



Vlaamse  
overheid

# DCAT-AP plug-in for GeoNetwork

Ispra, Italy  
3-4 July, 2019

Mathias De Schrijver

Geraldine Nolf

Dirk De Baere

**INFORMATIE  
VLAANDEREN**



[www.vlaanderen.be/informatievlaanderen](http://www.vlaanderen.be/informatievlaanderen)

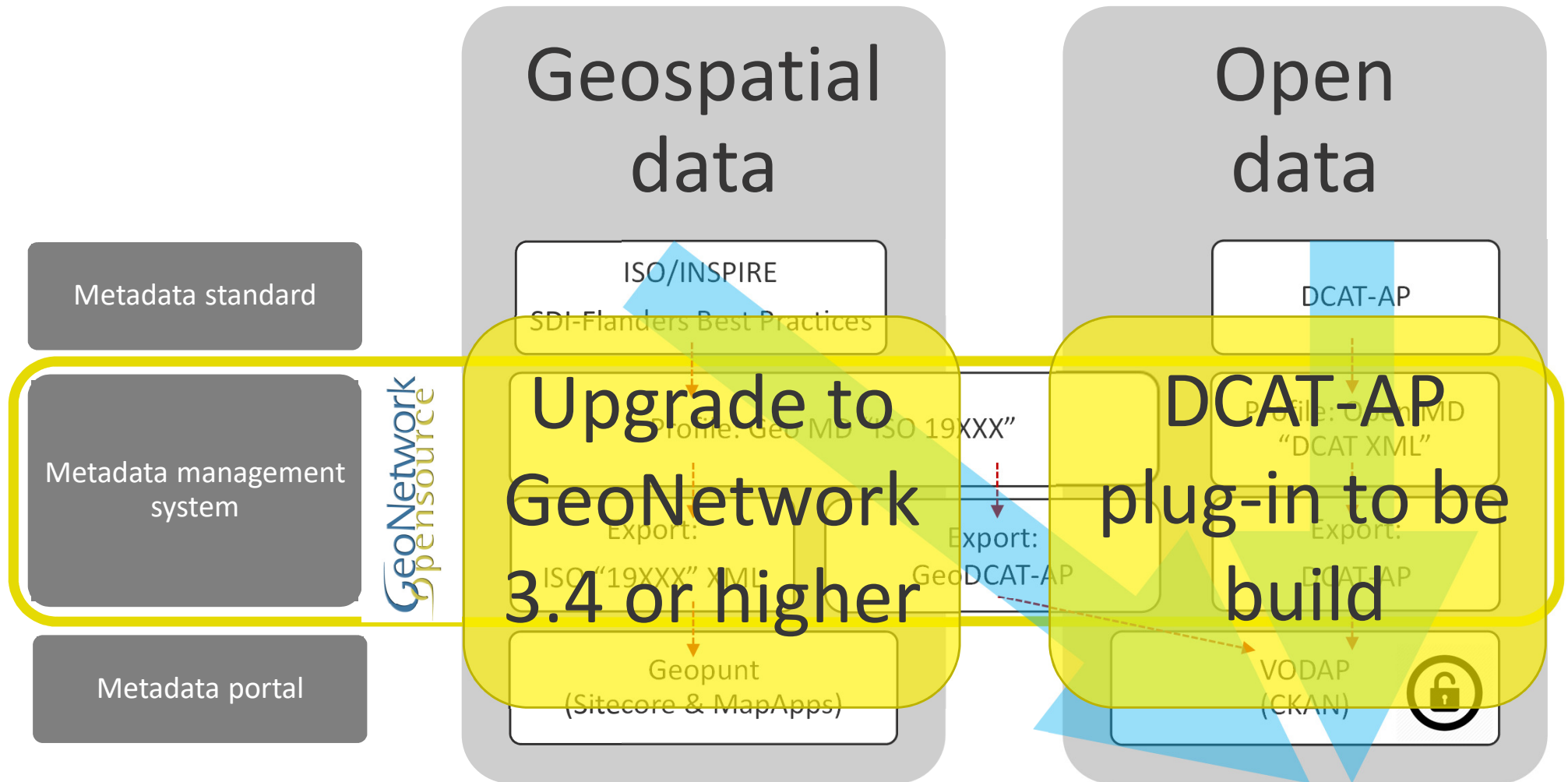
# LOTS OF METADATA STANDARDS, PORTALS AND SYSTEMS



# BRINGING MISS WORLD AND MISTER CUBE TOGETHER AND OPEN UP DATA



# BRIDGING THE GAP USING GEONETWORK



# XML SCHEMA PLUG-IN

- > GeoNetwork can only process XML-based metadata
  - So, an XML Schema plugin must be defined
- > Based on DCAT-AP v1.1
- > Developed by GIM, funded by Informatie Vlaanderen
  
- > Beta environment:  
<http://beta.metadata.vlaanderen.be/>
- > Available as open-source software (GPLv2 license):  
<https://github.com/metadata101/dcat-ap1.1>

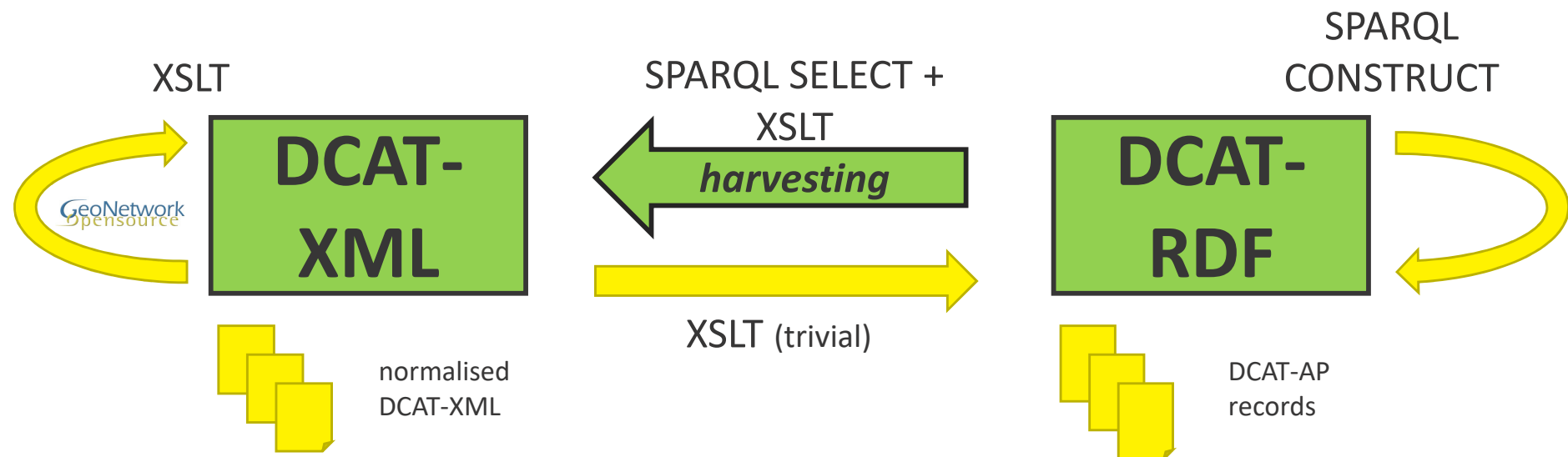


# PLUG-IN FEATURES

- > **XML Schema:** The plugin 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 plugin 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

Update harvester **visitflanders.org**

[x Delete](#) [Save](#) [Harvest](#)

---

**Identification** **Schedule** [Enable](#) [disable](#)

**Node name and logo**

▼

The name describing the remote node.

**Group**

▼

Group which owns the harvested records. Only the catalog administrator or user admin of this group can manage this node.

**User**

▼

User who owns the harvested records

---

**Configuration for protocol harvester-dcatap**

**dcatap-host**

dcatap-hostHelp



# 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)

The screenshot shows the 'Distributies' tab of the DCAT-AP Editor. It displays a list of records with the following fields:

- titel \***: Beer Museums in Flanders and Brussels
- taal**: nl
- aanbevolen waarden**: dropdown menu
- titel \***: Selectie van bier musea in Vlaanderen & Brussel.
- taal**: en
- aanbevolen waarden**: dropdown menu
- + titel**: button
- beschrijving \***: Selection of beer museums in Flanders & Brussels.
- taal**: en
- aanbevolen waarden**: dropdown menu
- beschrijving \***: Selectie van bier musea in Vlaanderen & Brussel.
- taal**: nl
- aanbevolen waarden**: dropdown menu
- + beschrijving**: button

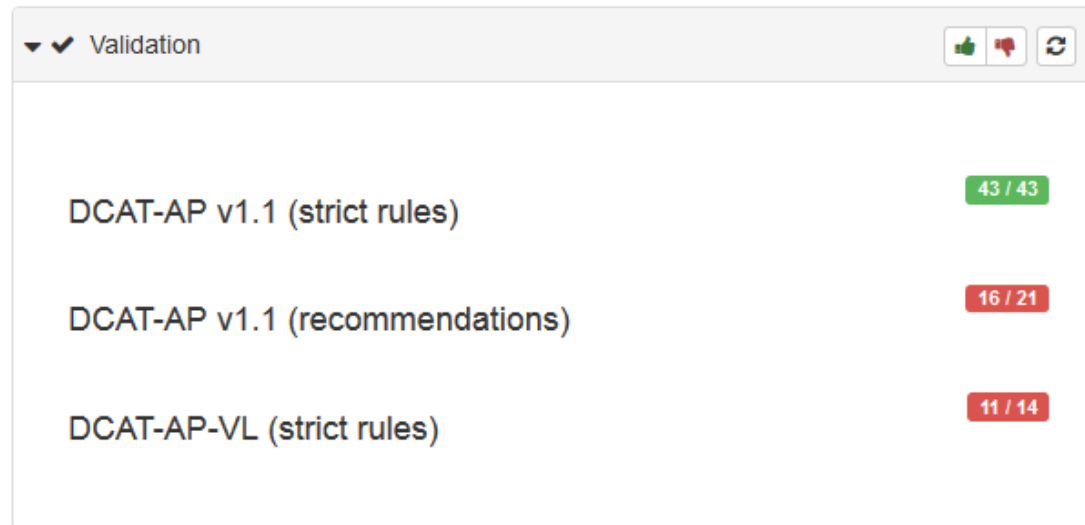
The screenshot shows the map interface of the DCAT-AP Editor. It displays a map of Antwerp, Belgium, with a red bounding box indicating the spatial extent. The map is centered on Antwerpen, Vlaanderen. The bounding box coordinates are:

- Top: 51,505081
- Bottom: 51,262502
- Left: 4,172293
- Right: 5,262502

The map is labeled with 'WGS84(EPGS:4326)' and 'Teken dekkingsgebied'. The map shows the city of Antwerpen and surrounding areas, including Breda, Tilburg, Noord-Brabant, Eindhoven, and Leuven.

# DCAT-AP VALIDATOR

- > DCAT-AP constraints have been transformed into three schematron rulesets:
  - DCAT-AP v1.1 mandatory rules
  - DCAT-AP v1.1 recommendations
  - DCAT-AP-VL (additional) mandatory rules for publishing open data

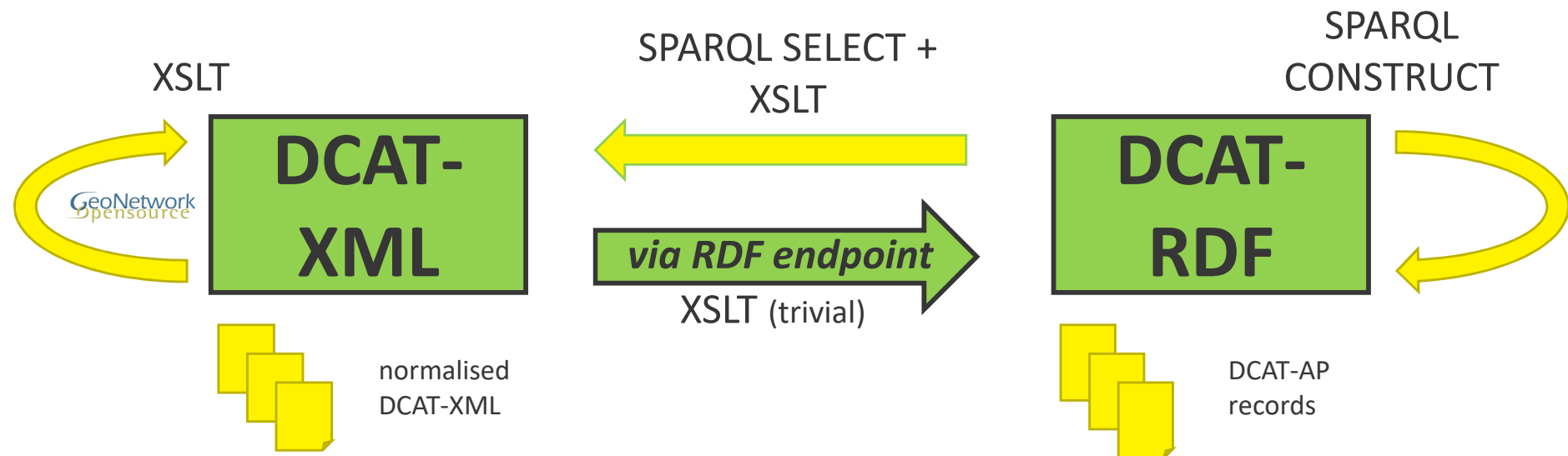


The screenshot shows a window titled "Validation" with a dropdown arrow and a checkmark icon. In the top right corner, there are three icons: a green thumbs up, a red thumbs down, and a refresh symbol. The main content area lists three validation rulesets with their respective counts:

Ruleset	Count
DCAT-AP v1.1 (strict rules)	43 / 43
DCAT-AP v1.1 (recommendations)	16 / 21
DCAT-AP-VL (strict rules)	11 / 14

# 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

- > Usability of validator (link with fields)
- > Harvester with paging support
- > RDF endpoint: multiple formats
- > Schema.org integration (JSON.LD on HTML pages)



Vlaamse  
overheid



## Contact?

Mathias.DeSchrijver@kb.vlaanderen.be

Geraldine.Nolf@kb.vlaanderen.be

Dirk.DeBaere@kb.vlaanderen.be

**INFORMATIE  
VLAANDEREN**

[www.vlaanderen.be/informatievlaanderen](http://www.vlaanderen.be/informatievlaanderen)