Analysing Linked Data principle for common directories

SeaDataCloud Technical Task Group

29th November 2016
Linked Data principles – A reminder...

- Use URIs as names for things
Linked Data principles – A reminder…

• Use URIs as names for things
  • Therefore, we need a robust naming scheme…
Linked Data principles – A reminder...

• Use URIs as names for things
• Use HTTP URIs so people can look those things up
Linked Data principles – A reminder...

- Use URIs as names for things
- Use HTTP URIs so people can look those things up
- When someone looks up a URI, provide some useful information using standards
Linked Data principles – A reminder...

- Use URIs as names for things
- Use HTTP URIs so people can look those things up
- When someone looks up a URI, provide some useful information using standards
  - RDF, SPARQL
Linked Data principles – A reminder...

• Use URIs as names for things
• Use HTTP URIs so people can look those things up
• When someone looks up a URI, provide some useful information using standards
• Provide links to other data / information
SeaDataCloud – Linked Data tasks

- 8.2.1 Updating metadata formats
8.2.1 Updating metadata formats

- Analyse appropriate “ontologies” for publication of SeaDataNet metadata
- Allows re-use of well-known patterns
- Some initial suggestions
  - W3C Organization ontology for EDMO
  - W3C DCAT / Google Science Datasets for EDMED
  - Oxford University ARPFO for EDMERP
  - O&M for CDI (Linkage with WP8.3)
  - INSPIRE EMF for EDIOS
  - Contact R2R for CSRs
- Map fields from SeaDataNet to well-known patterns
- Due M8
SeaDataCloud – Linked Data tasks

8.2.1 Updating metadata formats

<http://www.seadatanet.org/edmo/396> a org:Organization;
org:identifier "sdn::edmo::396";
skos:prefLabel "Marine Institute"@en;
skos:prefLabel "Foras na Mara"@ga;
sdn:organisationProfile "The Marine Institute is a ..."@en;
sdn:collatingCentre <http://www.seadatanet.org/edmo/396>;
org:hasSite <_:site>.

_:site a org:Site;
org:siteOf <http://www.seadatanet.org/edmo/396>;
org:siteAddress <:_address>. 
8.2.1 Updating metadata formats

`:address` a vcard:Home.
  vcard:street-address “Rinville, Oranmore”@en;
  vcard:locality “Co. Galway”@en;
  vcard:country-name “Ireland”@en.

- Address pattern need tidying
- More links to URLs required…
• 8.2.1 Updating metadata formats
  <http://www.seadatanet.org/edmed/1996> a dcat:Dataset;
  dct:title "Celtic Explorer Underway North Atlantic"@en;
  sdn:dataHoldingCentre <http://www.seadatanet.org/edmo/396>;
  dct:spatial <http://vocab.nerc.ac.uk/collection/C19/current/1_6>/;
  dct:spatial <http://vocab.nerc.ac.uk/collection/C19/current/1>/;
  dct:spatial <http://vocab.nerc.ac.uk/collection/C19/current/UKMDN015>/;
  dct:spatial <http://vocab.nerc.ac.uk/collection/C19/current/1_4>/;
  dct:temporal "2003-2015";
8.2.1 Updating metadata formats


dcat:theme <http://vocab.nerc.ac.uk/collection/P02/current/AYMD/>;
dcat:keyword "Date and time";
dcat:theme <http://vocab.nerc.ac.uk/collection/P02/current/TEMP/>;
dcat:keyword "Temperature of the water column";
dcat:theme <http://vocab.nerc.ac.uk/collection/L05/current/302/>;
dcat:keyword "Water body temperature sensor";
• 8.2.1 Updating metadata formats
  dct:publisher <http://www.seadatanet.org/edmo/396>;
dct:description "Sea surface temperature is collected..." @en;
rdfs:seeAlso <https://www.bodc.ac.uk/data/information_and_inventories/edmed/report/1966/>;
  sdn:dataCollatingCentre <http://www.seadatanet.org/edmo/396>;
  dct:identifier "sdn::edmed:396:f0c308a9-4c5c-11df-9879-0800200c9a66";
  dct:identifier "sdn::edmed:1996";
  dct:modified “2015-09-14”.

• Identifier schemes need to be tightened (see below)
• DCAT Catalog, CatlogRecord and Distribution entries need to be considered
SeaDataCloud – Linked Data tasks

• 8.2.1 Updating metadata formats
• 8.2.2 Developing pilot SPARQL endpoints for each catalogue
SeaDataCloud – Linked Data tasks

• 8.2.1 Updating metadata formats
• 8.2.2 Developing pilot SPARQL endpoints for each catalogue
  • Identify candidate datasets for the pilot
    • All catalogues to be targeted: EDMED, EDMERP, EDMO, CSR, EDIOS, CDI
  • Publish as RDF
  • Make the RDF accessible through a SPARQL query interface
  • Design a demonstrator app
    • Combining results from multiple catalogues & sources
  • Demonstrate the benefit of the approach
• Due M14
SeaDataCloud – Linked Data tasks

- 8.2.1 Updating metadata formats
- 8.2.2 Developing pilot SPARQL endpoints for each catalogue
- 8.2.3 Updating the REST services
  - Content negotiation – HTML, RDF, JSON(?)
  - REST services for RDF can call the SPARQL endpoints
  - Need to consider robust naming schemes
    - Handles?
    - e.g. hdl.seadatanet.org/edmed/1234
  - Due M18
# SeaDataCloud – Linked Data timeline

<table>
<thead>
<tr>
<th>M0</th>
<th>M3</th>
<th>M6</th>
<th>M9</th>
<th>M12</th>
<th>M15</th>
<th>M18</th>
</tr>
</thead>
<tbody>
<tr>
<td>Initial Linked Data recommendations for EDMERP, CDI, EDMO</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Initial Linked Data recommendations for EDIOS, CSR and EDMED</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>D8.3</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>M8</td>
<td></td>
</tr>
<tr>
<td>Identification of example datasets</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Delivery of RDF &amp; SPARQL</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>D8.4</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
## SeaDataCloud – Linked Data Data timeline

<table>
<thead>
<tr>
<th>Task</th>
<th>M0</th>
<th>M3</th>
<th>M6</th>
<th>M9</th>
<th>M12</th>
<th>M15</th>
<th>M18</th>
</tr>
</thead>
<tbody>
<tr>
<td>Scope and implement robust identification scheme</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Scope and specify upgrades to REST services</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Implement upgrades</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>D8.5 M18</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
**Next steps**
- First draft RDF patterns for each catalogue to be distributed by end of January
- Feedback on patterns and iterations during February
- Work on translations from plain-old-XML to RDF from beginning of March onwards