Vlaamse overheid

DCAT-AP plug-in for GeoNetwork

Ispra, Italy
3-4 July, 2019

Mathias De Schrijver

Geraldine Nolf

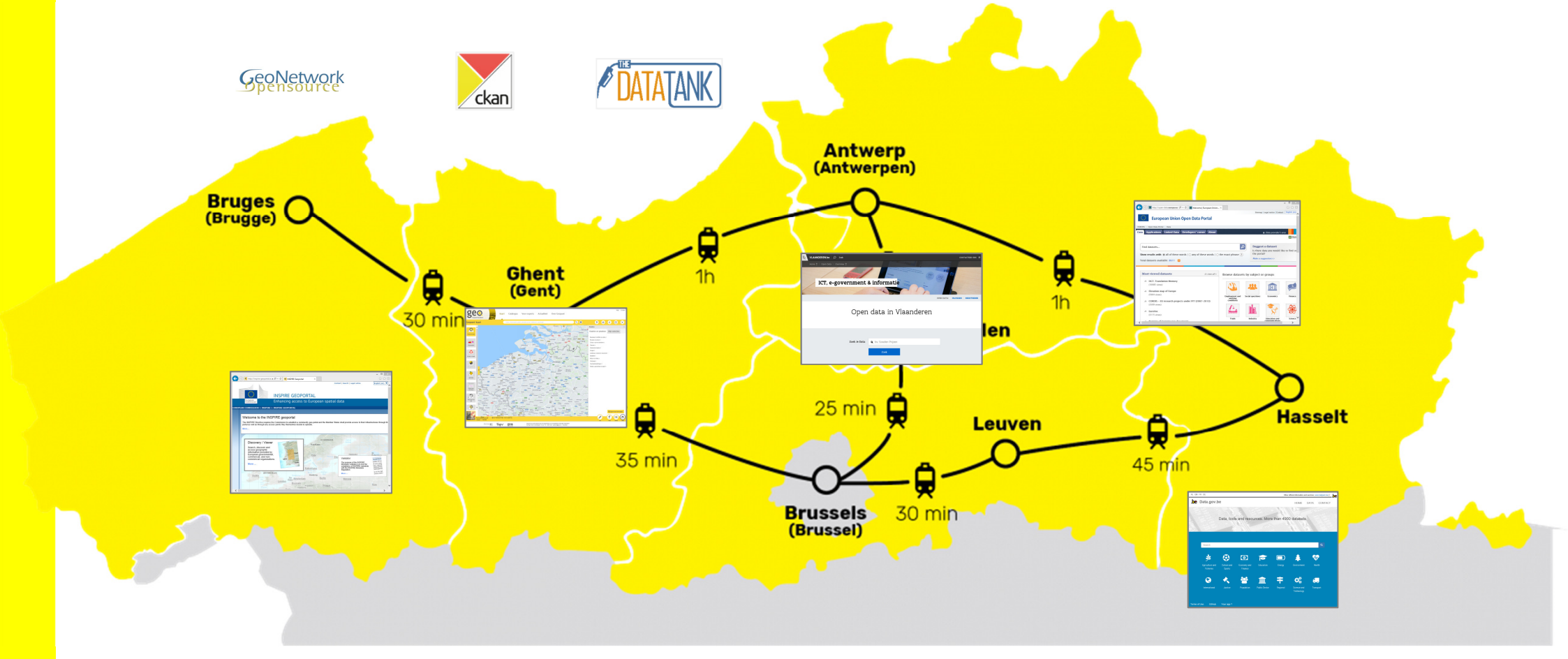
Dirk De Baere

INFORMATIE VLAANDEREN

gim SMART GEO INSIGHTS

www.vlaanderen.be/informatievlaanderen

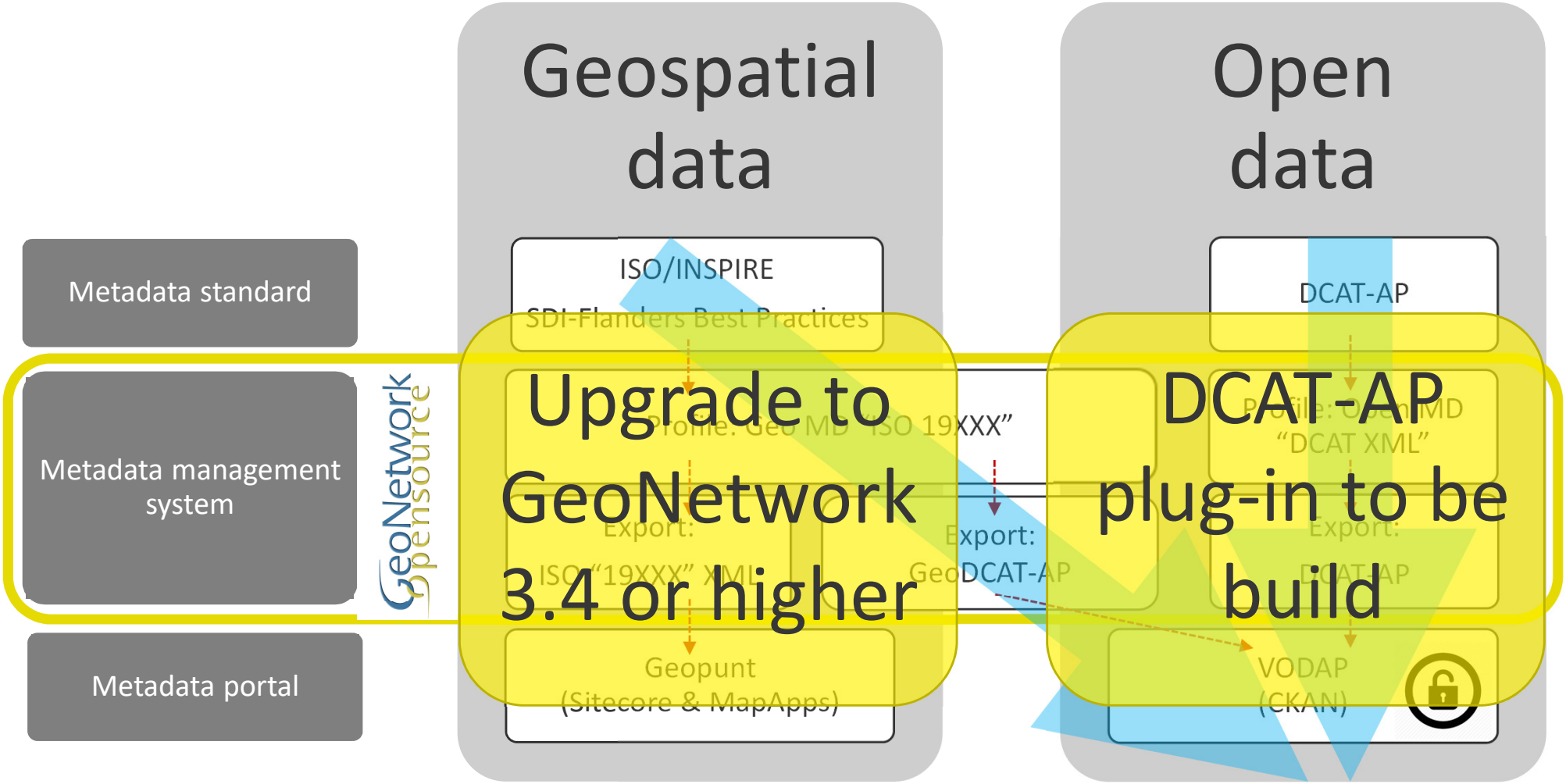
LOTS OF METADATA STANDARDS, PORTALS AND SYSTEMS



BRINGING MISS WORLD AND MISTER CUBE TOGETHER AND OPEN UP DATA

Copyright Miss Globe & Mr Cube:
Frans Knibbe, Geodan

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

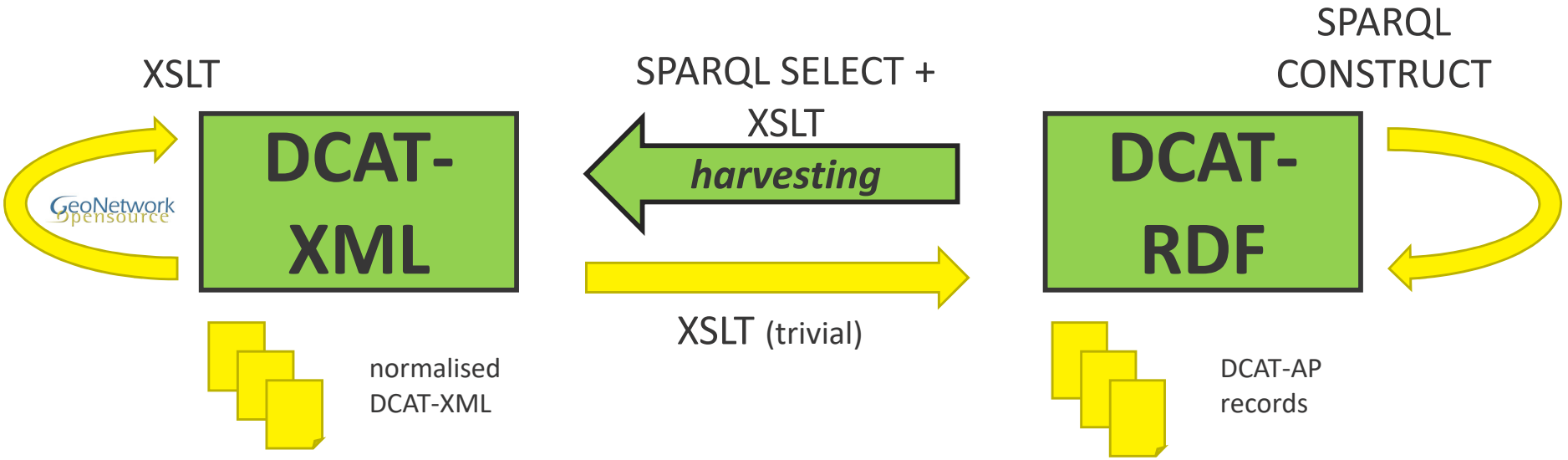
PLUG-IN FEATURES

- **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

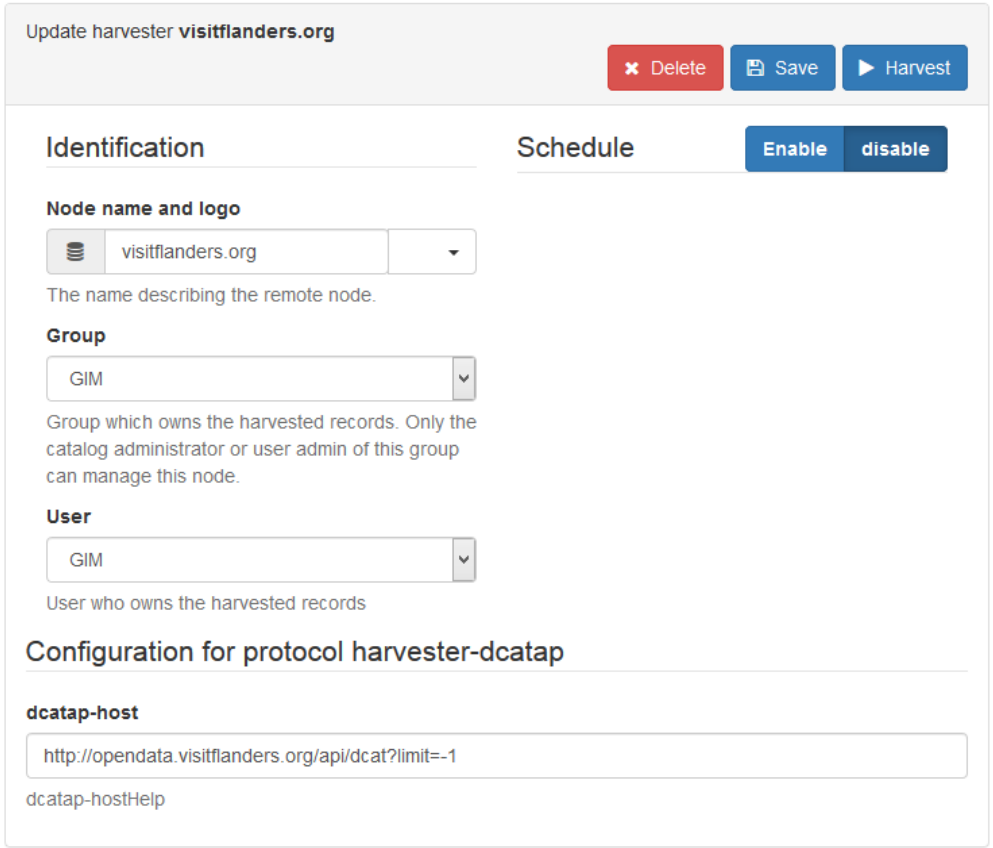
- XML is a nested data structure that can be queried with Xpath
- RDF is a graph-based data structure in different formats and arbitrarily structured that can be queried with SPARQL

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



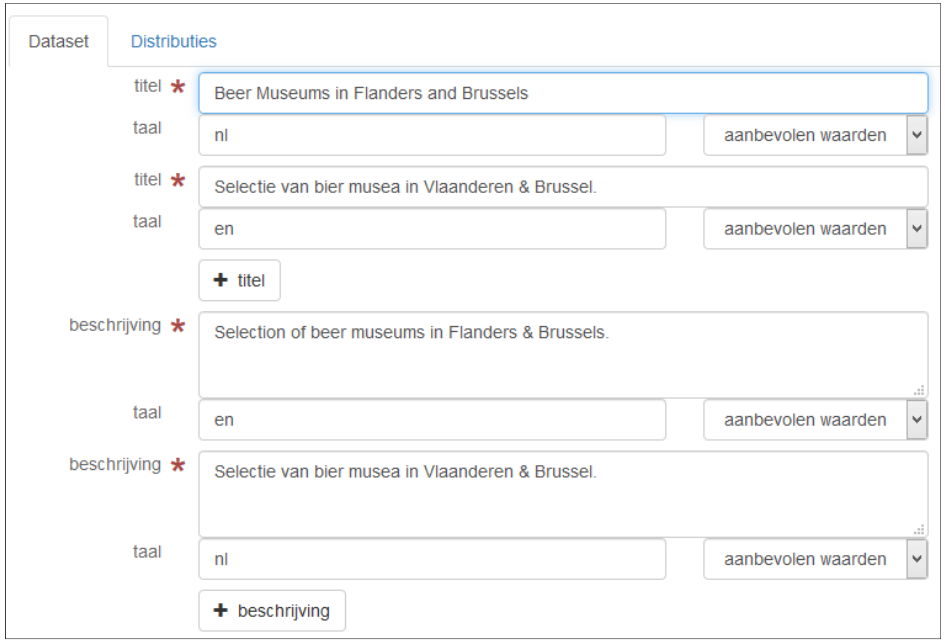
DCAT-AP HARVESTER

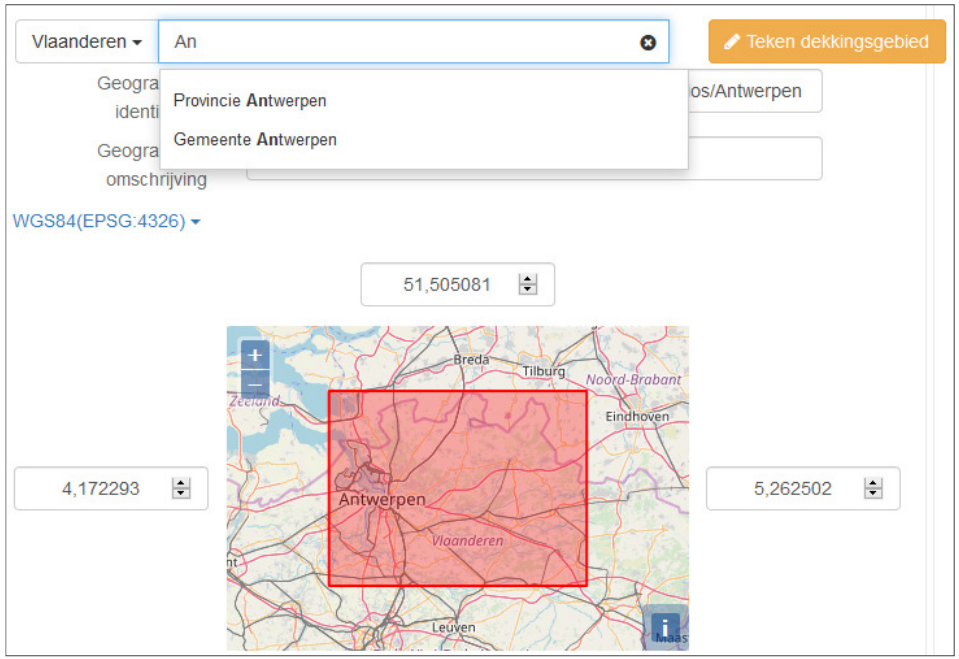
- Input: Turtle, RDF/XML, N3, JSON-LD



DCAT- AP EDITOR

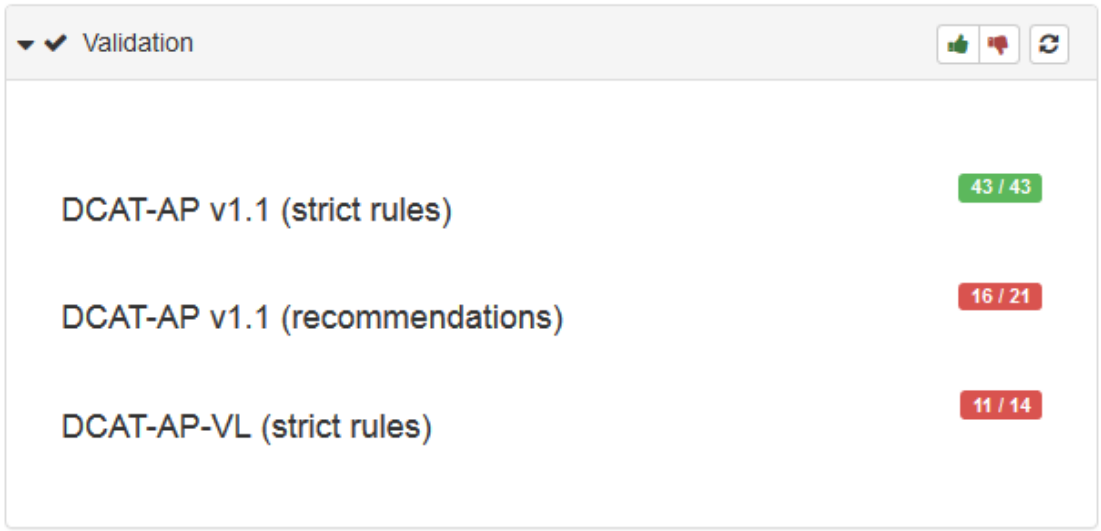
- Simple and advanced form for editing DCAT-AP records
- Multilingual
- Controlled vocabularies required by DCAT-AP as thesauri
- Reuse (with modification) of existing controls (e.g. spatial extent)





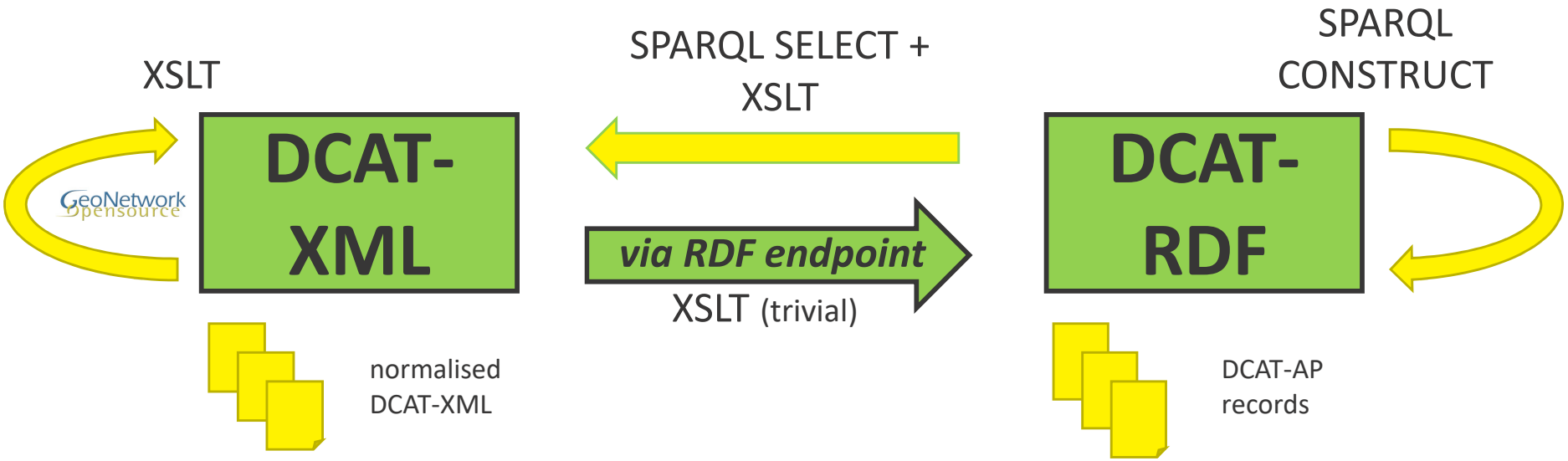
DCAT-AP VALIDATOR

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



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