

Integrated Spatio-temporal Data for Earth Observation

A RDF dataset of Territorial Units
with their Land Cover

Ba-Huy Tran, Nathalie Aussenac-Gilles, Catherine
Comparot, Cassia Trojahn

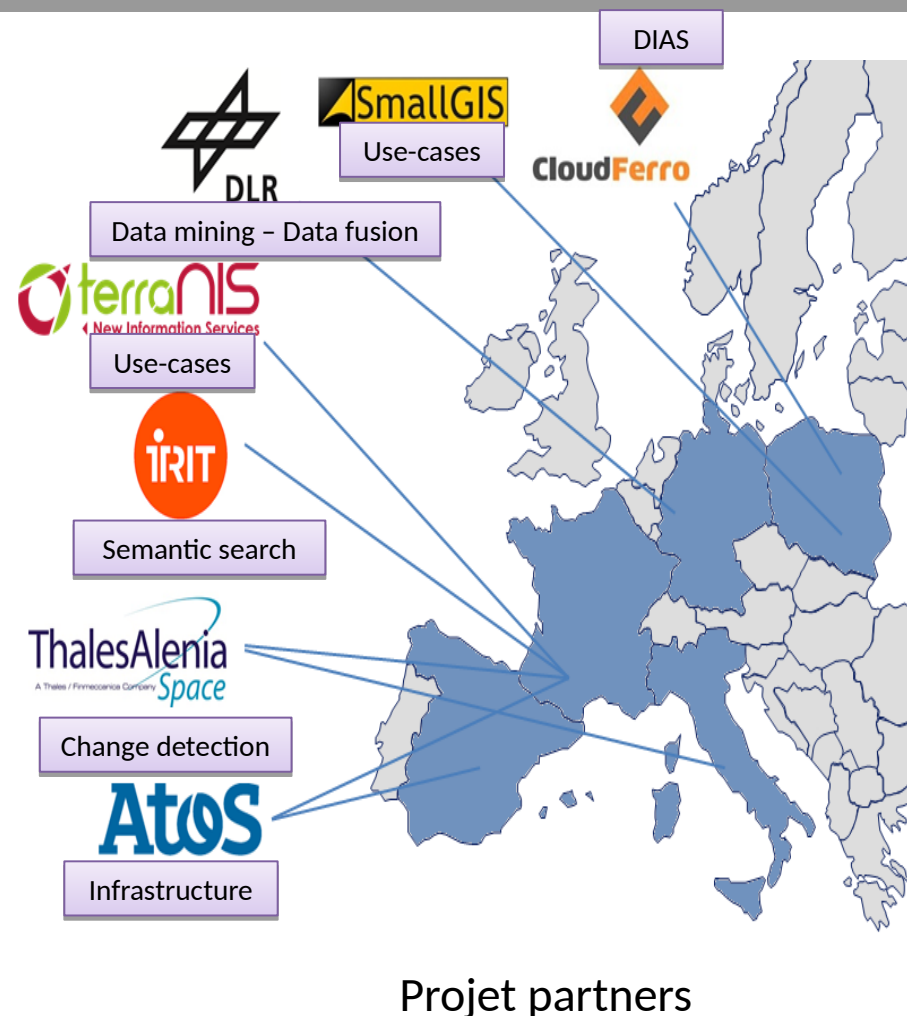
Prenom.Nom@irit.fr

Contents

1. Introduction
 - Context
 - Data sources
2. Semantic integration
 - Semantic representation
 - Data integration
 - System architecture
3. Application
4. Conclusion

1.1 Context

- CANDELA: Copernicus Access Platform Intermediate Layers Small Scale Demonstrator
- Objectives: Development of tools, which allow users to access and analyze of large volume of EO data (05/2018 – 11/2020)
- Semantic search:
 - Enable the search for enriched EO data through different heterogeneous data sources.
 - These sources are selected in keeping with the requirements and scenarios of use-cases.



1.1 Context

- Use case
 - **Vineyards, Nouvelle Aquitaine region**
 - **Urban expansion, Nouvelle Aquitaine region**
 - Primeval Forest, Białowieża – Poland
 - Natura 2000 reservation sites - Europe



1.2 Data sources

1. Administrative unit

- OpenStreetMap based dataset¹:
 - Updated yearly based on OpenStreetMap data
- GeoZones dataset²:
 - Comes from a certified public service;
 - Provides a common geospatial and administrative repository based on open data

GeoZones is chosen because:

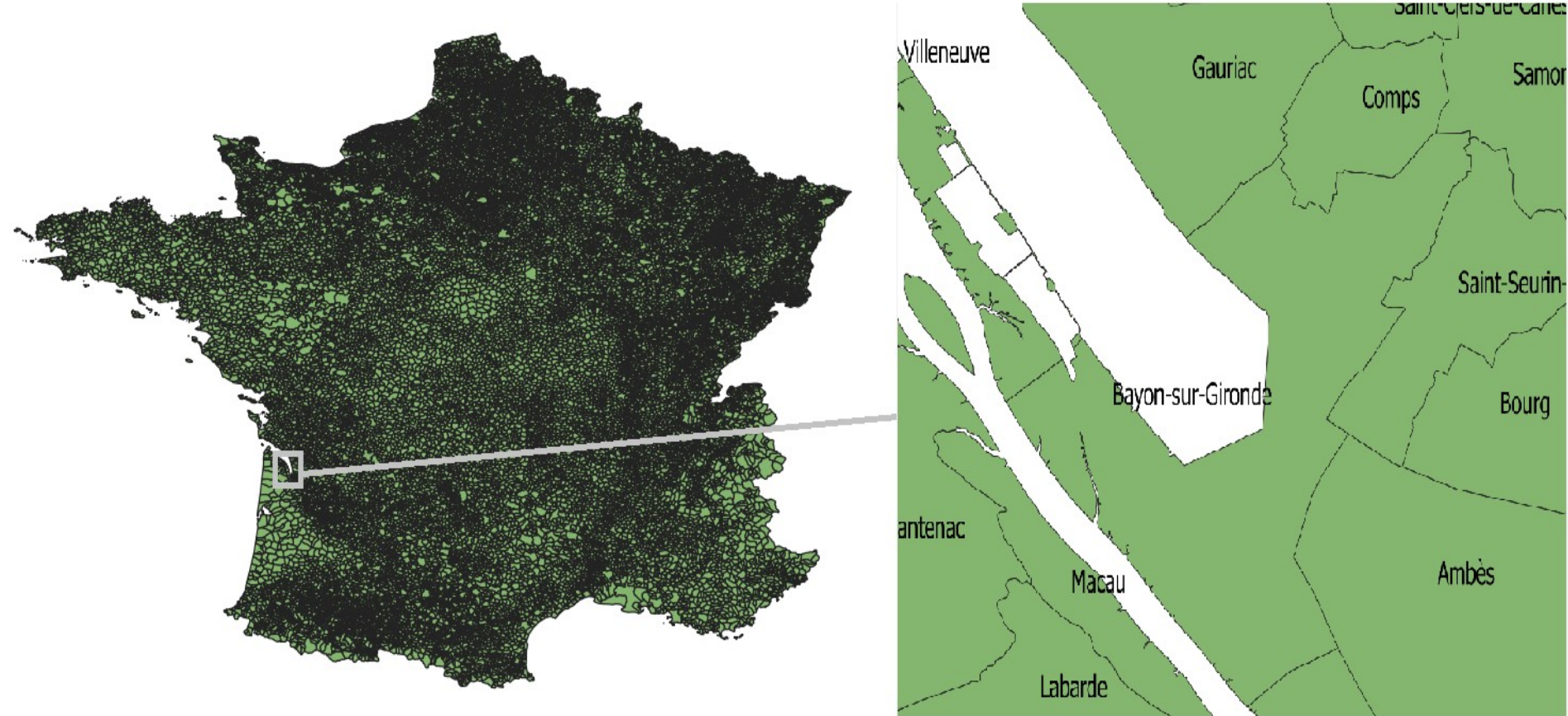
- More details about administrative units
- Additional information such as the area, population and especially links to other open datasets (for example, Geonames, INSEE, Wikipedia or Wikidata)

1. <https://www.data.gouv.fr/en/datasets/decoupage-administratif-communal-francais-issu-d-openstreetmap/>

2. <https://www.data.gouv.fr/en/datasets/geozones/>

1.2 Data sources

1. Administrative unit

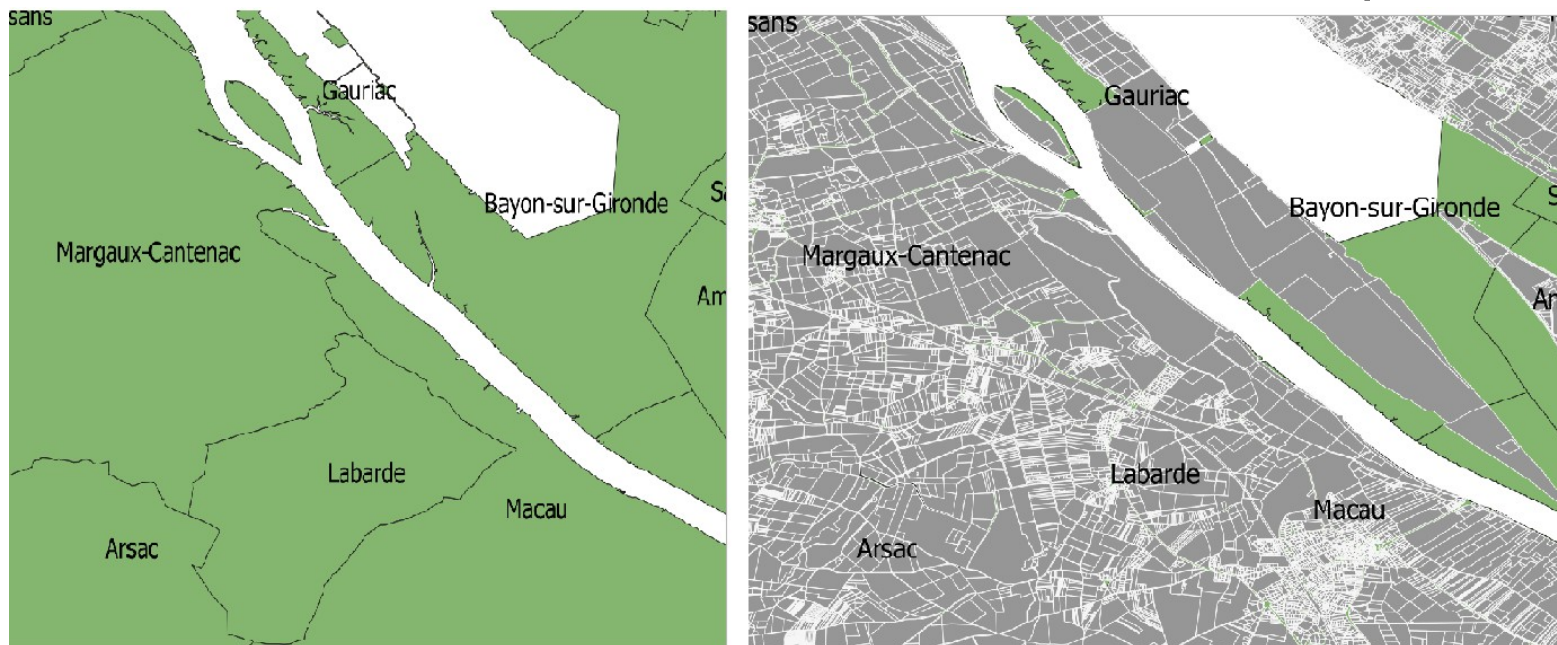


Visualization of French communes in 2019

1.2 Data sources

2. Land register

- Available from the French government data website¹
- Either in GeoJSON format or shapefiles
- Indicates the identification and the localization of parcels



A view of communes in the Gironde department without (left) and with (right) parcels

1. <https://cadastre.data.gouv.fr/datasets/cadastre-etalab>

1.2 Data sources

3. Land cover

- Global Land Cover SHARE dataset¹
 - Created by FAO in 2012, not updated from then
- Corine Land Cover dataset²
 - The most standard as it uses the Corine Land cover vocabulary.
 - The two most recent versions of the dataset were published in 2012 and 2018.
- Cesbio dataset³
 - Updated yearly
 - Only available for France

1. <http://www.fao.org/geospatial/resources/detail/en/c/1036591/>

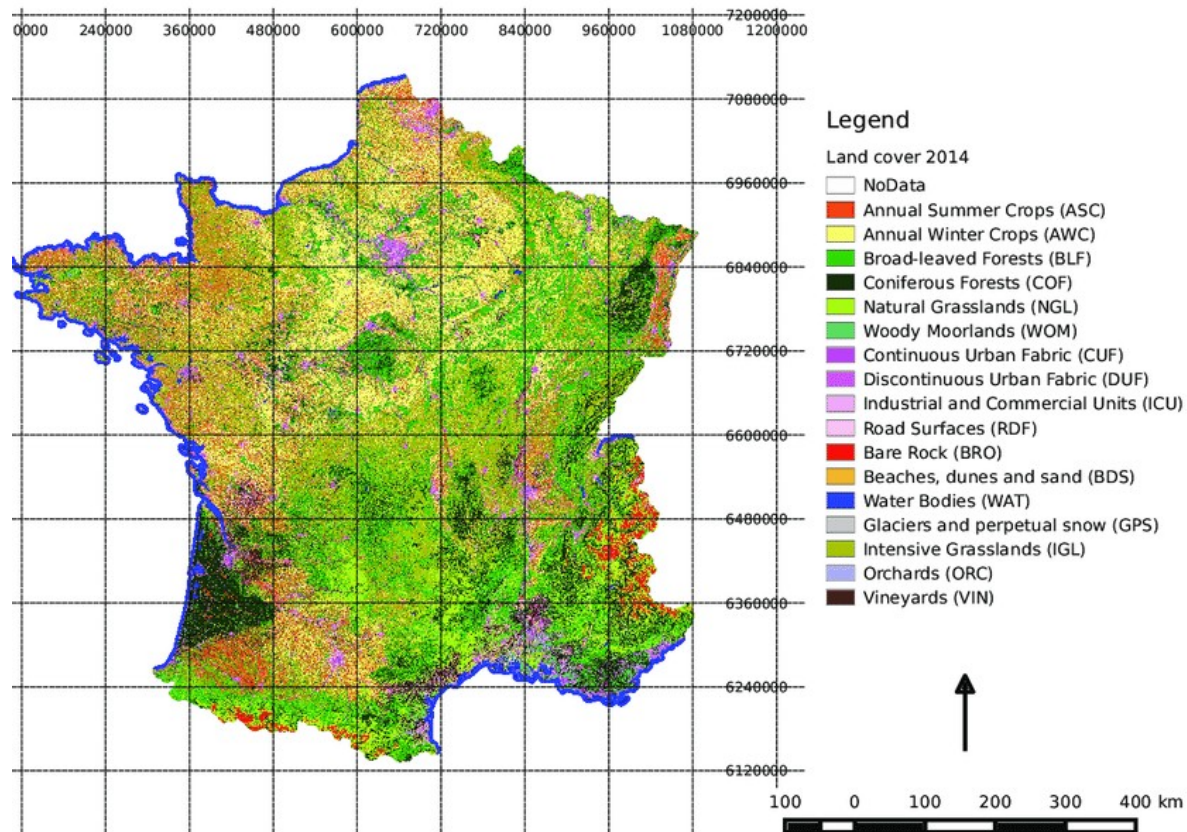
2. <http://osr-cesbio.ups-tlse.fr/oso/>

3. <https://www.data.gouv.fr/en/datasets/corine-land-cover-occupation-des-sols-en-france>

1.2 Data sources

3. Land cover

- Cesbio dataset



Land cover map for France for the year 2014 [CesInglada Jordi et al. ,2017]

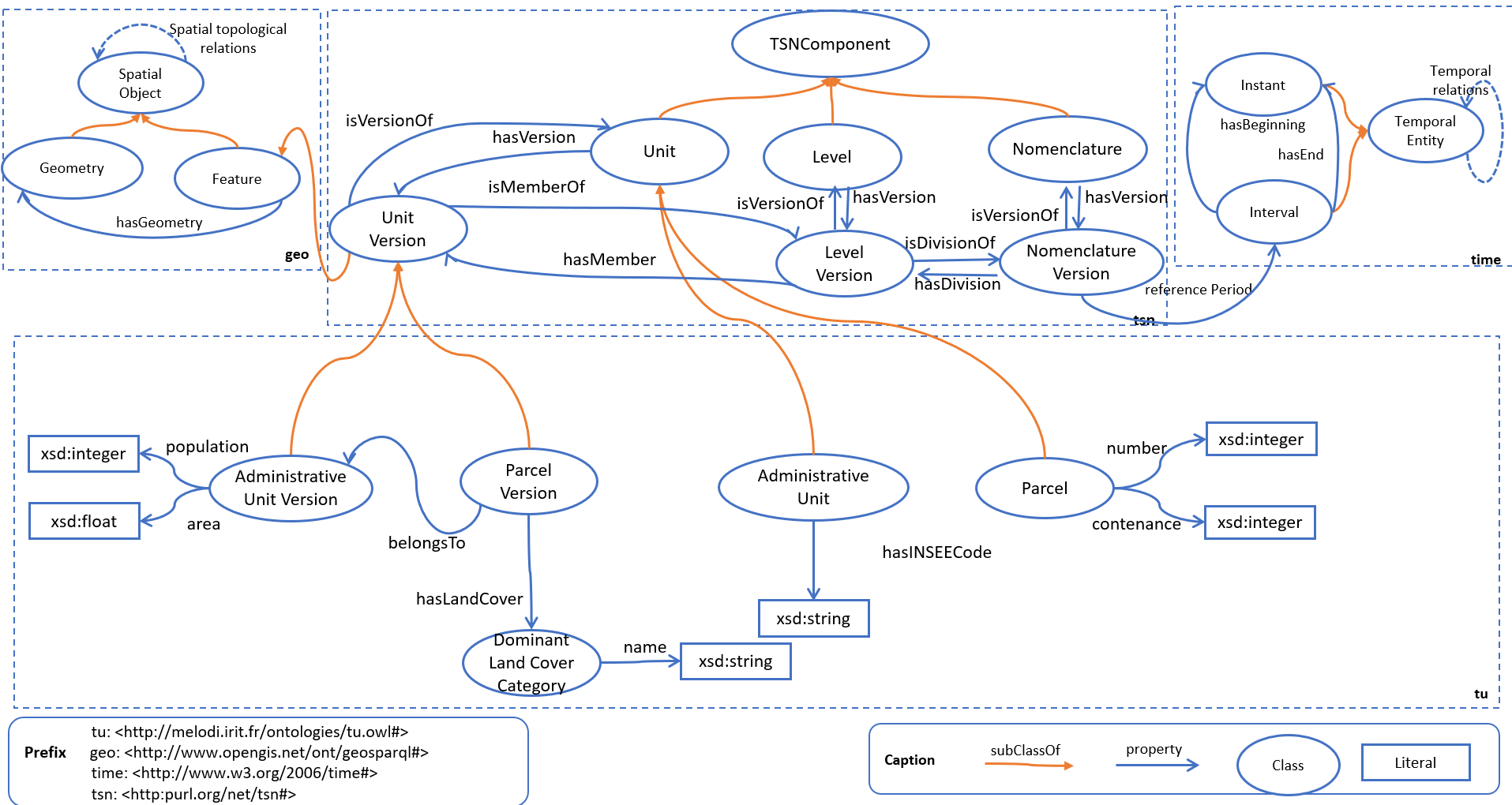
2. Semantic integration

- Relies on defining an ontology that serves as basis for the integration together with a process that converts the different formats and data to instances of this ontology
- Can be divided into two main stages:
 - Semantic representation: Build a modular ontology with specific parts adapted from each source schemes
 - Data integration: Integrate the data sources based on this ontology and a set of transformation rules

2.1 Semantic representation

- Territorial unit ontology: composed of two parts
 - Generic part: reuses existing vocabularies
 - ✓ TSN Ontology
 - ✓ OWL-Time ontology
 - ✓ GeoSPARQL ontology
 - Specific part: dedicated to data to be integrated

2.1 Semantic representation

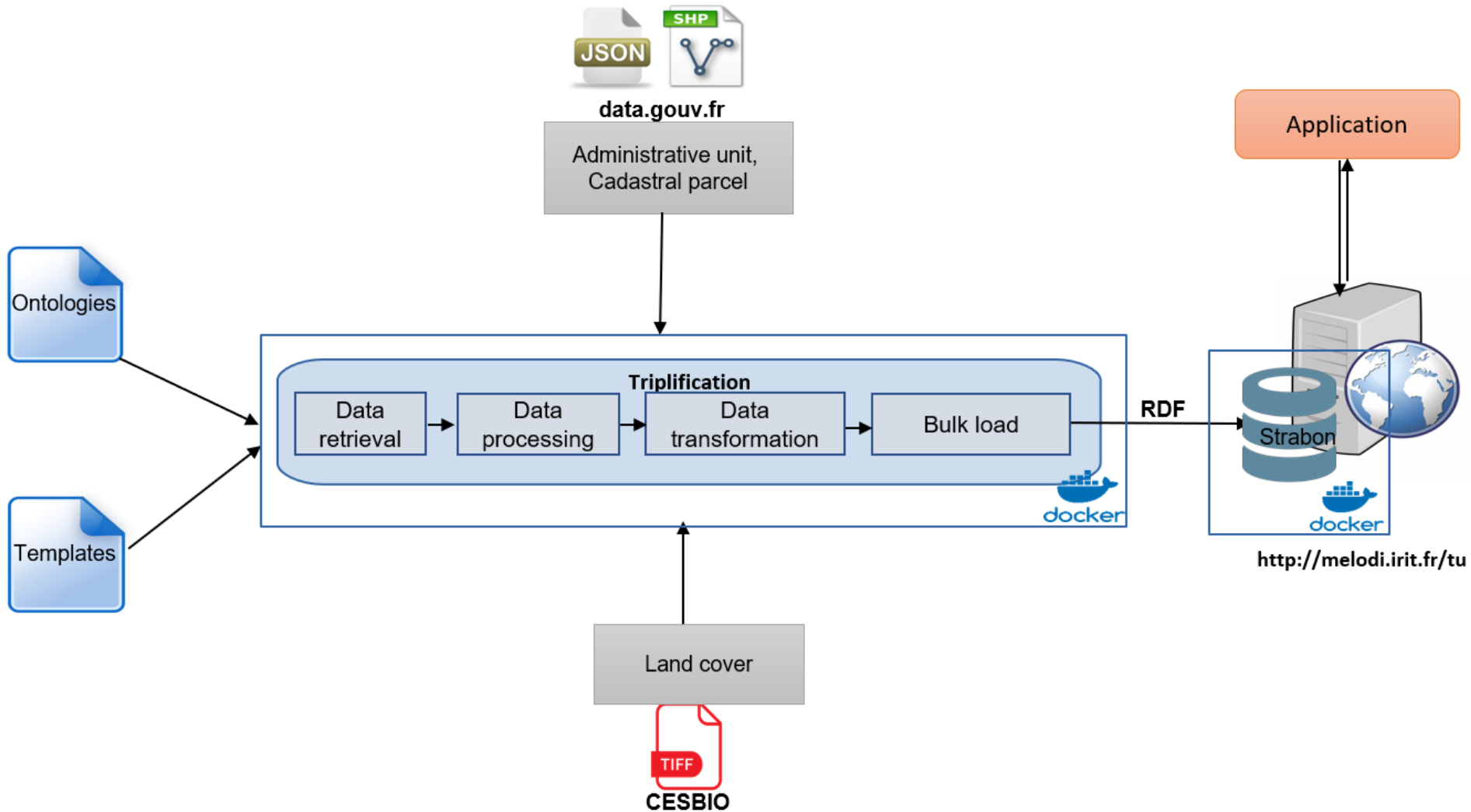


A modular ontology to represent French administrative units, land registry parcel, their dominant land cover, and their history.

2.2 Data integration

- Two approaches:
 - Mediator (on-demand mapping)
 - Data materialization
- Materialization approach is more appropriated:
 - It isn't easy to accomplish an on-demand mapping due to the format of the original sources.
 - The stable performances of the approach in querying and reasoning tasks.
 - Can perform data enrichment
 - Federated queries that perform spatial joins spanning different geospatial endpoints are not yet supported in any federated system.

2.3 System architecture



The system architecture based on docker technology

3. Application

Non sécurisé | melodi.irit.fr/tu/Query

Discovery Queries

- Find all triples in the dataset.
- Select all distinct subjects that appear in the dataset.
- Select all distinct predicates that appear in the dataset.
- Select all distinct objets that appear in the dataset.
- Find all distinct classes of the dataset.
- Find the number of triples that appear in the dataset.
- Present the first ten triples of the dataset.

Explore/Modify operations

You must be logged in to perform update queries, or run in localhost.

stSPARQL Query:

```
prefix rdfs: <http://www.w3.org/2000/01/rdf-schema#>
prefix time: <http://www.w3.org/2006/time#>
prefix tsn: <http://purl.org/net/tsn#>
prefix tu: <http://melodi.irit.fr/ontologies/tu.owl#>
prefix xml: <http://www.w3.org/XML/1998/namespace>
prefix xsd: <http://www.w3.org/2001/XMLSchema#>
PREFIX geof: <http://www.opengis.net/def/function/geosparql/>
Select ?uv ?name
where
{
  ?uv a tu:AdminUnitVersion.
  ?uv geo:hasGeometry ?geo.
  ?geo geo:asWKT ?wkt.
  filter(geof:sfContains("POLYGON((1.24 43.78,1.67 43.78,1.67 43.45,1.24 43.45,1.24 43.78))"^^geo:wktLiteral, ?wkt))
  ?uv tu:name ?name.
}
```

Output Format: HTML

View Result: Plain

Map Bounds:

Query Update

uv	name
http://melodi.irit.fr/resource/UnitVersion/V2017_L3_31230	"Gratentour"^^<http://www.w3.org/2001/XMLSchema#string>
http://melodi.irit.fr/resource/UnitVersion/V2016_L3_31230	"Gratentour"^^<http://www.w3.org/2001/XMLSchema#string>
http://melodi.irit.fr/resource/UnitVersion/V2017_L3_31186	"Fonbeauzard"^^<http://www.w3.org/2001/XMLSchema#string>
http://melodi.irit.fr/resource/UnitVersion/V2016_L3_31186	"Fonbeauzard"^^<http://www.w3.org/2001/XMLSchema#string>
http://melodi.irit.fr/resource/UnitVersion/V2017_L3_31003	"Aigrefeuille"^^<http://www.w3.org/2001/XMLSchema#string>

The SPARQL endpoint hosted on IRIT server

3. Application

When

From 01/04/2017 To 01/05/2017

Where

Search by village



Parcel: 33393000YE0037(LC: tubercules-racines)

Export CSV

Cadastral

contenance	number	landcover	village	area_ha
84956.0	37	tubercules-racines	http://melodi.irit.fr/resource/Commune/33393	8.478958129882812

Export CSV

Administrative unit

village	name	population	area	insee
http://melodi.irit.fr/resource/Commune/33393	Saint-Denis-de-Pile	5302	28.27	http://id.insee.fr/geo/commune/33393
http://melodi.irit.fr/resource/Commune/33393	Saint-Denis-de-Pile	5302	28.27	http://sws.geonames.org/6432384
http://melodi.irit.fr/resource/Commune/33393	Saint-Denis-de-Pile	5302	28.27	https://fr.wikipedia.org/wiki/3ASaint-Denis-de-Pile
http://melodi.irit.fr/resource/Commune/33393	Saint-Denis-de-Pile	5302	28.27	https://www.wikidata.org/wiki/Q720555

A part of the Semantic search interface that retrieves data from the endpoint

4. Conclusion

- An ontology-based approach that aims at integrating various spatiotemporal data source.
- The approach were used first to integrate three French datasets.
- Future work:
 - Apply the approach on other available datasets, such as the land use, vegetation index or change detection.
 - Perform inferences on the constructed knowledge-base.

References and useful links

- Jordane Dorne, Nathalie Aussenac-Gilles, Catherine Comparot, Romain Hugues, Jean-Guy Planès, et al.. Une approche sémantique pour représenter l'indice de végétation d'images Sentinel-2 et son évolution. Spatial Analytics and GEOmatics (SAGEO 2018), Nov 2018, Montpellier, France. pp.49-54
- Inglada, Jordi & Vincent, Arthur & Arias, Marcela & Tardy, Benjamin & Morin, David & Rodes, Isabel. (2017). Operational High Resolution Land Cover Map Production at the Country Scale Using Satellite Image Time Series. Remote Sensing. 9. 95. 10.3390/rs9010095.
- Project deliverables: <http://www.candela-h2020.eu/deliverables>
- Semantic search interface : <http://platform.candela-h2020.eu/semsearch/> and <http://melodi.irit.fr/semantic-search/>

References and useful links

- Jordane Dorne, Nathalie Aussenac-Gilles, Catherine Comparot, Romain Hugues, Jean-Guy Planès, et al.. Une approche sémantique pour représenter l'indice de végétation d'images Sentinel-2 et son évolution. Spatial Analytics and GEOmatics (SAGEO 2018), Nov 2018, Montpellier, France. pp.49-54
- Inglada, Jordi & Vincent, Arthur & Arias, Marcela & Tardy, Benjamin & Morin, David & Rodes, Isabel. (2017). Operational High Resolution Land Cover Map Production at the Country Scale Using Satellite Image Time Series. Remote Sensing. 9. 95. 10.3390/rs9010095.
- Project deliverables: <http://www.candela-h2020.eu/deliverables>
- Semantic search interface : <http://platform.candela-h2020.eu/semsearch/>
and <http://melodi.irit.fr/semantic-search/>