Data Catalog Vocabulary (DCAT)

A very brief introduction

Gilles Landais (CDS)

DCAT

Motivation , history and description

- **DCAT** (**D**ata **Cat**alogue vocabulary) is a W3C standard created in 2014, regularly updated (last version March 2023: version3)

  "DCAT is an RDF vocabulary designed to facilitate interoperability between data catalogs published on the Web. "

- Describe **Dataset** and **Services**

  - gather datasets and services into **catalogue**

  - Link dataset with service

  - Link datasets

  - Aggregate catalogues

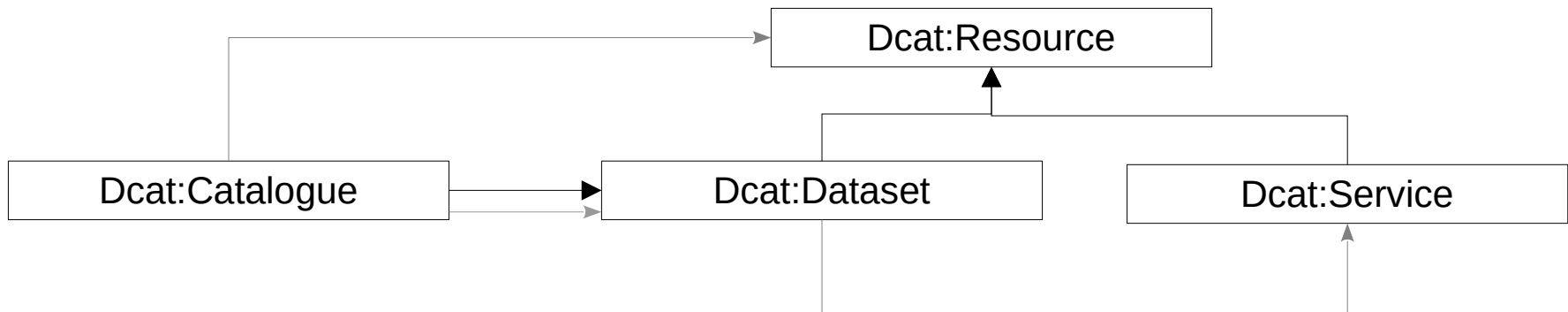- Open Science semantic (used for instance in Google dataset search)

DCAT

Motivation , history and description

▪ A semantic mapped in a Data-model that exploits other namespaces like Prov, DublinCore or FOAF (friend-of-friend)

▪ Serialisation : RDF/XML, Turtle, JSON-LD

Prefix	Namespace IRI	Source
adms	http://www.w3.org/ns/adms#	[VOCAB-ADMS]
dc	http://purl.org/dc/elements/1.1/	[DCTERMS]
dcat	http://www.w3.org/ns/dcat#	[VOCAB-DCAT]
dcterms	http://purl.org/dc/terms/	[DCTERMS]
dctype	http://purl.org/dc/dcmitype/	[DCTERMS]
foaf	http://xmlns.com/foaf/0.1/	[FOAF]
locn	http://www.w3.org/ns/locn#	[LOCN]
odrl	http://www.w3.org/ns/odrl/2/	[ODRL-VOCAB]
owl	http://www.w3.org/2002/07/owl#	[OWL2-SYNTAX]
prov	http://www.w3.org/ns/prov#	[PROV-O]
rdf	http://www.w3.org/1999/02/22-rdf-syntax-ns#	[RDF-SYNTAX-GRAMMAR]
rdfs	http://www.w3.org/2000/01/rdf-schema#	[RDF-SCHEMA]
skos	http://www.w3.org/2004/02/skos/core#	[SKOS-REFERENCE]
spdx	http://spdx.org/rdf/terms#	[SPDX]
time	http://www.w3.org/2006/time#	[OWL-TIME]
vcard	http://www.w3.org/2006/vcard/ns#	[VCARD-RDF]
xsd	http://www.w3.org/2001/XMLSchema#	[XMLSCHEMA11-2]

DCAT

The model

- **dcat:Resource** : basis class represents a dataset or a data service
  attributes: Title, landingpage, rights, date, publisher, relations, versions, ...

- **dcat:Dataset** is a Resource of data, published or curated by a single agent
  eg: a VizieR table

- **dcat: DataService**: A collection of operations that provides access to one or more datasets
  eg: conesearch, ftp, ..

- **Dcat:Catalogue**: A curated collection of metadata about resources.
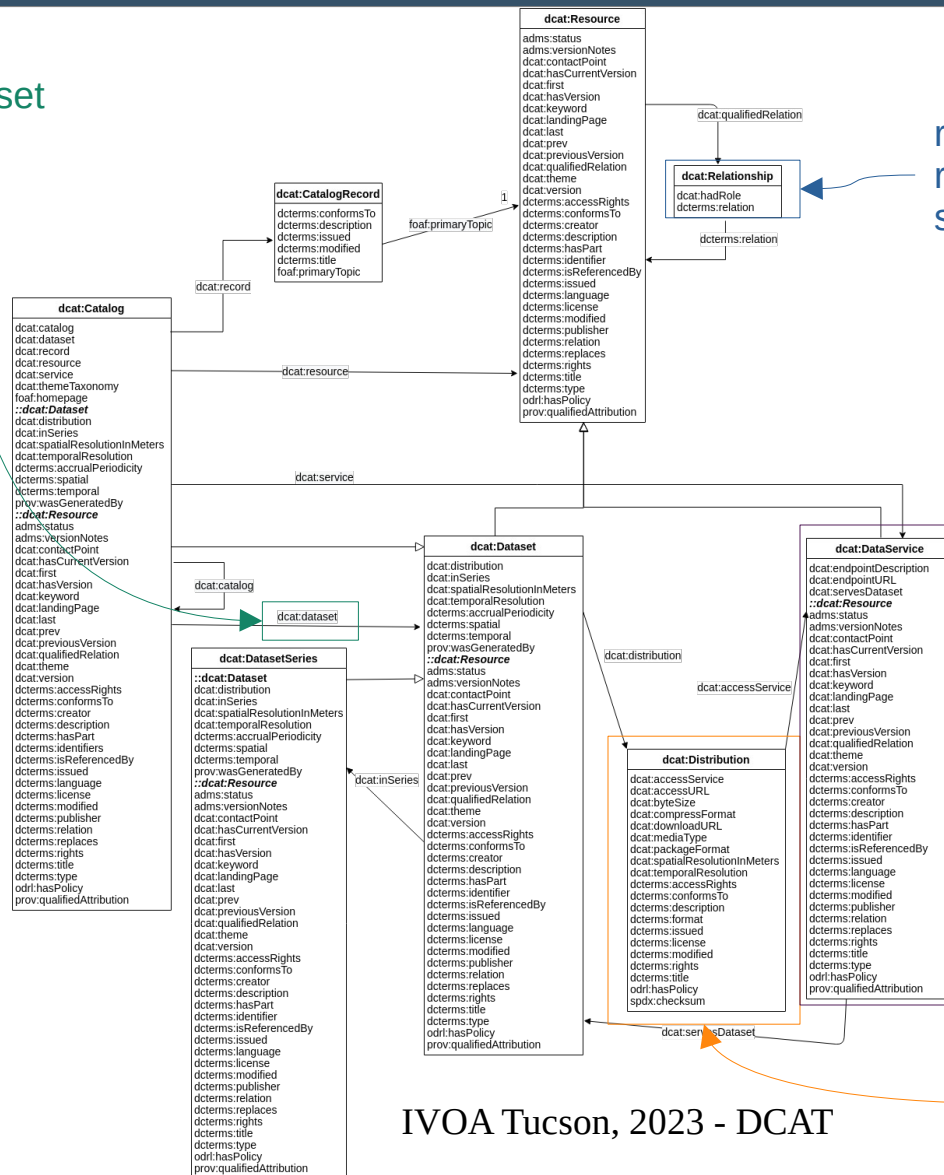 *eg: VizieR catalogue*

DCAT

IVOA Tucson, 2023 - DCAT

DCAT

Implementation example

- National Extension of DCAT

  - Open data initiative of the Government of Spain
    https://datos.gob.es/sites/default/files/doc/file/report_dcat-ap_and_its_extensions
    .pdf

  - European comission
    https://joinup.ec.europa.eu/collection/semantic-interoperability-community-semic/
    solution/dcat-application-profile-data-portals-europe/release/11

  - (French) Research data Gouv
    https://doc.data.gouv.fr/moissonnage/dcat/

  - US resources data gouv
    https://resources.data.gov/resources/dcat-us/

- Google dataset search:
"We can understand structured data in web pages about datasets, using either
schema.org Dataset markup, or equivalent structures represented in W3C's Data
Catalog Vocabulary (DCAT) format"

- Ifremer : french institute about ocean science
https://sextant.ifremer.fr/geonetwork/
https://sextant.ifremer.fr/geonetwork/api/collections/520d5344-c030-449c-886f-a5ac09
7fddf9/items/96d1523c0a86ac418a861b20e1b93b36f5e723bc?f=dcat

DCAT

Conclusion

▪ Complex model with possible (?) mapping with vo-registry

▪ Show the need of interoperability in Open Science

▪ Increasing usage of Semantic, RDF, JSON-LD

▪ A lot of concepts, semantics that I am not aware of!..

Example of XML/RDF DCAT output provided by Ifremer
https://sextant.ifremer.fr/geonetwork/api/collections/520d5344-c030-449c-886f-a5ac097fddf9/items/96d1523c
0a86ac418a861b20e1b93b36f5e723bc?f=dcat

A short selection of DCAT
terms that completes
DublinCore semantic applied
to cataloge and service

serialization mixes concepts
from different semantics

```xml
<dcat:CatalogRecord xmlns:adms="http://www.w3.org/ns/adms#" xmlns:dcatap="http://data.europa.eu/r5r/"
xmlns:owl="http://www.w3.org/2002/07/owl#" xmlns:skos="http://www.w3.org/2004/02/skos/core#"
xmlns:rdfs="http://www.w3.org/2000/01/rdf-schema#" xmlns:vcard="http://www.w3.org/2006/vcard/ns#"
xmlns:dct="http://purl.org/dc/terms/" xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
xmlns:ns14="http://spdx.org/rdf/terms#" xmlns:locn="http://www.w3.org/ns/locn#" xmlns:dcat="http://www.w3.org/ns/dcat#"
xmlns:prov="http://www.w3.org/ns/prov#" xmlns:foaf="http://xmlns.com/foaf/0.1/" xmlns:dc="http://purl.org/dc/elements/1.1/"
rdf:about="https://sextant.ifremer.fr/geonetwork/srv/metadata/96d1523c0a86ac418a861b20e1b93b36f5e723bc">
  <dct:identifier>96d1523c0a86ac418a861b20e1b93b36f5e723bc</dct:identifier>
  <dct:created>2022-04-10T18:45:00Z</dct:created>
  <dct:modified>2022-04-10T18:45:00Z</dct:modified>
  <dct:language>
    <skos:Concept rdf:about="http://publications.europa.eu/resource/authority/language/FRE">
      <skos:prefLabel>fre</skos:prefLabel>
      <rdf:type rdf:resource="http://purl.org/dc/terms/LinguisticSystem"/>
    </skos:Concept>
  </dct:language>
  <foaf:primaryTopic>
    <dcat:DataService
    rdf:about="https://sextant.ifremer.fr/geonetwork/srv/metadata/96d1523c0a86ac418a861b20e1b93b36f5e723bc#service">
      <dct:title>Serveur wms public de l'Ifremer - Référentiels géographiques du Système d'informations halieutiques
      SIH</dct:title>
      <dct:description>Serveur wms public de l'Ifremer, pour les référentiels du SIH</dct:description>
      <dct:modified>2023-10-18T22:00:00Z</dct:modified>
      <dcat:contactPoint>
        <vcard:Kind>
          <vcard:title>Ifremer - SIH - Sextant</vcard:title>
          <vcard:role>pointOfContact</vcard:role>
          <vcard:hasEmail/>
        </vcard:Kind>
      </dcat:contactPoint>
      <dcat:theme>
        <skos:Concept>
          <skos:prefLabel>Sextant</skos:prefLabel>
        </skos:Concept>
      </dcat:theme>
      <dcat:theme>
        <skos:Concept>
          <skos:prefLabel>Inspire</skos:prefLabel>
        </skos:Concept>
      </dcat:theme>
      <dct:type>
        <skos:Concept>
          <skos:prefLabel>DataService</skos:prefLabel>
        </skos:Concept>
      </dct:type>
      <dcat:landingPage>
        <foaf:Document rdf:about="https://sextant.ifremer.fr/geonetwork/srv/metadata/96d1523c0a86ac418a861b20e1b93b36f5e723bc">
          <dct:title>Serveur wms public de l'Ifremer - Référentiels géographiques du Système d'informations halieutiques
          SIH</dct:title>
        </foaf:Document>
      </dcat:landingPage>
      <dcat:endpointURL rdf:resource="https://sextant.ifremer.fr/services/wms/sih_referentiels?"/>
    </dcat:DataService>
  </foaf:primaryTopic>
</dcat:CatalogRecord>
```