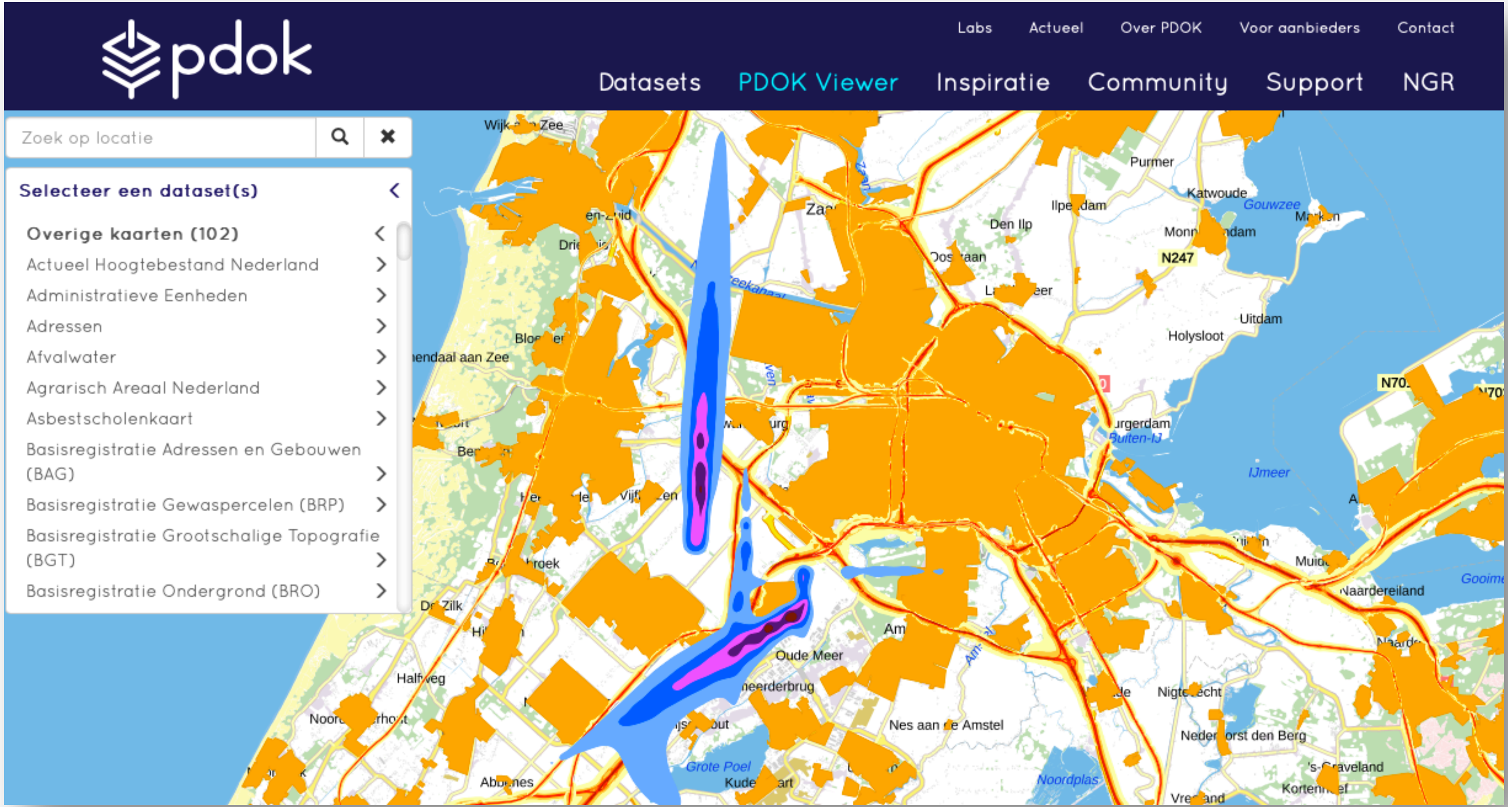
The Geospatial Knowledge Graph

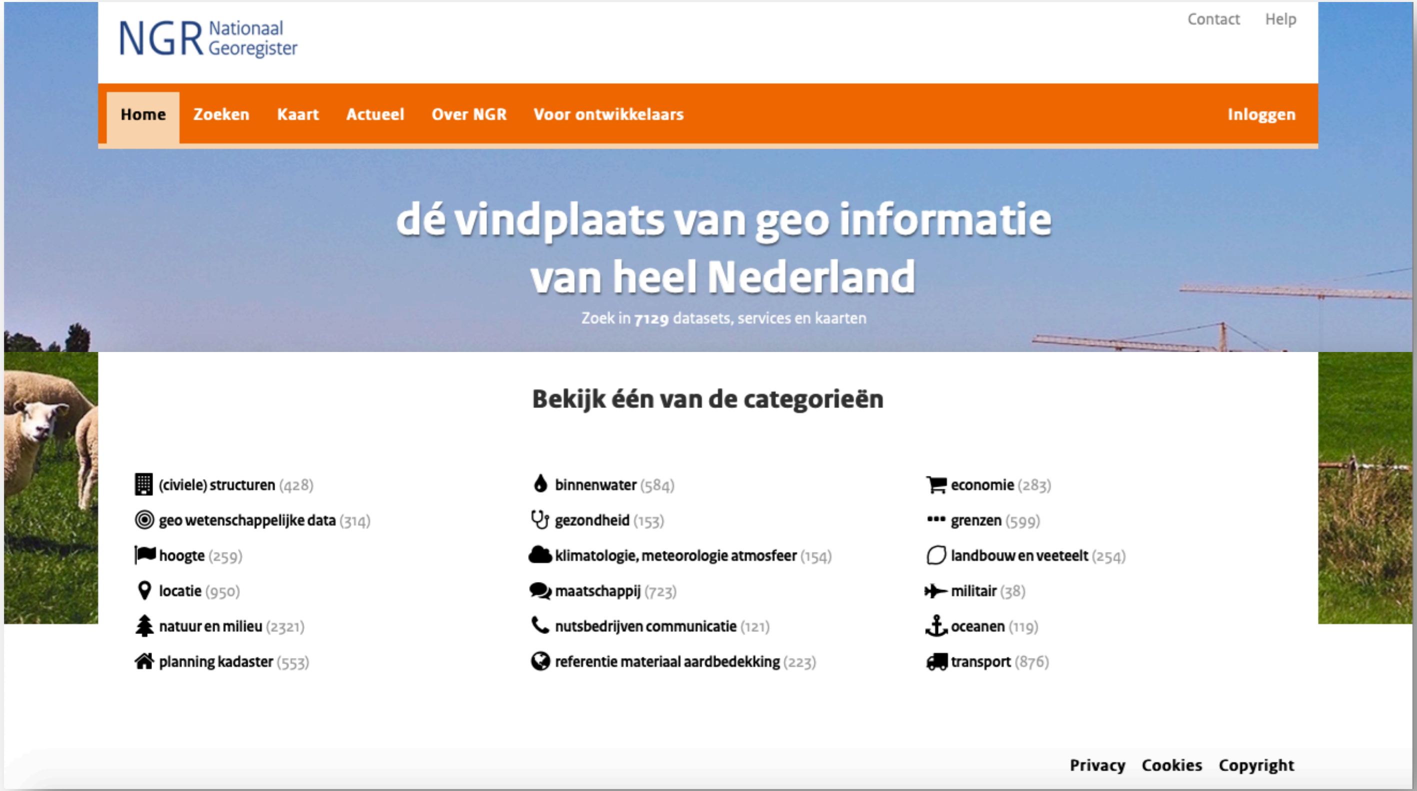
From traditional UML defined datasets to Linked Data

Marco Brattinga (Ordina) and Pano Maria (Skemu)
Semantics 2019, September 11th

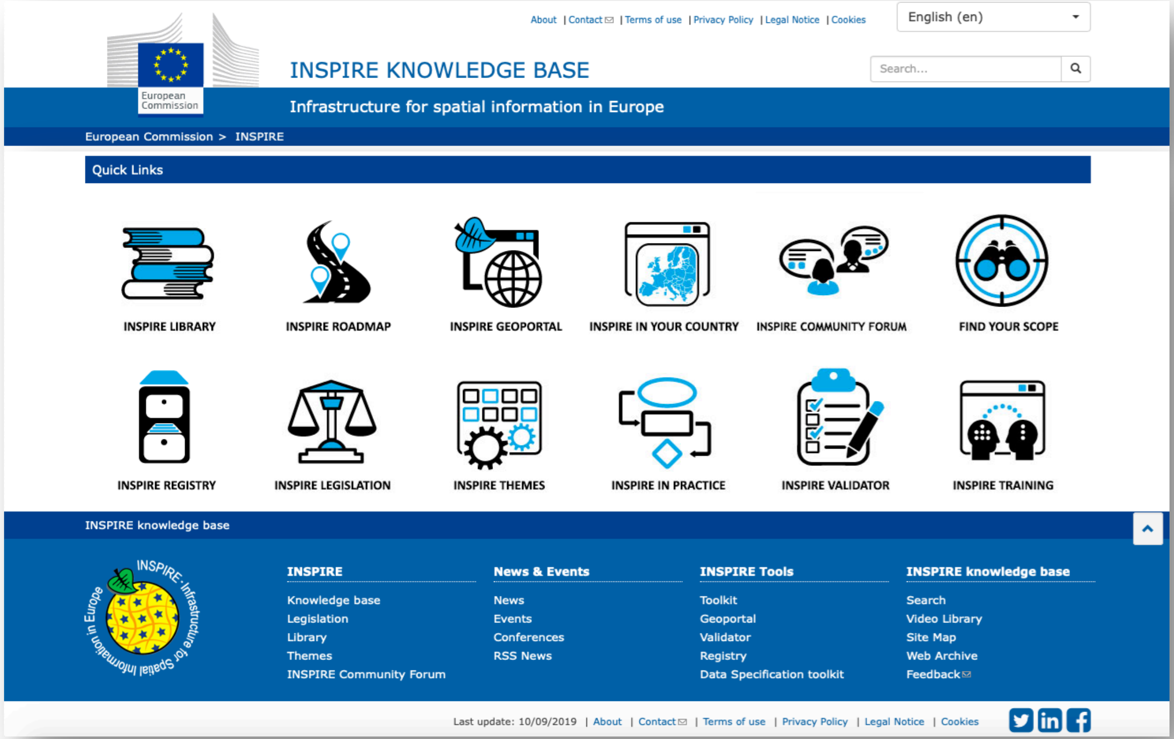
Geospatial data in the Netherlands



Geospatial data in the Netherlands



Geospatial data in ~~the Netherlands~~ Europe



Geospatial data in ~~the Netherlands~~ Europe

DIRECTIVES

DIRECTIVE 2007/2/EC OF THE EUROPEAN PARLIAMENT

of 14 March 2007

establishing an Infrastructure for Spatial Information in the Eu

METADATA

Article 5

1. Member States shall ensure that metadata are created for the spatial data sets and services corresponding to the themes listed in Annexes I, II and III, and that those metadata are kept up to date.

INSPIRE GEOPORTAL

INSPIRE IN YOUR COUNTRY

INSPIRE THEMES

INSPIRE IN PRACTICE

News & Events

News

Events

Conferences

RSS News

INSPIRE

Toolkit

Geoporta

Validator

Registry

Data Spe

Library

Themes

INSPIRE Community Forum

Last update: 10/09/2019 | About | Contact | Terms of u

Article 11

1. Member States shall establish and operate a network of the following services for the spatial data sets and services for which metadata have been created in accordance with this Directive:

(a) discovery services making it possible to search for spatial data sets and services on the basis of the content of the corresponding metadata and to display the content of the metadata;

(b) view services making it possible, as a minimum, to display, navigate, zoom in/out, pan, or overlay viewable spatial data sets and to display legend information and any relevant content of metadata;

(c) download services, enabling copies of spatial data sets, or parts of such sets, to be downloaded and, where practicable, accessed directly;

(d) transformation services, enabling spatial data sets to be transformed with a view to achieving interoperability;

(e) services allowing spatial data services to be invoked.

The promising opportunity!

If so much geospatial data is available and…

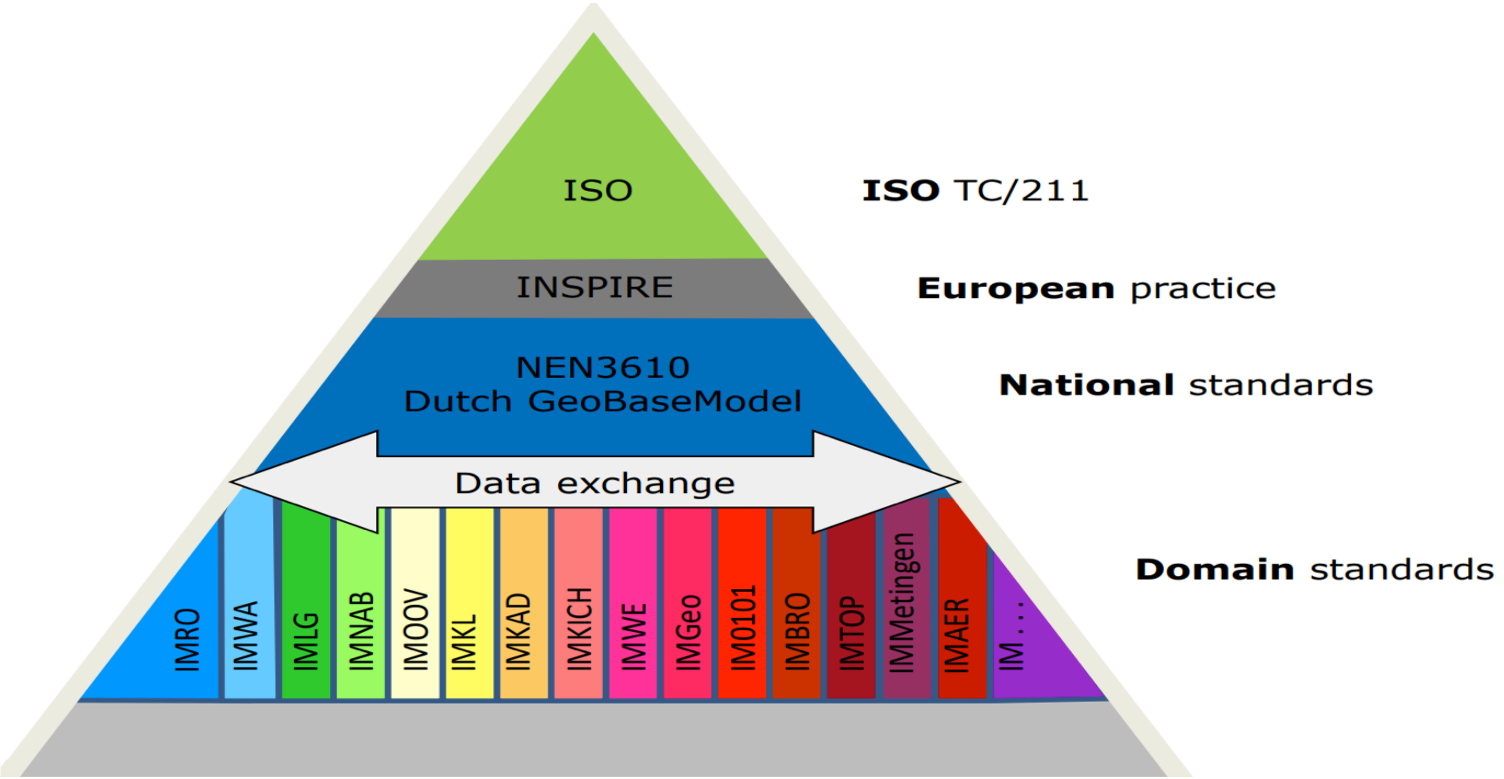
- It's metadata is available;
- The data itself is uniform…by law.

…it *should* be easy to use these datasets within a (RDF) knowledge graph!

So… what do we need?

1. Obtain a RDF model from the current metadata;
2. Use the RDF model to transform the original data to RDF;
3. Classify entities from different datasets under a common upper ontology;
4. Create links between entities that are more or less the "same".

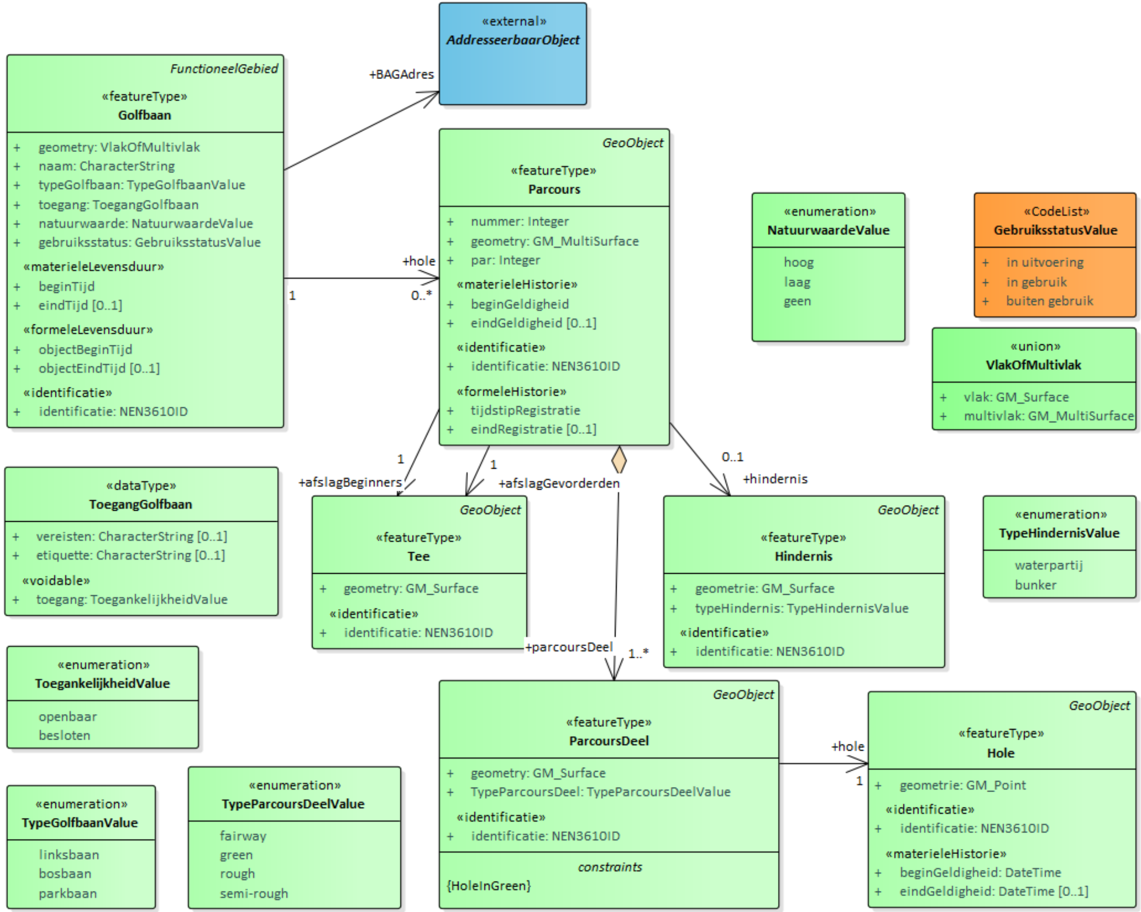
The current metadata model



NEN

- NEN supports standardization process in the Netherlands;
- Manages over 31000 standards accepted in the Netherlands;
  - International (ISO, IEC), European (EN) and national (NEN)
- NEN 3610 – Basic schema for geo-information
  - Several iterations since 1995;
  - Aims to simplify exchange of geo-information between parties by defining mutual concepts to describe the world;
  - UML based.

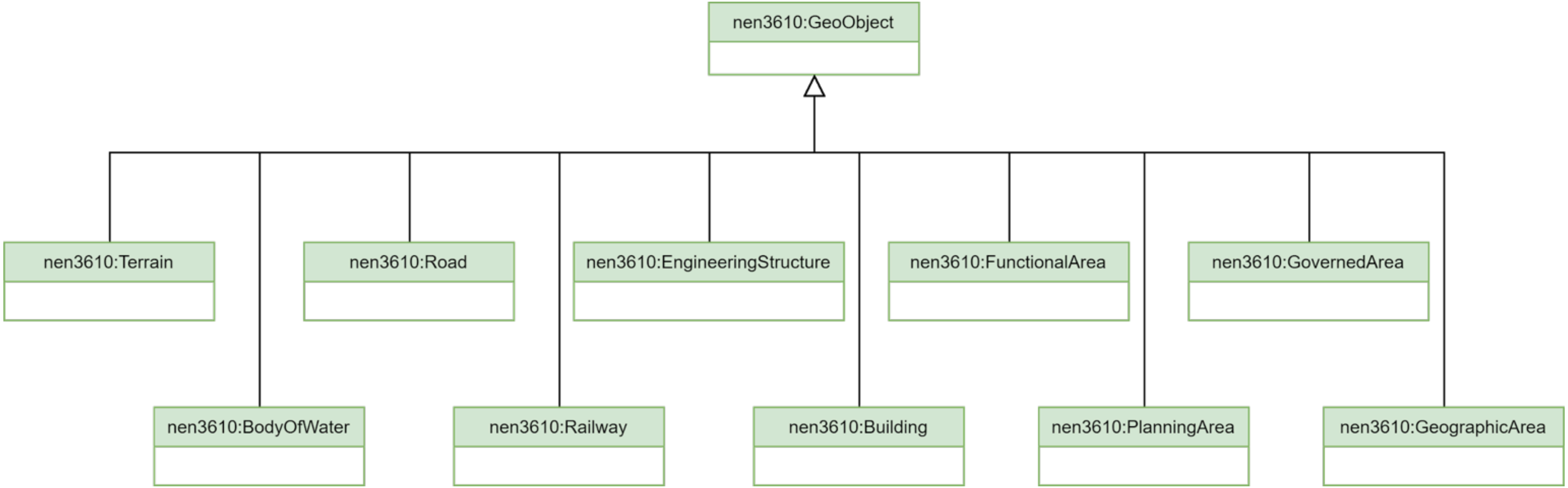
UML



ORDINA

skemu

NEN 3610 Top model (in UML)



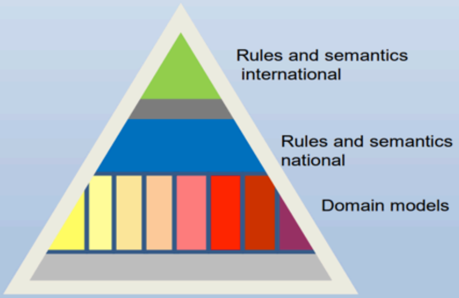
ORDINA skemu

Added value for a Linked Data model – from the perspective of the geo guy

UML-OO GeoBaseModel (NEN 3610)

Pyramid of specialzation

Silos extending common rules
Stack of profiles and extensions



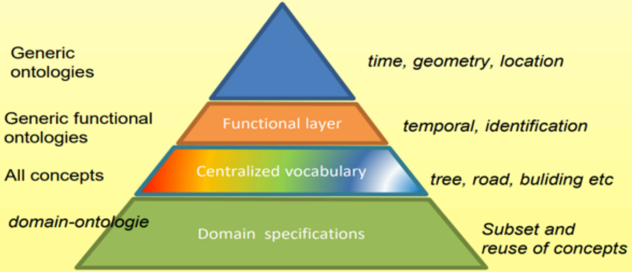
Use case defines universe of discourse

Domain defines vocabulary – domain standard

Linked data

Pyramid of reuse and reference

reuse and references of ontologies



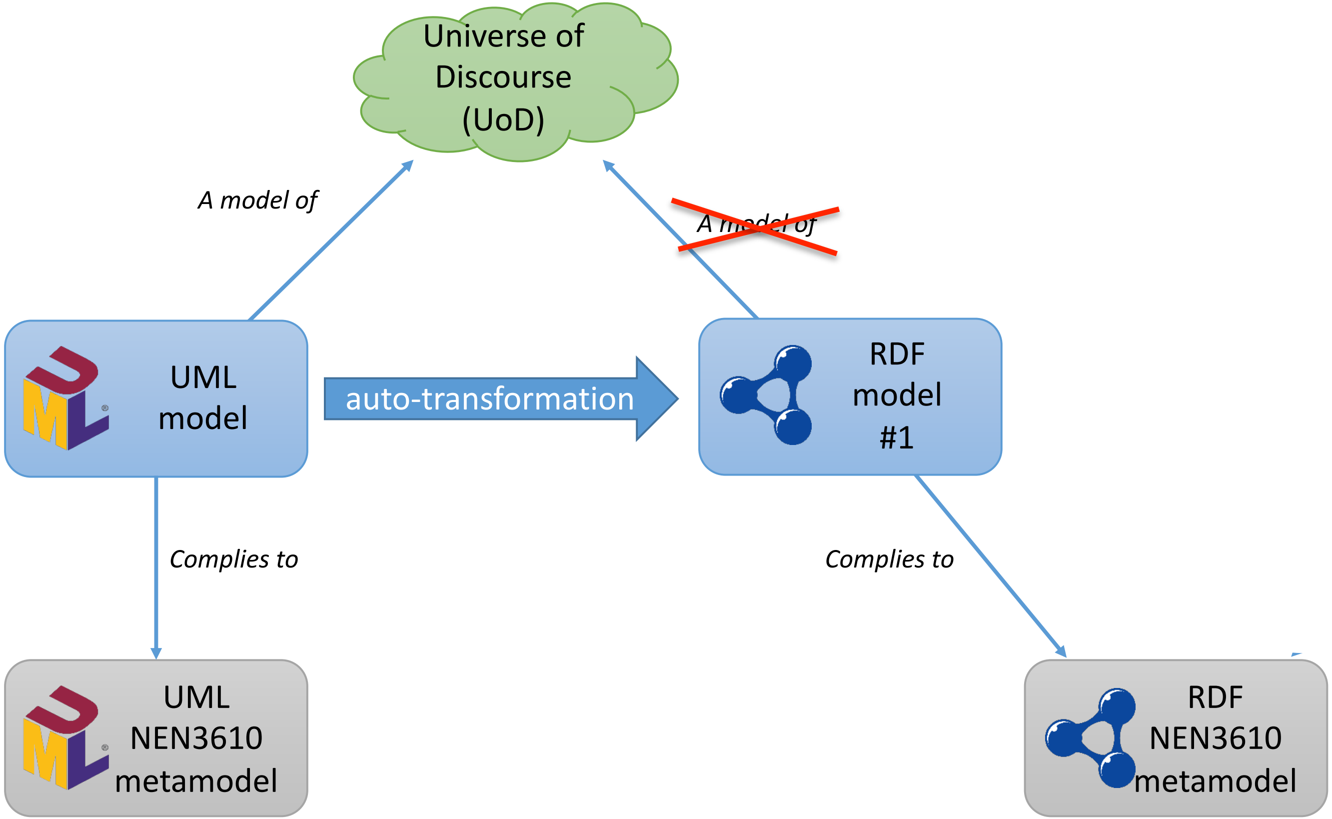
The Universe is the Universe of discourse

Anybody can say anything about anything

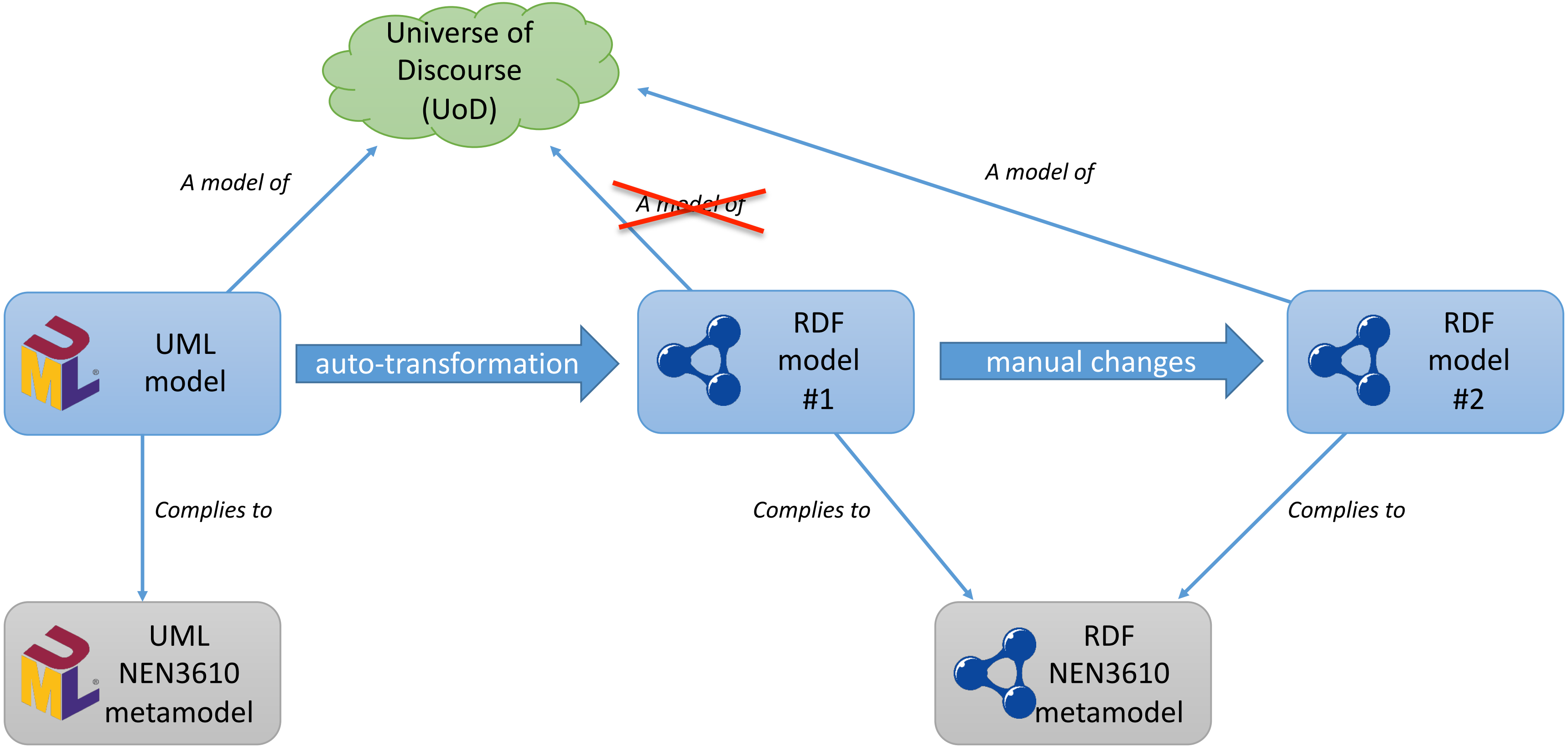
Step 1

obtaining the RDF model from the current metadata

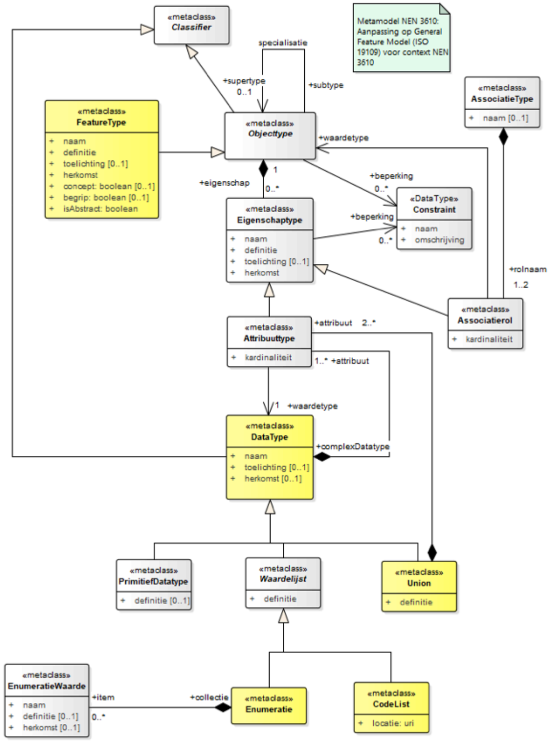
Step 1: obtaining the RDF model from the current metadata

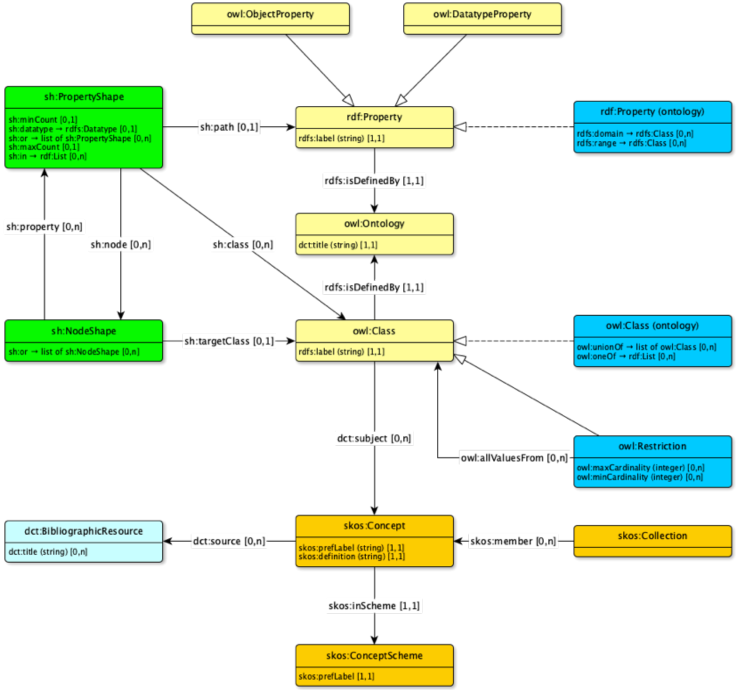


Step 1: obtaining the RDF model from the current metadata



UML and RDF NEN 3610 metamodel





ORDINA

skemu

Building on existing work

- ISO 19150-2: 2015 - Geographic information -- Ontology –
    - Part 2: Rules for developing ontologies in the Web Ontology Language (OWL)
- INSPIRE - Guidelines for the RDF encoding of spatial data (ARE3NA)
- OSLO² - Open standards for linked organizations
    - Enterprise Architect RDF Conversion Tool
- MIM - Metamodel for Information Modeling
    - https://docs.geostandaarden.nl/mim/mim10

ORDINA skemu

Challenges when transforming UML models to RDF

- There is no 1 to 1 mapping possible between UML information models and "good" RDF ontologies for linked data

- What is a "good" RDF ontology for linked data?
  - One that results in linkable data!
  - Recognizable/relatable things

- Why is this a problem?
  - UML information models model data, RDF ontologies model semantics
  - UML information models have implicit semantics
  - UML information models often reflect some degree of denormalization
  - UML information models often model registrations of real world objects, whereas RDF ontologies model real world objects "directly"

ORDINA skemu

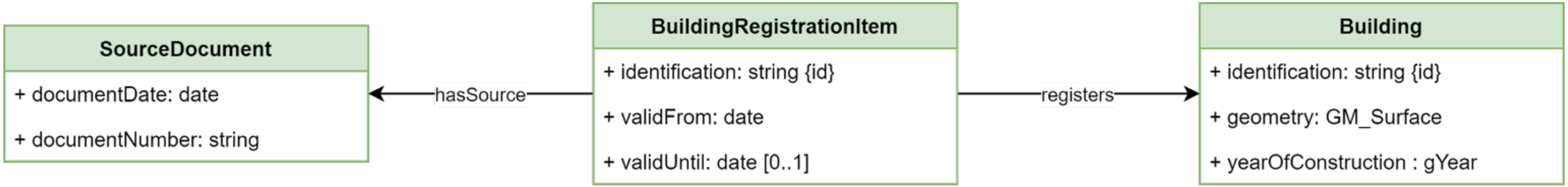
Example

Building
+ identification: string {id}
+ geometry: GM_Surface
+ yearOfConstruction : gYear
+ documentDate: date
+ documentNumber: string
+ validFrom: date
+ validUntil: date [0..1]

Example

Building
+ identification: string {id}
+ geometry: GM_Surface
+ yearOfConstruction : gYear
+ documentDate: date
+ documentNumber: string
+ validFrom: date
+ validUntil: date [0..1]

Example


SourceDocument
+ documentDate: date
+ documentNumber: string
hasSource
BuildingRegistrationItem
+ identification: string {id}
+ validFrom: date
+ validUntil: date [0..1]
registers
Building
+ identification: string {id}
+ geometry: GM_Surface
+ yearOfConstruction : gYear

Automatic translation of UML to RDF ontology?

- It is possible to create a (syntactically) *correct* translation;

- It is **not** possible to create a *valid* translation
  - Semantic information is (almost always) missing in the original model
  - Human interpretation is necessary.

Approach taken:

- Standardised automatic translation as a starting point.

Step 2

creating a common upper ontology

Step 2: creating a common upper ontology

- Not that hard: an upper ontology is already available in NEN 3610!

- Only thing missing:
  - An RDF vocabulary;
  - Strict rules as part of the standard to use this RDF vocabulary.

Approach taken:

- https://definities.geostandaarden.nl/def/nen3610#GeoObject
- https://definities.geostandaarden.nl/def/nen3610#Water
- etc…

Step 3

linking across information models

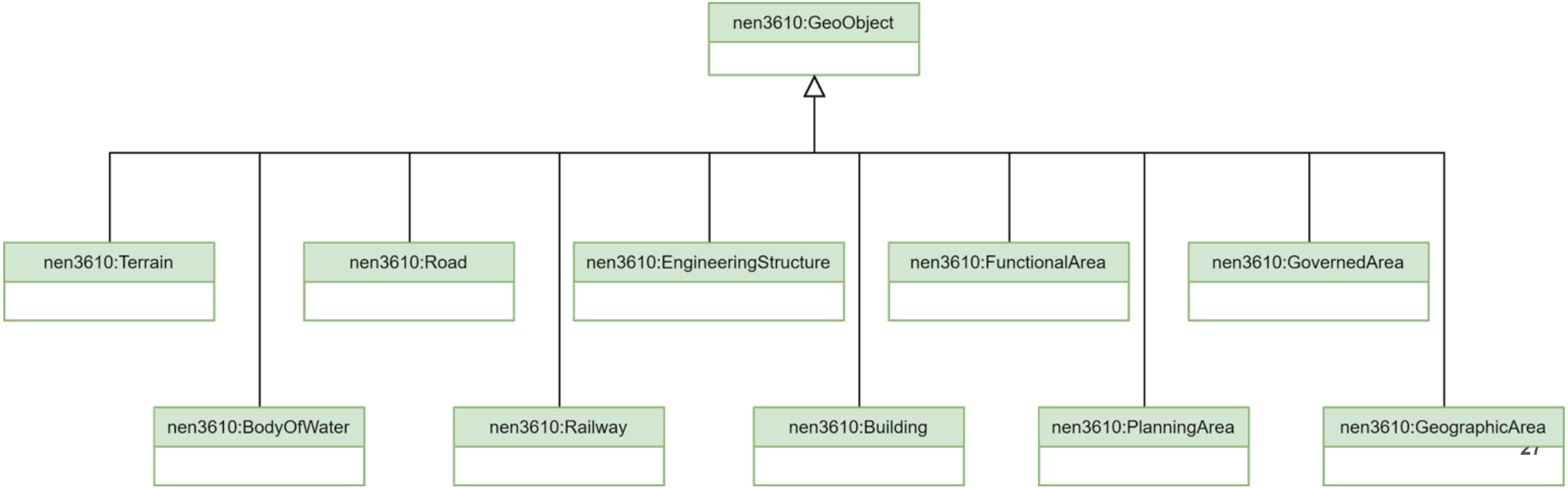
Step 3: linking across information models

- Linking to related objects;

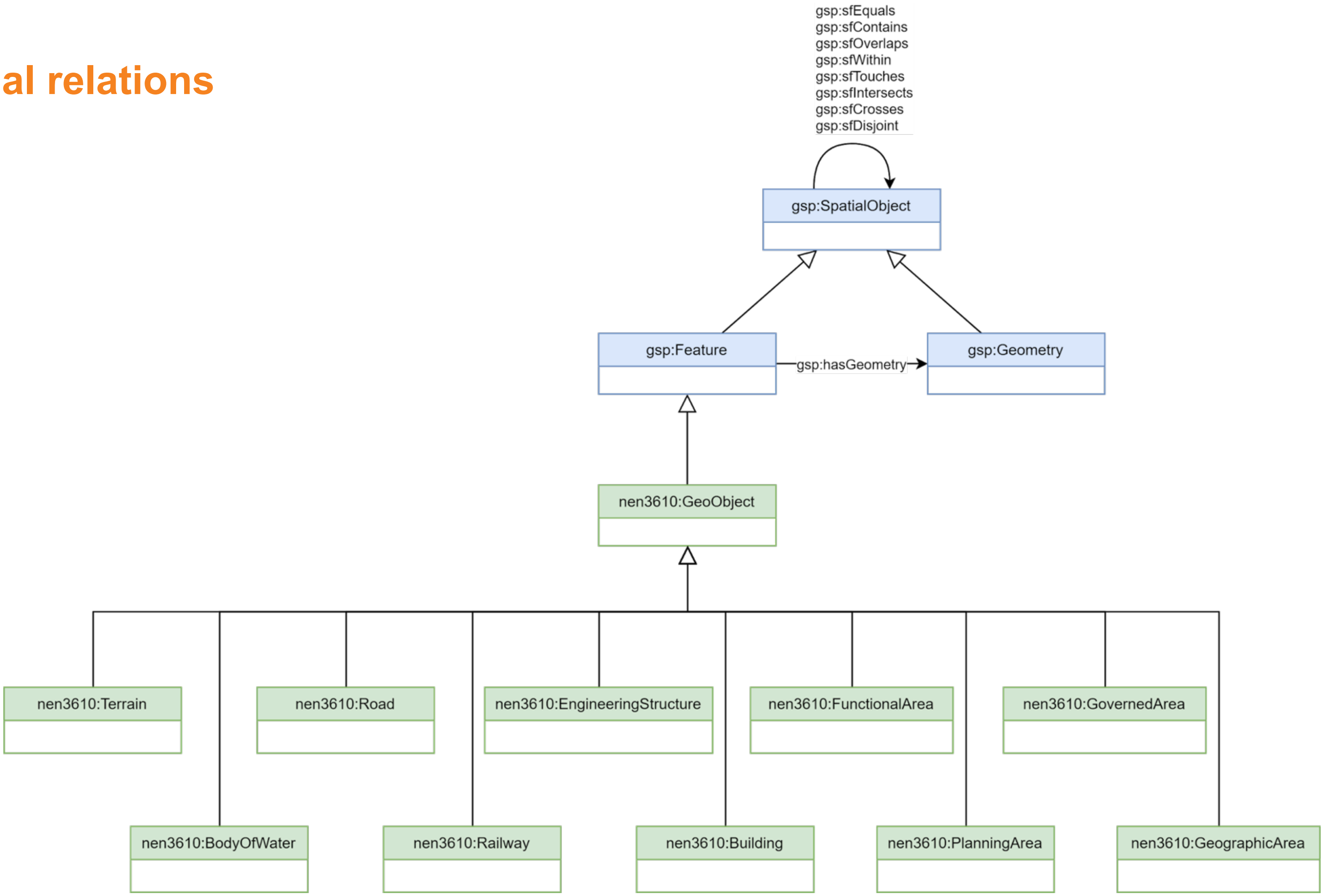
- Linking objects that are more or less “the same”, handling sameness:
  - Exactly the same as;
  - Almost the same as;
  - Somewhat the same as, but not really
  - …

- The NEN 3610 approach:
  - Using spatial relations for “sameness”!

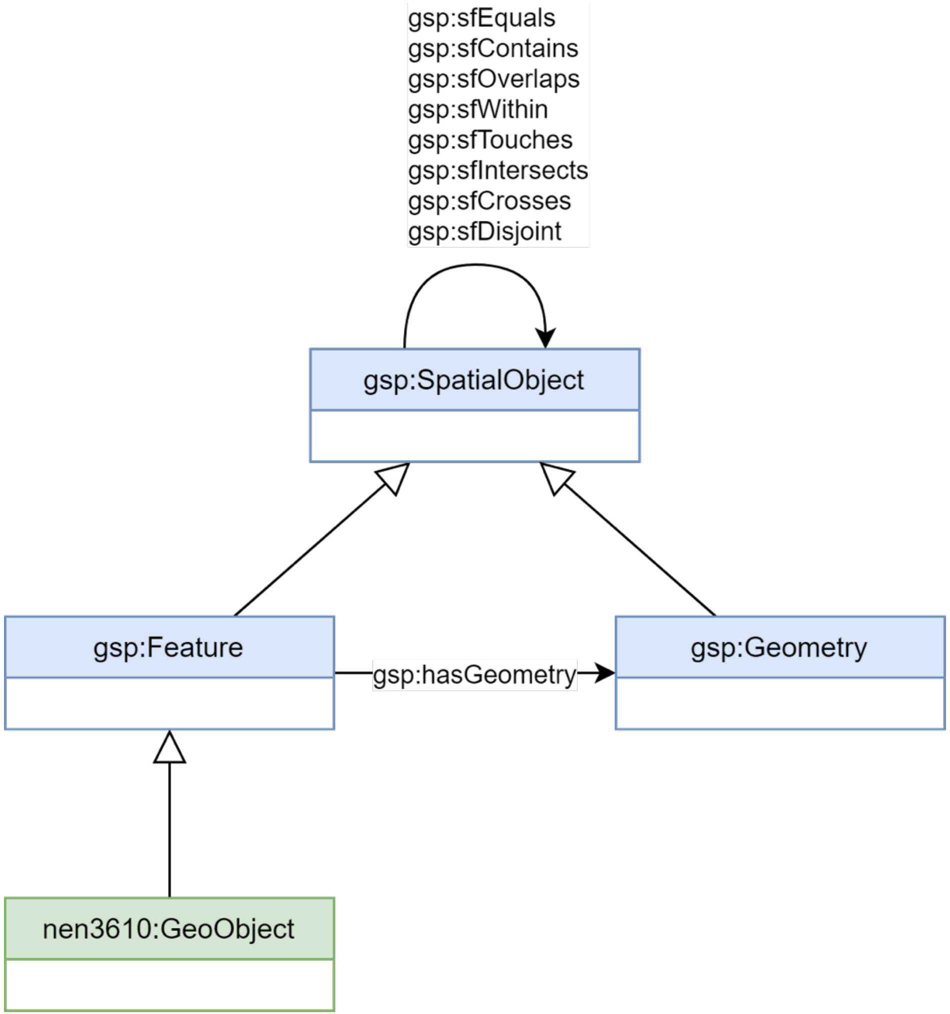
Spatial relations



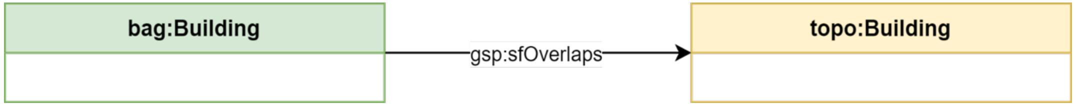
Spatial relations



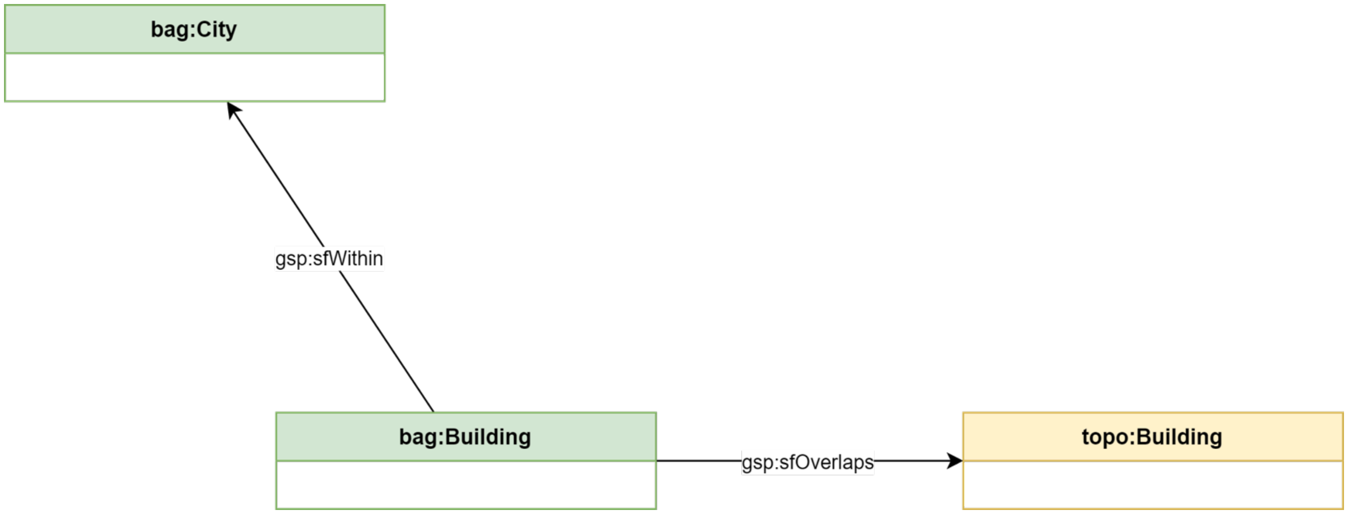
Spatial relations



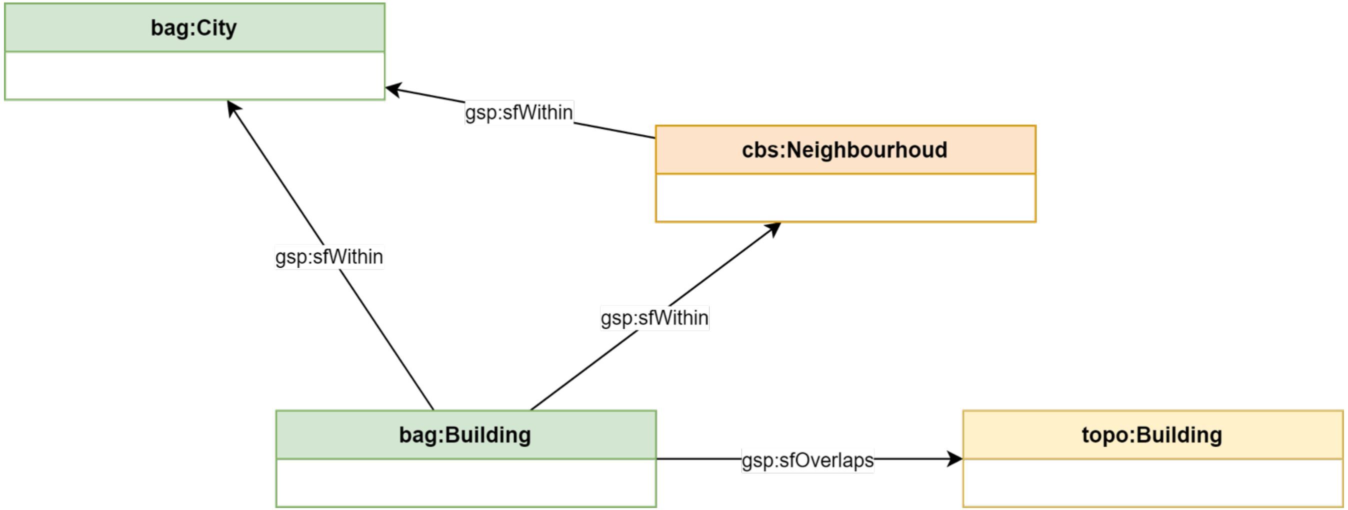
Spatial relations



Spatial relations



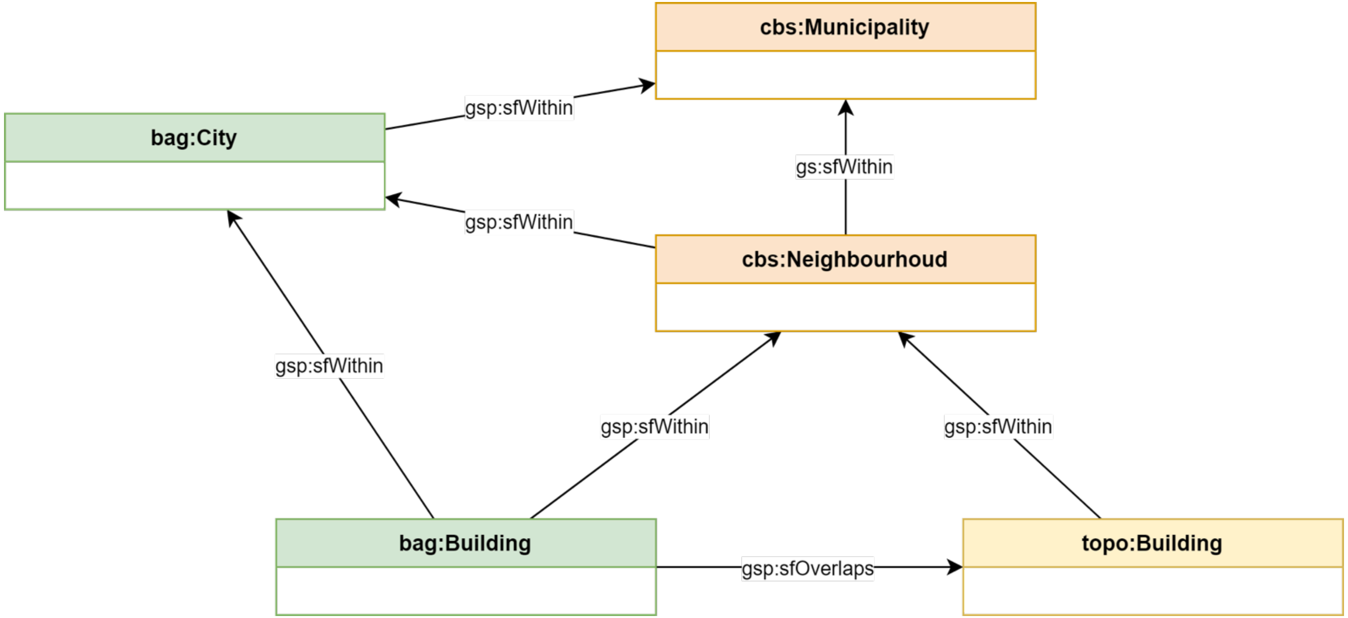
Spatial relations



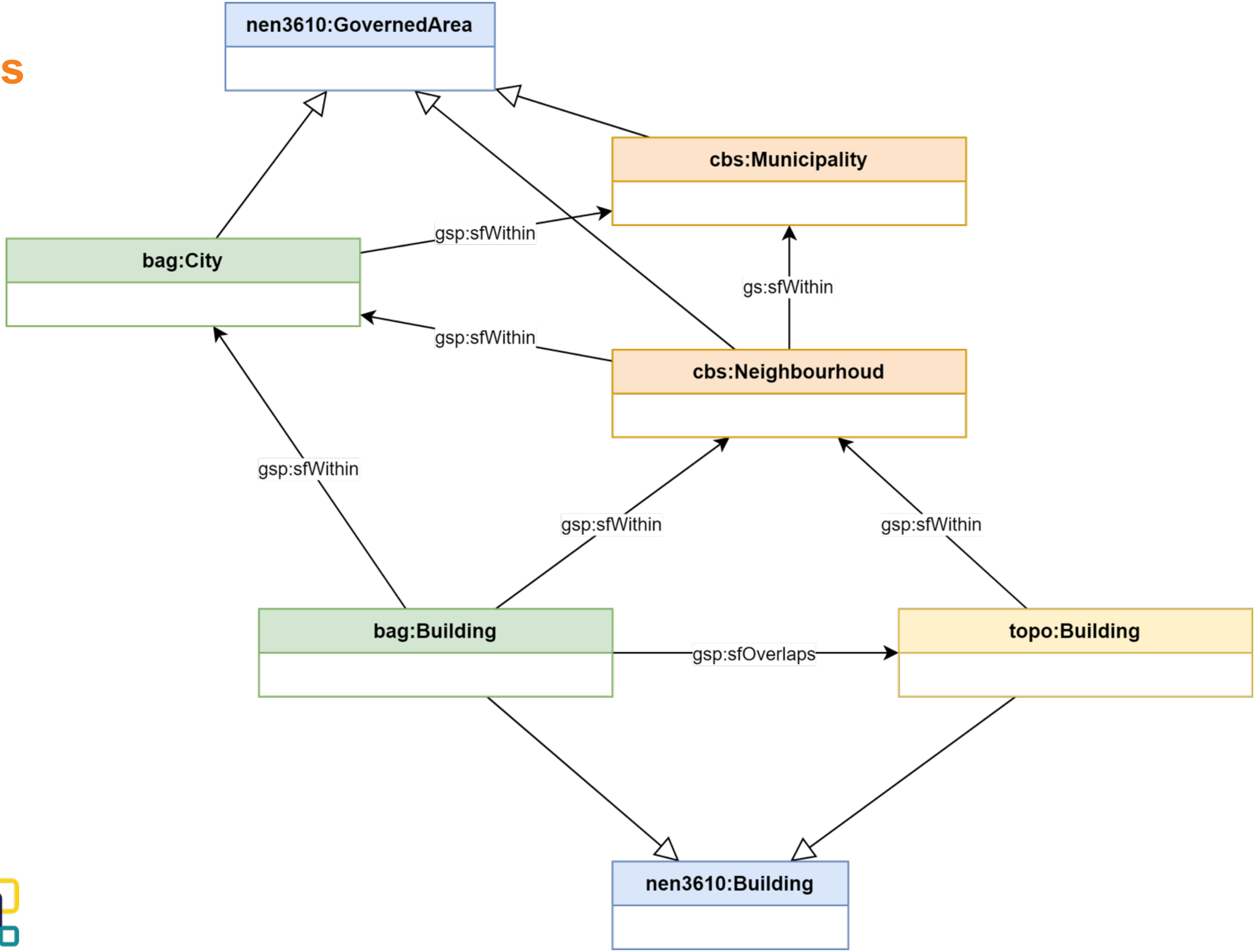
ORDINA

skemu

Spatial relations



Spatial relations

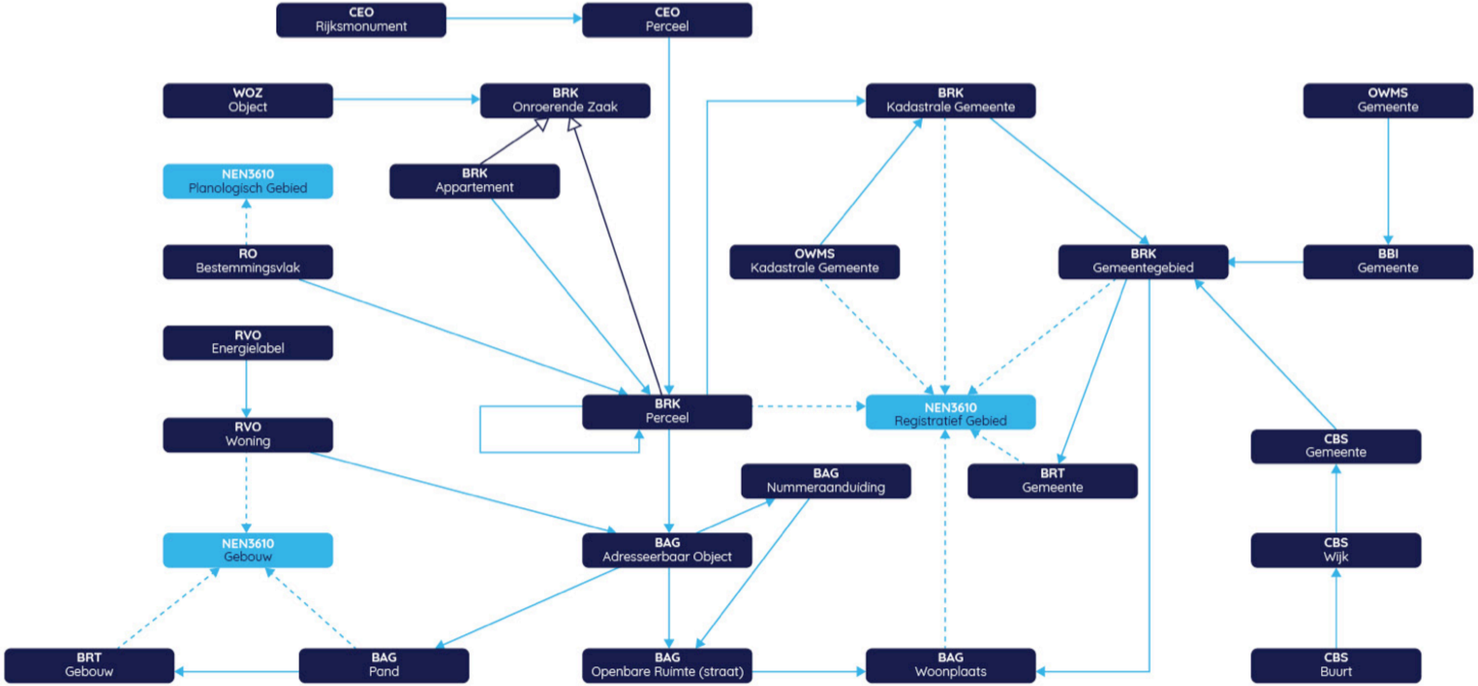


Demonstration

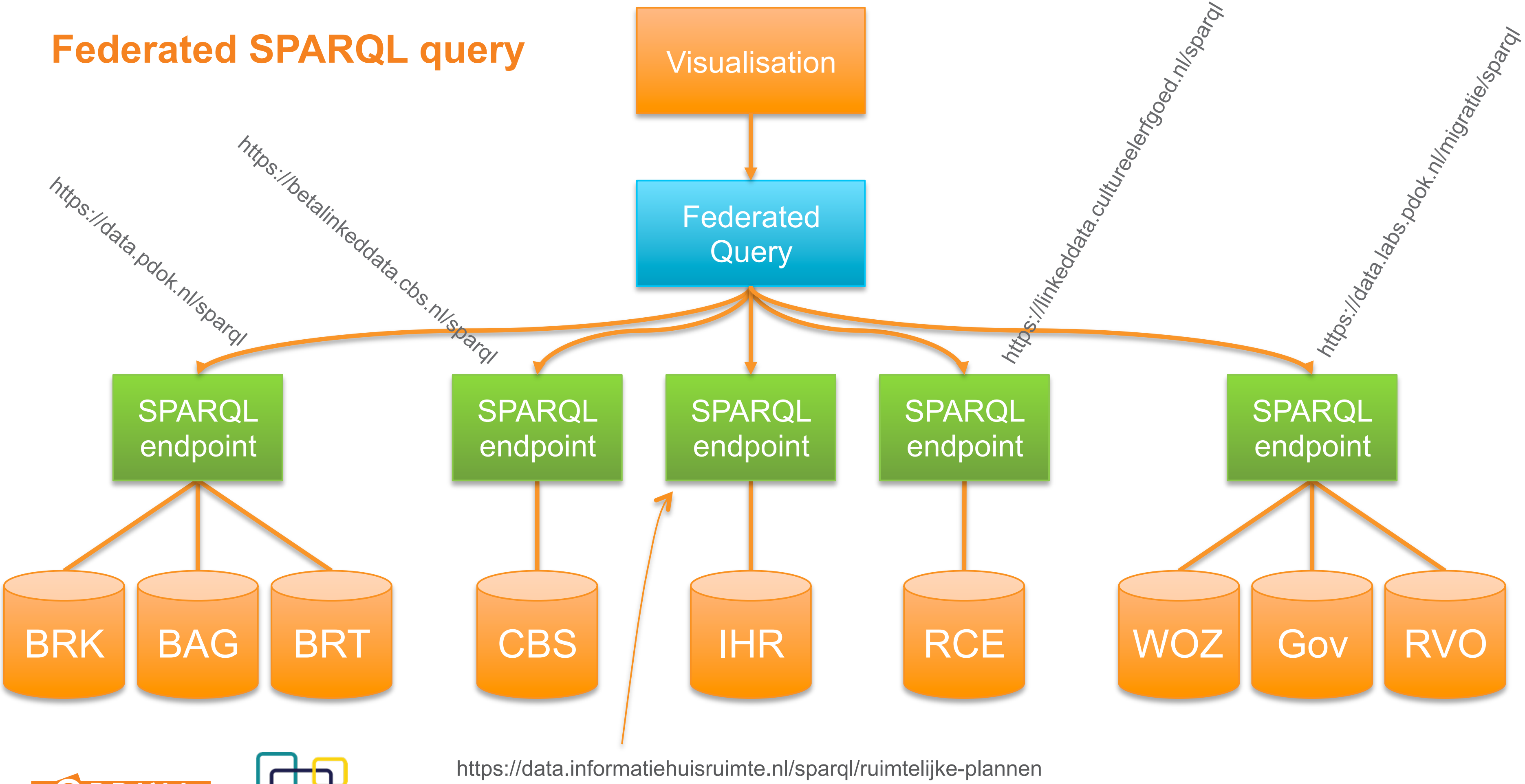
The Geospatial Knowledge Graph

Demonstration of the Geospatial Knowledge Graph

- Using NEN 3610 upper ontology to harmonise different classes;
- Using standard geospatial relations to mitigate the “sameness” problem.



Federated SPARQL query

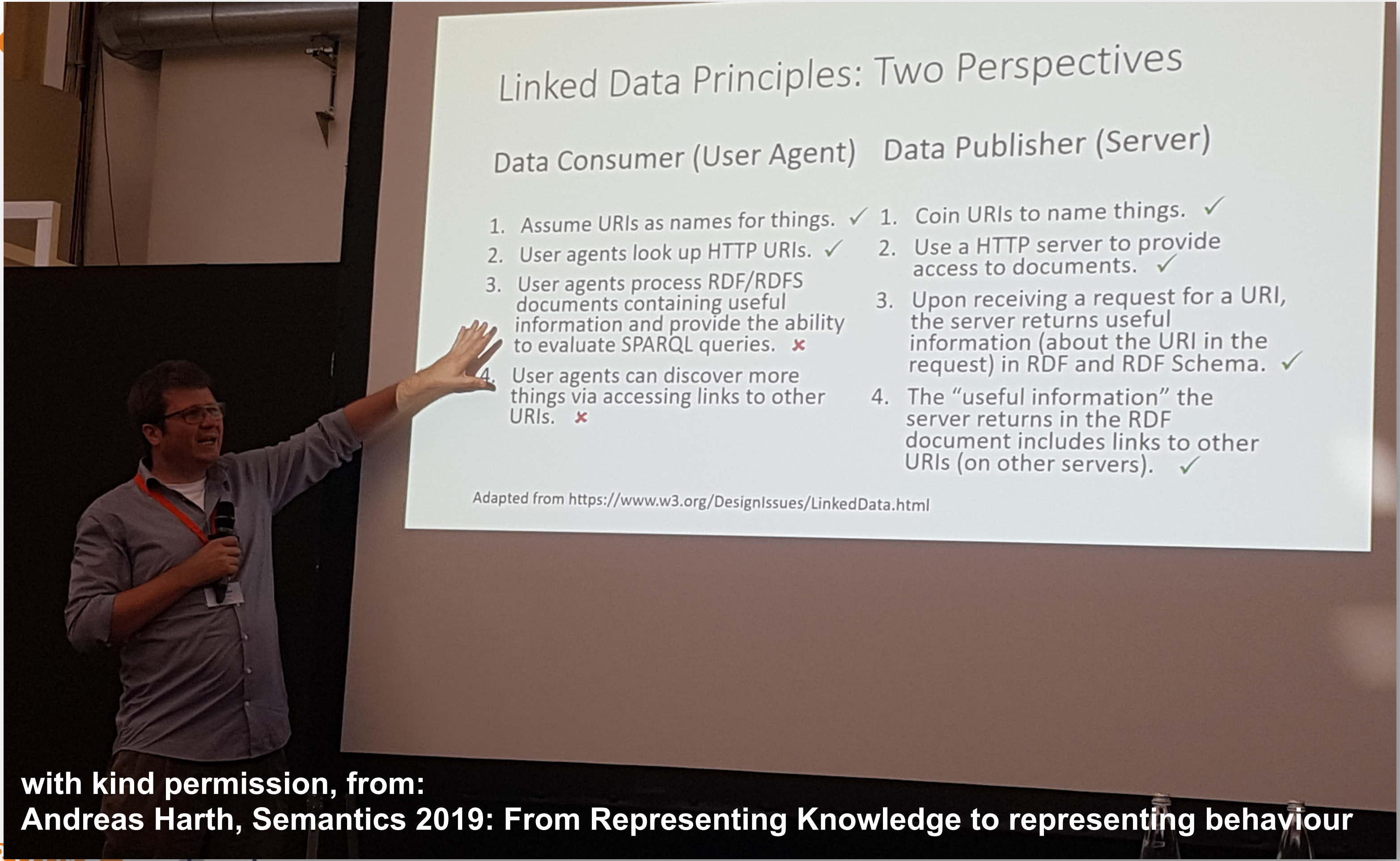


ORDINA skemu

Federated SPARQL query

```
CONSTRUCT {
        <@SUBJECT@> ?p ?o.
        ?o rdfs:label ?olabel
}
WHERE {
        {
                SERVICE <http://linkeddata.cultureelerfgoed.nl/sparql> {
                        <@SUBJECT@> ?p ?o
                        OPTIONAL {?o rdfs:label ?olabel}
                }
        } UNION
        {
                SERVICE <https://data.pdok.nl/sparql> {
                        <@SUBJECT@> ?p ?o
                        OPTIONAL {?o rdfs:label ?olabel}
                }
        } UNION
        {
                SERVICE <https://data.labs.pdok.nl/migratie/sparql> {
                        <@SUBJECT@> ?p ?o
                        OPTIONAL {?o rdfs:label ?olabel}
                }
        } UNION
        {
                SERVICE <https://data.informatiehuisruimte.nl/sparql/ruimtelijke-plannen> {
                        <@SUBJECT@> ?p ?o
                        OPTIONAL {?o rdfs:label ?olabel}
                }
        } UNION
        {
                SERVICE <https://betalinkeddata.cbs.nl/sparql> {
                        <@SUBJECT@> ?p ?o
                        OPTIONAL {?o rdfs:label ?olabel}
                }
        }
}
```

Fe



with kind permission, from:
Andreas Harth, Semantics 2019: From Representing Knowledge to representing behaviour

skemu

Wrap-up

- Translating geospatial metadata to corresponding RDF models is possible and feasible, with some human help;
- Translating geospatial data can be done automatically, if the model is available;
- An standardised upper ontology for geospatial features is now available, directly derived from already used standards (NEN 3610);
- Using spatial features for “sameness” works, creating them might be hard work;
- Federated SPARQL queries work for traversing a knowledge graph, but SLA dependencies dictate more solutions.

Marco Brattinga
Senior principal expert

marco.brattinga@ordina.nl

Pano Maria
Information architect

pano@skemu.com

@panomaria