

DELIVERABLE D4: METADATA OF THE RESOURCES DEALT WITH IN THE PROJECT

Author:

Victor de Boer (VU Amsterdam)

Deliverable type: DATA

Version	Date	Author	Description
0.1	14-09-2013	Victor de Boer	First version
1.0	23-4-2014	Victor de Boer	Final version
1.1	23-5-2014	Victor de Boer	Minor edits

This deliverable describes the CMDI metadata of the resources.

CMDI PROFILE AND MAPPING TO ISOCAT

The CMDI profile was created for this project with the identifier *clarin.eu:cr1:p_1396012485132*. We reuse DC terms components and introduce two new datacategories (Linked_Data_Set and Named_Graph). These two new categories are linked to two new ISOCAT categories (<http://www.isocat.org/datcat/DC-6561> and <http://www.isocat.org/datcat/DC-6568>). (see screenshot). The xsd for the profile is found at http://catalog.clarin.eu/ds/ComponentRegistry/rest/registry/profiles/clarin.eu:cr1:p_1396012485132/xsd

The screenshot shows the ISOCAT web interface. On the left is a navigation tree with 'My Workspace' containing 'Private' and 'Public' categories. The main area displays a table of 'Private Data Categories' with columns for ID, Name, Version, Administration status, Registration status, Check, Type, Owned by, and Scope. Two entries are visible: ID 6561 'Linked Data Set' and ID 6568 'Named Graph'. Below the table, the details for the 'Linked Data Set - 1:0' are shown, including its PID, Type, Owner, and Scope. The interface also displays sections for 'Administration Information Section', 'Description Section', and 'Conceptual Domain'.

#	Name	Version	Administration st	Registration stat	Check	Type	Owned by	Scope
6561	Linked Data Set	1:0	private	private		open	de Boer, Victor	private
6568	Named Graph	1:0	private	private		open	de Boer, Victor	private

1. Administration Information Section

[+] 1.1. Administration Record

2. Description Section

Profile Terminology

2.1 Data Element Name Section

Data Element Name Linked data set

Source

[+] 2.2 English Language Section

Language English (en)

2.2.1 Name Section

Name Linked Data Set

Name Status preferred name

2.2.2 Definition Section

Definition A Linked Dataset (being one or more RDF files) It can be serialized as RDF/XML, Turtle or any other W3C admitted format

Source

3. Conceptual Domain

Data Type anyURI

CMDI FILES

For the four datasets and the DSS schema, a total of five CMDI metadata files were created, in accordance with CLARIN standards. Since our RDF data contains only metadata, the CMDI files are in a sense 'meta-metadata'. They list the individual named graphs that make up a dataset. We furthermore use DC-terms data categories. The CMDI files are found at <https://github.com/biktorrr/dss/tree/master/cmd%20isocat>

An example of a CMDI file (for Noordelijke Monsterrollen) is found below.

```
<CMD xmlns="http://www.clarin.eu/cmd/" xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
CMDVersion="1.1" xsi:schemaLocation="http://www.clarin.eu/cmd/
http://catalog.clarin.eu/ds/ComponentRegistry/rest/registry/profiles/clarin.eu:cr1:p_1396012485132/xsd">
<Header>
  <MdCreator>V. de Boer</MdCreator>
  <MdCreationDate>2014-04-20</MdCreationDate>
  <MdSelfLink/>
  <MdProfile>clarin.eu:cr1:p_1396012485132</MdProfile>
  <MdCollectionDisplayName/>
</Header>

<Resources>
  <ResourceProxyList>
    <ResourceProxy id="mdb_data">
      <ResourceType mimetype="application/xml">Resource</ResourceType>
      <ResourceRef>${easy-dataset}/dss/mdb_data.ttl.gz.xml</ResourceRef>
    </ResourceProxy>
    <ResourceProxy id="mdb_thes_generated">
      <ResourceType mimetype="application/xml">Resource</ResourceType>
      <ResourceRef>${easy-dataset}/dss/mdb_thes_generated.ttl.xml</ResourceRef>
    </ResourceProxy>
    <ResourceProxy id="mdb_schema">
      <ResourceType mimetype="application/xml">Resource</ResourceType>
      <ResourceRef>${easy-dataset}/dss/mdb_schema.ttl.xml</ResourceRef>
    </ResourceProxy>
    <ResourceProxy id="mdb_thes_places">
      <ResourceType mimetype="application/xml">Resource</ResourceType>
      <ResourceRef>${easy-dataset}/dss/mdb_thes_places.ttl.xml</ResourceRef>
    </ResourceProxy>
    <ResourceProxy id="ranks_and_shiptypes_1">
      <ResourceType mimetype="application/xml">Resource</ResourceType>
      <ResourceRef>${easy-dataset}/dss/ranks_and_shiptypes_1.ttl.xml</ResourceRef>
    </ResourceProxy>
    <ResourceProxy id="mdb_2_kb">
      <ResourceType mimetype="application/xml">Resource</ResourceType>
      <ResourceRef>${easy-dataset}/dss/mdb_2_kb.ttl.xml</ResourceRef>
    </ResourceProxy>
    <ResourceProxy id="mdb_ship_sameas">
      <ResourceType mimetype="application/xml">Resource</ResourceType>
      <ResourceRef>${easy-dataset}/dss/mdb_ship_sameas.ttl.xml</ResourceRef>
    </ResourceProxy>
  </ResourceProxyList>
  <JournalFileProxyList/>
  <ResourceRelationList/>
</Resources>

<Components>
  <Linked_Data_Set>
    <dc-terms>
```

```

    <contributor>Leinenga, J.</contributor>
    <contributor>Boer, V. de</contributor>
    <contributor>Hoekstra, R.</contributor>
    <contributor>van Rossum, M.</contributor>
    <publisher>Huygens Instituut voor Nederlandse Geschiedenis</publisher>
    <relation>Dutch Ships and Sailors Linked Data [Resource is a collection of 5
parts][Relationship described is IsPartOf]</relation>
    <rights>This work is licensed under a Creative Commons Attribution 3.0 Unported
License</rights>

    <source>http://www.github.com/biktorrr/dss</source>
    <title>Noordelijke Monsterrollen Database als Linked Data</title>
    <format>RDF/XML</format>
    <coverage>1803-1936</coverage>
    <description>Dataset met alle in havens in Friesland, Groningen, Drenthe en
Overijssel geregistreerde bemanningen van zeeschepen in de periode 1803-1936, uiteindelijk lopend t/m 1956.
De bronnen voor dit overzicht worden gevormd door bemanningslijsten in de vorm van monsterrollen en registers
van aanmonstering, voor zover aanwezig in openbare archieven, maritieme musea of particuliere collecties in de
vier noordelijke provincies. De structuur van een monsterrol is bepaald in enkele Koninklijke Besluiten, maar
hiervan is – vooral in de eerste helft van de negentiende eeuw - vaak afgeweken. Monsterrollen en registers
bevatten altijd informatie over het schip, de bemanning en de datum van aanmonstering. De monsterrollen en
registers bevatten de volgende kerngegevens: naam van het schip, naam van de schipper, tonnage van het
schip, ligplaats en bestemming, bemanning (uitgesplitst in verschillende categorieën).</description>
  </dc-terms>
  <Named_Graph ref="mdb_data">
    <description>MDB Main data named graph</description>
  </Named_Graph>
  <Named_Graph ref="mdb_thes_generated">
    <description>MDB SKOS thesaurus generated from literal values</description>
  </Named_Graph>
  <Named_Graph ref="mdb_schema">
    <description>MDB RDFS Schema</description>
  </Named_Graph>
  <Named_Graph ref="mdb_thes_places">
    <description>MDB Places consolidated from literal values</description>
  </Named_Graph>
  <Named_Graph ref="ranks_and_shiptypes_1">
    <description>MDB Ranks and Shiptypes, based on input by J. Leinenga</description>
  </Named_Graph>
  <Named_Graph ref="mdb_2_kb">
    <description>MDB Links to Koninklijke bibliotheek (KB) (A. Bravo
Balado)</description>
  </Named_Graph>
  <Named_Graph ref="mdb_ship_sameas">
    <description>MDB OWL:sameAs links identifying recurring ships (R.
Ponstein)</description>
  </Named_Graph>
</Linked_Data_Set>
</Components>
</CMD>

```