From niche to mainstream, loveable (meta)data

Geraldine Nolf
Previously...
Nowadays
LOTS OF METADATA STANDARDS, PORTALS AND SYSTEMS
BRINGING MISS WORLD AND MISTER CUBE TOGETHER AND OPEN UP DATA

Copyright Miss Globe & Mr Cube:
Frans Knibbe, Geodan

INSPIRE 2020: From niche to mainstream, loveable (meta)data - Geraldine Nolf
BRIDGING THE GAP USING GEONETWORK

**Geospatial data**
- ISO/INSPIRE
- SDI-Flanders Best Practices
- Profile: Geo MD ISO “19XXX”
- Export: ISO “19XXX” XML
- Geopunt Portal (Sitecore & MapApps)

**Open data**
- DCAT-AP
- Profile: Open MD “DCAT-AP XML”
- Export: DCAT-AP RDF

**Metadata standard**
- Metadata management system
- Metadata portal

**Upgrade to GeoNetwork 3.8**

**DCAT-AP plug-in to be build**

INSPRIE 2020: From niche to mainstream, loveable (meta)data - Geraldine Nolf
XML SCHEMA PLUG-IN FOR DCAT-AP

> GeoNetwork can only process XML-based metadata
  ▪ So, an XML Schema plug-in must be defined
> Based on DCAT-AP v1.1; DCAT-AP v1.2
> Plug-in for GeoNetwork 3.8
> Developed by GIM, funded by Informatie Vlaanderen

> Beta environment:
  https://beta.metadata.vlaanderen.be/
> Available as open-source software (GPLv2 license):
  https://github.com/metadata101/dcat-ap1.1
DCAT-AP PLUG-IN FEATURES

> **XML Schema:** The plug-in defines its own XML Schema. The XML Schema was designed to fully resemble an XML/RDF syntax of DCAT-AP.

> **Harvester:** A DCAT-AP harvester was written to "normalize" the RDF metadata, such that it fits in the XML Schema for DCAT-AP.

> **Editor:** A custom form was created that uses the controlled vocabularies required by DCAT-AP.

> **Adding licenses:** Sample licenses templates can be imported via the custom form for editing.

> **Multilingual:** The editor, view, and search benefit from the already existing multilingual capabilities of GeoNetwork.

> **Validator:** Validation steps are XSD validation followed by the schematron validation.

> **RDF endpoint:** The plug-in exports DCAT-AP RDF metadata using the GeoNetwork API, which can in turn be harvested by e.g. CKAN.
DCAT-AP HARVESTER

- XML is a nested data structure that can be queried with Xpath
- RDF is a graph-based data structure in different formats and arbitrarily structured that can be queried with SPARQL

- On import, DCAT-RDF metadata must be “normalised”
DCAT-AP HARVESTER

> Input: Turtle, RDF/XML, N3, JSON-LD
DCAT-AP EDITOR

> Simple and advanced form for editing DCAT-AP records
> Multilingual
> Controlled vocabularies required by DCAT-AP as thesauri
> Reuse (with modification) of existing controls (e.g. spatial extent)
DCAT-AP VALIDATOR

> DCAT-AP constraints have been transformed into three schematron rulesets:
  - DCAT-AP v1.2 mandatory rules
  - DCAT-AP v1.2 recommendations
  - DCAT-AP-VL (additional) mandatory rules for publishing open data
DCAT-AP DATASET VIEW

Dataset  Distributies

Basisinformatie

Titel  havengebied

Beschrijving

Uitgever

URI  https://antwerpen.be

Naam  Stad Antwerpen

Trefwoord  Vlaamse Open data

Identificator  996601F8-3907-3F30-80cB-1C99d2749333

Versie informatie

Wijzigingsdatum  05 aug. 2019 08:39

Publicatiedatum  06 dec. 2018 22:27

Gebruiksinformatie

Bijsluiter  http://portaal-stadantwerpen.opendata.arcgis.com/datasets/cebf8819bc584485284d2e9df8132e52c_183
DCAT-AP RDF ENDPOINT

> Exports DCAT-AP RDF using standard GeoNetwork API
> Supports W3C LDP paging and hydra paging
> Output: RDF/XML
ROOM FOR IMPROVEMENT

> Harvester (input side) with paging support
> GeoNetwork RDF endpoint: multiple formats
> Schema.org integration (JSON.LD on HTML pages)

TIMELINE

> BETA-release: Q2 2019
> Production-release: together with GeoNetwork upgrade & Geo migration (± Q3 2020)
TOWARDS ONE CATALOGUE
TOWARDS ONE CATALOGUE

- **Geospatial data & services**
  - ISO/INSPIRE
  - GeoNetwork
    - CKAN
    - GeoNetwork
  - DCAT-AP

- **Open data**
  - Analyzing DCAT-AP v.2.0
  - (to be – GeoNetwork ?)
  - MAGDA-online
  - (to be)

- **Data concerning natural persons & enterprises**
  - Data/Services
    - Data
    - Services API’s
    - Statistics/Services

- **API’s**
  - Analyzing DCAT-AP v.2.0
  - (to be – GeoNetwork ?)

- **Statistic data & services**
  - SDMX / StatDCAT ?
  - (to be – GeoNetwork ?)

- **Archival Information**
  - EAD
  - (to be – GeoNetwork ?)

---

**Scope Information Catalogue**

INSPIRE 2020: From niche to mainstream, loveable (meta)data - Geraldine Nolf
Information Catalogue

Copyright:
Frans Knibbe, Geodan / ICPSR / 123RF / Geraldine Nolf
INSPIRE 2020: From niche to mainstream, loveable (meta)data - Geraldine Nolf
Contact?

Geraldine.Nolf@vlaanderen.be

& team:

Mathias De Schrijver, Dirk De Baere, Bart Cosyn, Loes Deventer, Peter Haelvoet, Koen De Baets, Severien D’Hoker, Joeri Vleurick;
GIM: Stijn Goedertier, Gustaaf Vandeboel & Mathieu Chaussier