td></td>
</tr>
</tbody>
</table>

**Instance**

<table>
<thead>
<tr>
<th>Field</th>
<th>Description</th>
<th>Example</th>
</tr>
</thead>
<tbody>
<tr>
<td>id</td>
<td>URI that must be <a href="http://bibfra.me/">http://bibfra.me/</a>...</td>
<td></td>
</tr>
<tr>
<td>type</td>
<td>String from controlled vocabulary</td>
<td></td>
</tr>
<tr>
<td>edition</td>
<td>String</td>
<td></td>
</tr>
<tr>
<td>volumes</td>
<td>String integer</td>
<td></td>
</tr>
<tr>
<td>pages</td>
<td>String integer</td>
<td></td>
</tr>
<tr>
<td>date</td>
<td>Date using Bibframe date-time format</td>
<td></td>
</tr>
<tr>
<td>publisher</td>
<td>URI</td>
<td></td>
</tr>
<tr>
<td>ISBN</td>
<td>String that must be valid ISBN</td>
<td></td>
</tr>
<tr>
<td>image</td>
<td>URI</td>
<td></td>
</tr>
</tbody>
</table>
Singapore Framework for Dublin Core Application Profiles (2007)

Application Profile

Functional Requirements

Domain Model

Description Set Profile

Data Format

Usage Guidelines

annotates

Domain Standards

Community Domain Model

Metadata Vocabularies

DCMI Abstract Model

DCAM-compatible Syntaxes

Foundation Standards

RDFS

RDF

Wednesday, September 4, 2013
An updated Description Set Profile templating language?
Application-specific Context

Application Profile

- Functional Requirements
- Domain Model
- Description Set Profile
- Data Format

Usage Guidelines

annotates

Domain Standards

- Community Domain Model
- Metadata Vocabularies
- DSP Language
- Compatible Syntaxes

Foundation Standards

- RDFS
- RDF

RDFS compaßle Syntaxes

RDF compaßle Syntaxes

Wednesday, September 4, 2013
Model “reality” with a light touch
Principle of Minimal Semantic Commitment

Application Profile

- Usage Guidelines
- Data Format

Domain Standards

- Functional Requirements
- Domain Model
- Description Set Profile
- Metadata Vocabularies
- DSP Language
- Compatible Syntaxes

Foundation Standards

- RDFS
- RDF

Wednesday, September 4, 2013
Constrain the Data
For consistency and quality control
Good Linked Data

• Design for peaceful coexistence
  – Expect unexpected uses

• Emphasize pragmatic control of data quality
  – Constraints on data rather than sweeping generalizations about the world.

• Keep it simple
  – Simple model have larger audiences
Requirements for application profiles

• Ordinary people can read and author
• Nice if constraints are reusable by machines
• Nice if constraints usable for data validation