

## DELIVERABLE D8: METADATA MADE AVAILABLE ON A RECOGNIZED CLARIN SERVER

Author:

Victor de Boer (VU Amsterdam)

Deliverable type: DATA

Version	Date	Author	Description
0.1	26-03-2014	Victor de Boer	First version
1.0	23-04-2014	Victor de Boer	Final version

This deliverable describes the availability of data and metadata on CLARIN Servers. It copies information from Deliverable D8 and D4, since in our case there is no clear distinction between data and metadata.

### INFRASTRUCTURE (LIVE LINKED DATA)

The infrastructure, as described in Deliverable D5 was transferred to Huygens ING – a CLARIN center. Huygens ING will maintain the live linked data in the form of a running ClioPatria Triple store available at <http://dutchshipsandsailors.nl/data>. Here the (meta)data as well are available in RDF.

### OFFLINE DATA AND METADATA

The datasets were also exported to RDF/XML format and are submitted to the Data Archiving and Networking Services (DANS) of the KNAW through Easy. The screenshot below shows the EASY interface and the five datasets submitted.

The screenshot shows the DANS website interface. At the top, there is a navigation bar with the DANS logo, a search bar, and user information for 'V. de Boer'. Below the navigation bar, there is a section titled 'My deposited datasets'. This section includes a list of five datasets, each with its title, author, date, and status. To the right of the dataset list is a 'Refine' sidebar with a search bar and filters for Audience, Access, and Status. The footer of the page contains links for 'Using EASY', 'Contact', 'Disclaimer', 'Legal information', 'Property rights statement', and 'Data Seal of Approval for EASY', along with version and build date information.

Dataset Title	Author	Date	Status	Audience	Access	Submitted
Generale Zeemonsterrollen VOC als RDF (Dutch Ships and Sailors)	de Boer, Dr. V. (VU University Amsterdam)	(2014-04-19)	submitted 2014-05-23	Modern history	Open	2014-05-23
VOC Opvarenden als RDF (Dutch Ships and Sailors)	de Boer, Dr. V. (VU University Amsterdam)	(2014-04-19)	submitted 2014-05-23	Modern history	Open	2014-05-23
Noordelijke Monsterrollen Database as RDF (Dutch Ships and Sailors)	de Boer, Dr. V. (VU University Amsterdam), Leinenga, Dr. J.	(2014-04-24)	submitted 2014-05-23	Modern history	Open	2014-05-23
Dutch Ships and Sailors interoperability layer as RDF (Dutch Ships and Sailors)	de Boer, Dr. V. (VU University Amsterdam)	(2014-04-23)	submitted 2014-05-23	Modern history	Open	2014-05-23
Dutch Asiatic Shipping as RDF (Dutch Ships and Sailors)	de Boer, Dr. V. (VU University Amsterdam)	(2014-04-19)	submitted 2014-05-23	Modern history	Open	2014-05-23

The datasets consist of the individual named graphs in RDF XML format as described in Deliverable D3 and D10 and the CMDI metadata as described in Deliverable D4. The Easy dataset identifiers are listed below.

DAS [urn:nbn:nl:ui:13-az7h-zh](https://easy.dans.knaw.nl/ui/datasets/id/easy-dataset:57620) <https://easy.dans.knaw.nl/ui/datasets/id/easy-dataset:57620>

DSS [urn:nbn:nl:ui:13-korn-zi](https://easy.dans.knaw.nl/ui/datasets/id/easy-dataset:57622) <https://easy.dans.knaw.nl/ui/datasets/id/easy-dataset:57622>

MDB [urn:nbn:nl:ui:13-czhm-ug](https://easy.dans.knaw.nl/ui/datasets/id/easy-dataset:57617) <https://easy.dans.knaw.nl/ui/datasets/id/easy-dataset:57617>

VOCOPV [urn:nbn:nl:ui:13-gupf-pd](https://easy.dans.knaw.nl/ui/datasets/id/easy-dataset:57623) <https://easy.dans.knaw.nl/ui/datasets/id/easy-dataset:57623>

GZMVOC [urn:nbn:nl:ui:13-xms7-0y](https://easy.dans.knaw.nl/ui/datasets/id/easy-dataset:57621) <https://easy.dans.knaw.nl/ui/datasets/id/easy-dataset:57621>

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 data categories (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](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 categories like 'Private', 'Shared', and 'Public'. The main area displays a table of datasets:

#	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

Below the table, the details for the 'Linked Data Set - 1:0' are shown:

- PID: <http://www.isocat.org/datcat/DC-6561>
- Type: complex/open
- Owner: [de Boer, Victor](#)
- Scope: private

The details are organized into sections:

- Administration Information Section
  - 1.1. Administration Record
- Description Section
  - 2.1 Data Element Name Section
    - Data Element Name: Linked data set
    - Source: Terminology
  - 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: Terminology
- 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>
  </ResourceProxyList>
</Resources>
```

```

<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 kernegegevens: 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>
```