Cartography M.Sc.

Semantic-driven Geospatial Data Visualization Approach to Agriculture
Use case: Apple-growing in South Tyrol, Italy

Darya Lapo

27.09.2022

Outline

- Introduction and Motivation
- Research Objective
- Theoretical Background
- Methodology
- Case Study
- Results
- Limitations and Outlook
- Conclusion

Introduction and Motivation

2 ZERO HUNGER

To end hunger, achieve food security and improved nutrition, and promote sustainable agriculture

SUSTAINABLE DEVELOPMENT GOALS

Introduction and Motivation

Research Objective

The main goal is to design a **semantic-driven geospatial data integration and visualization approach** for the needs of the Agri-Food domain, with a particular focus on apple growing in South Tyrol, Italy.

Research Objective

RO 1

To review the current requirements and methods of semantic integration of geospatial data as well as the visualization of domain knowledge using a semantic-driven approach.

RO 2

To propose a semantic-driven geospatial data visualization approach to agriculture, particularly in the apple-growing domain.

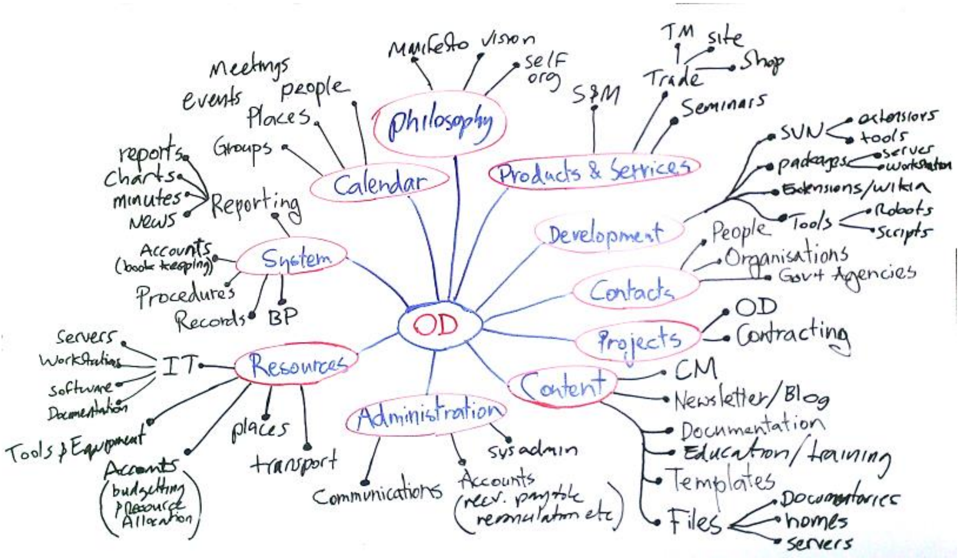
RO 3

To implement and explore the effectiveness of the developed semantic-driven geospatial data integration and visualization framework for the use cases of apple growing in South Tyrol, Italy.

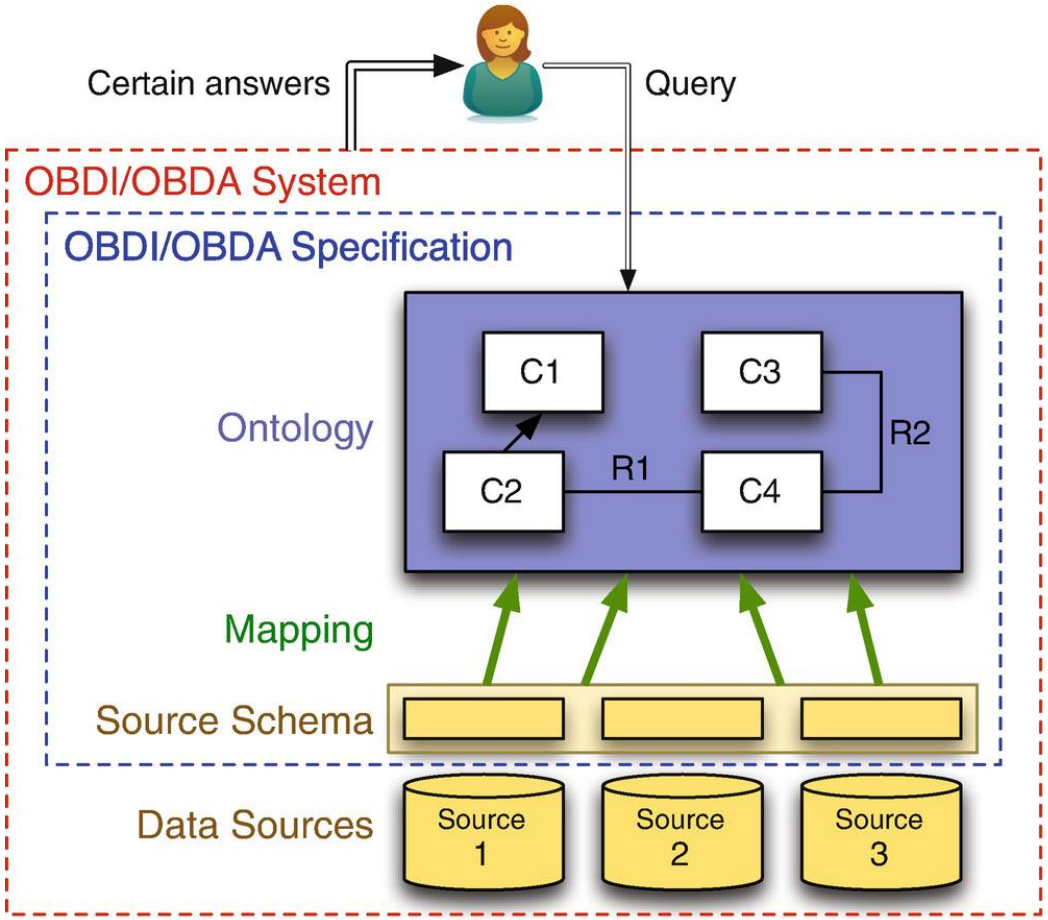
Theoretical Background

Ontology is a data model which describes the sorts of objects, properties of objects, and relations between objects that are possible in a specified domain of knowledge.

(Chandrasekaran et al., 1999)



Theoretical Background



OBDI/OBDA specification and system as depicted by Calvanese et al. (2018)

- **Data layer**

- **Ontology** gives a formal, orderly, and high level representation of the domain of interest. It is the part of the information system that clients (both people and computer programs) interact with.

- **Mapping** is a description of the relationship between the data sources and the ontology.

Theoretical Background

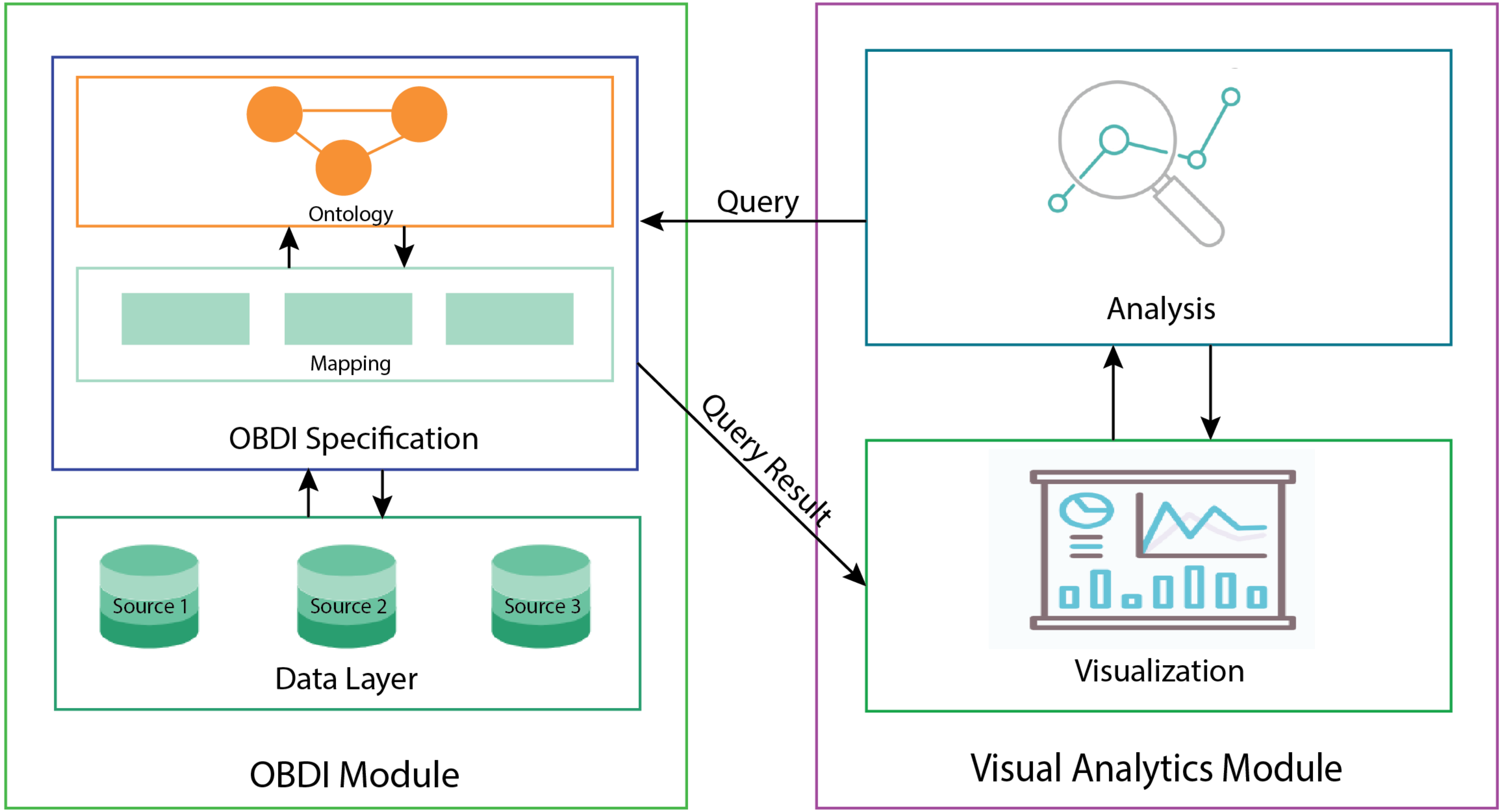
eurac research

Materialization vs Virtualization

Materialization = Extract-Transform-Load (ETL):
extract data from heterogeneous data sources -> transform -> integrate -> and materialize it to a certain target (triple store, a file, DB);

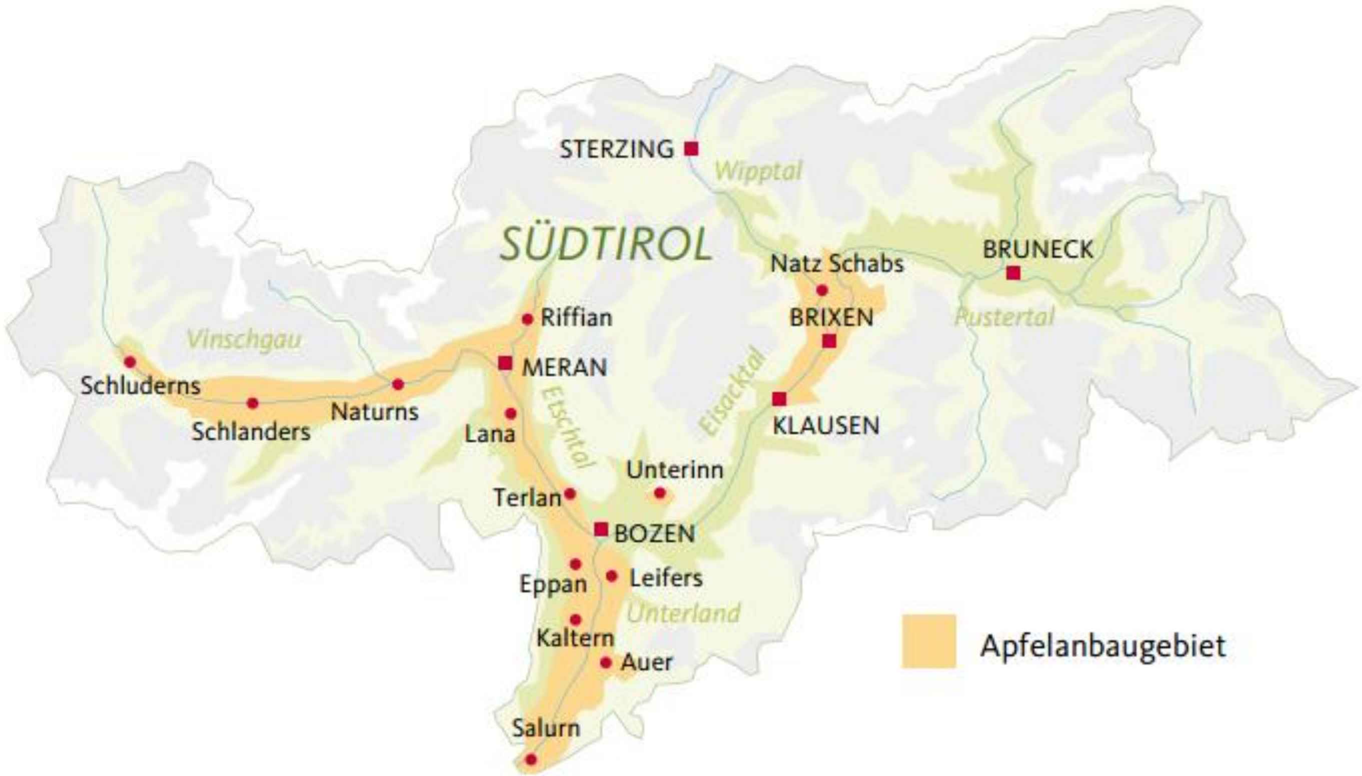
Virtualization = Ontology Based Data Access (OBDA):
Provides access to a **virtual** knowledge graph on top of the heterogeneous data. Only the data necessary to answer the query is used and transformed by the producer.

Methodology



Case Study

Study Area | Data Collection and Processing | Ontology-based Data Integration | OBDI-enabled Visual Analytics | Evaluation



Apple-growing regions of South Tyrol
(Thuile, 2022)

50% of the Italian apple market

15% of the European apple market

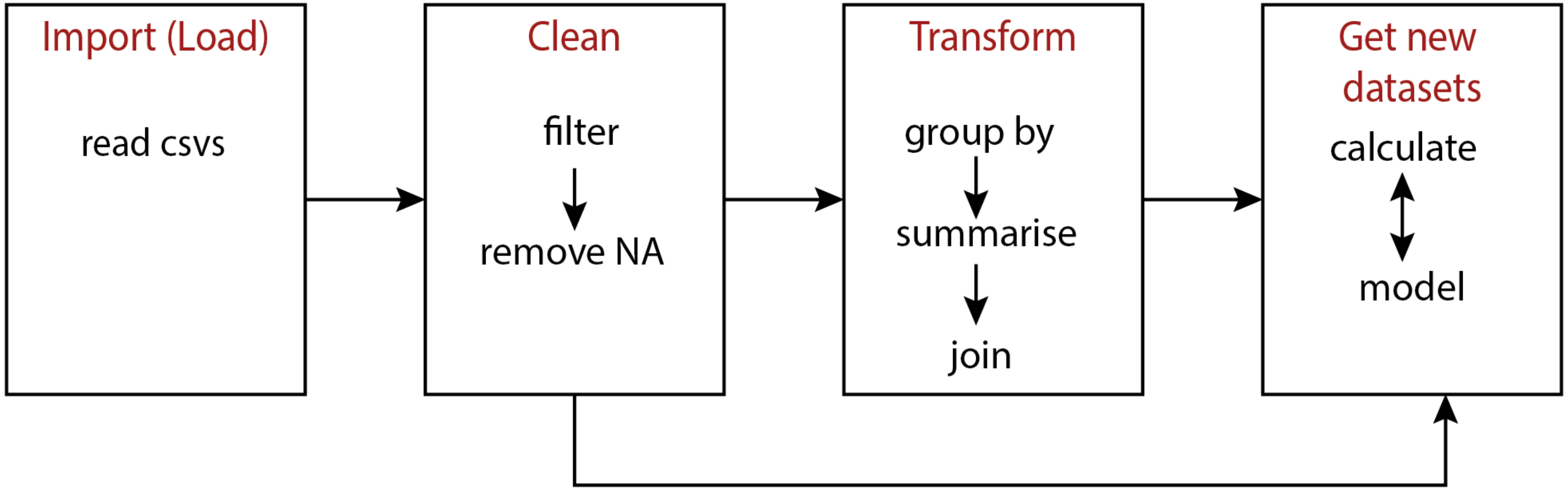
2% of the global apple market

Case Study

Dataset	Source	Format	Spatial?
Weather stations information	EDP	csv	yes
Daily meteorological records	EDP	csv	no
Monthly climatologies	EDP	csv	no
Solar Irradiation	EDP	Geotiff	yes
Land Cover	EDP	shp	yes
NDVI	EDP	API, Geotiff	yes
EU-DEM	CLMS	Geotiff	yes
Apple production in South Tyrol	ASTAT	csv	no
Bloom and harvest start dates	Laimburg Research Center	csv	no

Case Study

Study Area | **Data Collection and Processing** | Ontology-based Data Integration | OBDI-enabled Visual Analytics | Evaluation



R ‘tidyverse’
R ‘pollen’

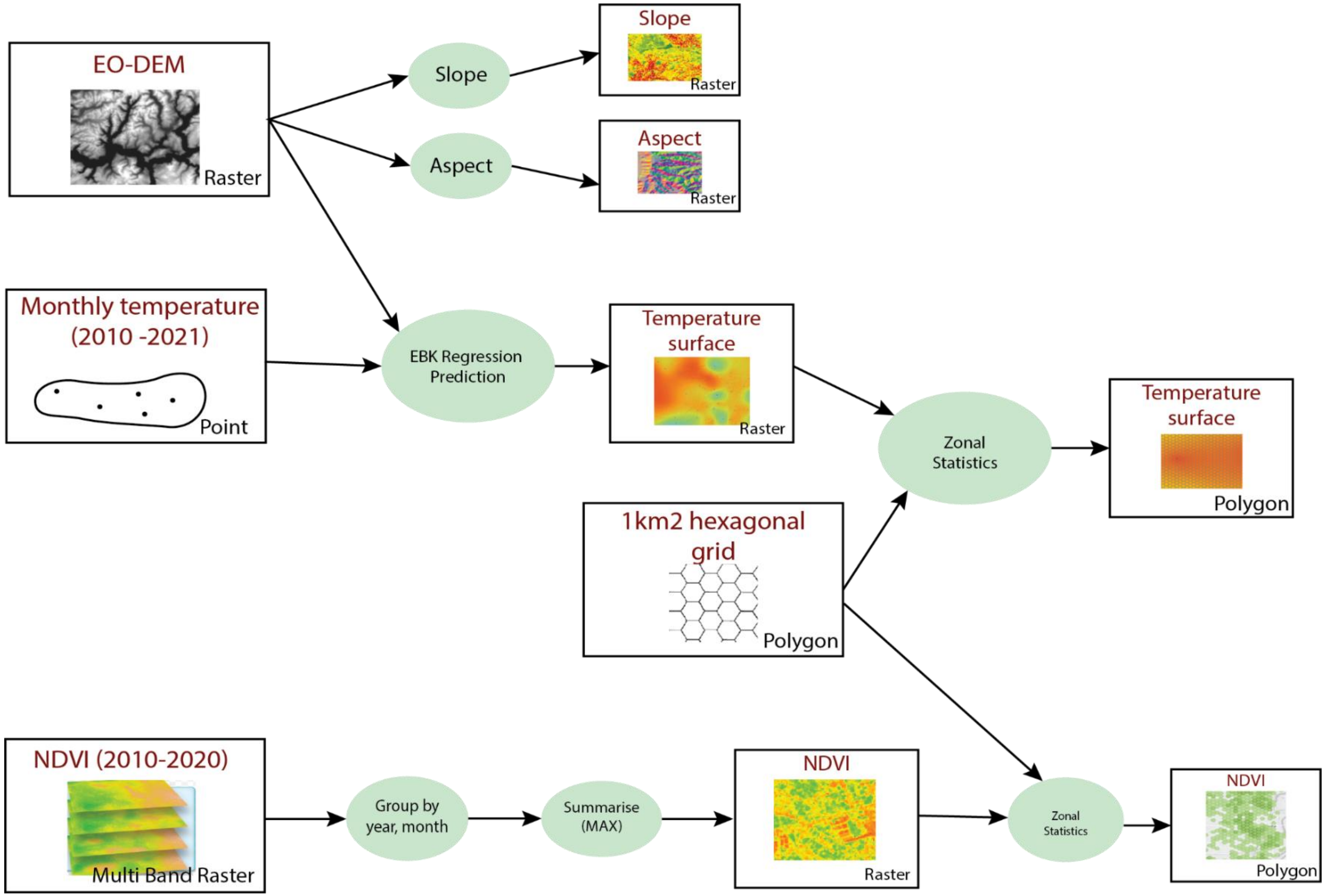
- GDD Calculation

$$GDD = \int (T(t) - T_{\text{base}})dt$$

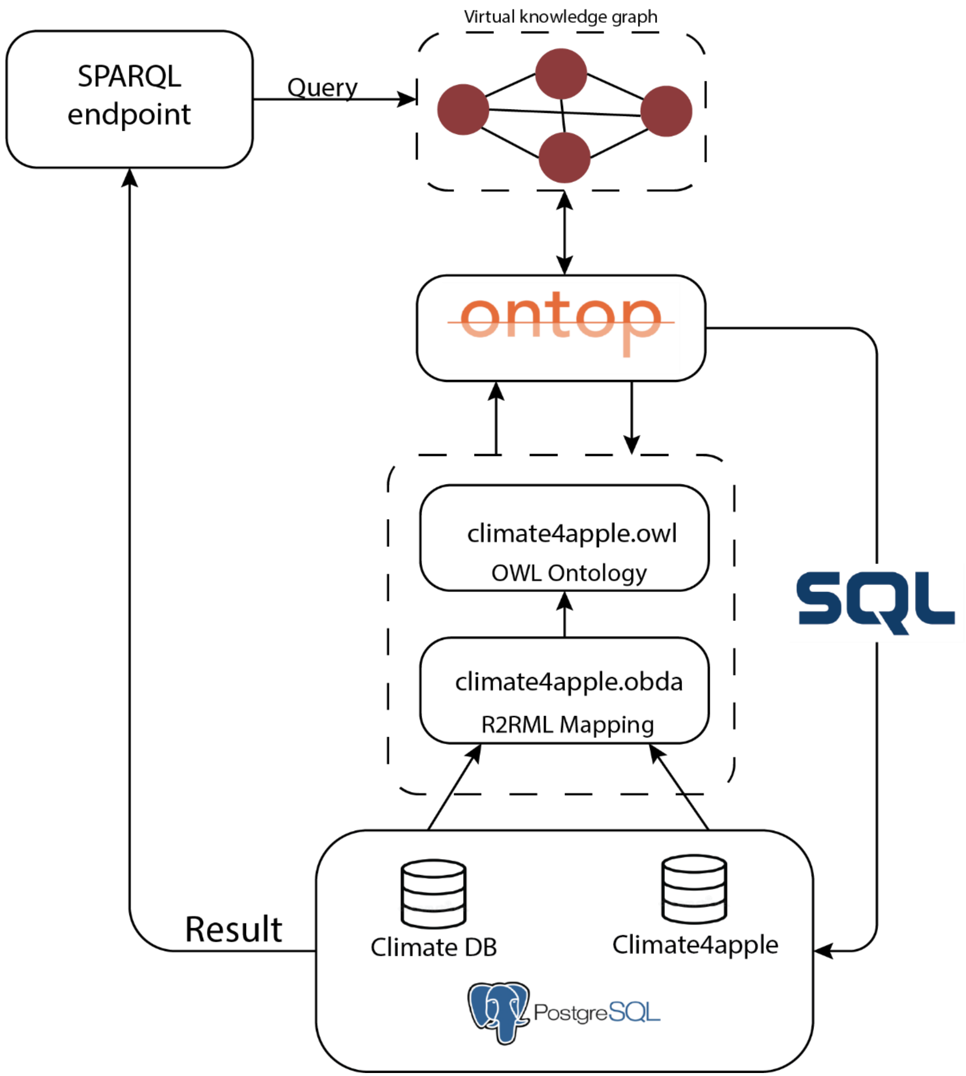
- Bloom/harvest start
- Chill units

Case Study

Study Area | **Data Collection and Processing** | Ontology-based Data Integration | OBDI-enabled Visual Analytics | Evaluation



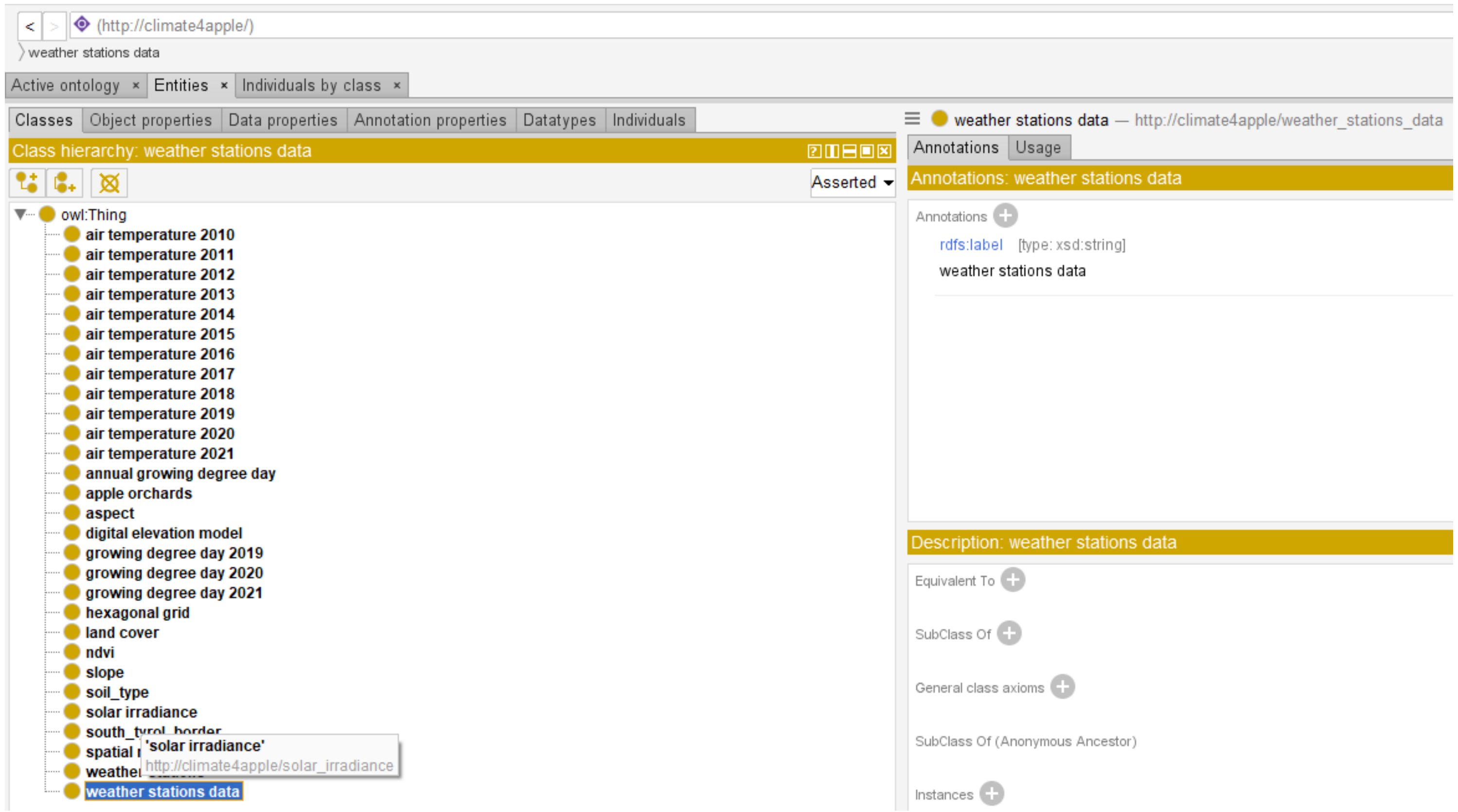
Case Study



To accomplish **interoperability**, the OBDI module depends on standard formats, including R2RML for mapping, OWL for ontology, RDF for the virtual graph, and SPARQL for queries.

Case Study


```
{
  "tableName" : "stations_daily",
  "tableAlias" : "weather_stations_data",
  "tableLabels" : ["weather stations data"],
  "tableSchema" : "public",
  "attAliases" : [
    {
      "attName" : "date",
      "attAlias" : "weather_observation_date",
      "attLabels" : ["weather observation date"]
    },
    {
      "attName" : "tmin",
      "attAlias" : "min_air_temperature",
      "attLabels" : ["minimum air temperature "]
    },
    {
      "attName" : "tmax",
      "attAlias" : "max_air_temperature",
      "attLabels" : ["maximum_air temperature"]
    },
    {
      "attName" : "tmean",
      "attAlias" : "mean_air_temperature",
      "attLabels" : ["mean air temperature"]
    },
    {
      "attName" : "prec",
      "attAlias" : "precipitation",
      "attLabels" : ["precipitation"]
```

Designing ontologies and mappings may be seen as a process of documenting/annotating the data source.

Case Study

Ontop SPARQL endpoint

endpoint address: http://localhost:8080/sparql | ontop v4.1.0-beta-1-SNAPSHOT

Query 1 x +

```
PREFIX rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#>
PREFIX xsd: <http://www.w3.org/2001/XMLSchema#>
PREFIX rdfs: <http://www.w3.org/2000/01/rdf-schema#>
PREFIX station: <http://climete4apple/weather_stations#>
PREFIX obs: <http://climete4apple/weather_stations_data#>
PREFIX gdd2019: <http://climete4apple/growing_degree_day_2019>
SELECT  ?station ?year (COUNT(?observation) as ?NOBS)
WHERE {

  ?observation obs:weather_station_name ?station.
  ?observation obs:air_temperature_max ?tmax.
  ?observation obs:weather_observation_date ?date.
  ?observation obs:weather_observation_date ?date.
  FILTER (?tmax>35)
```

Table | Response | Pivot Table | Google Chart | Geo | </>

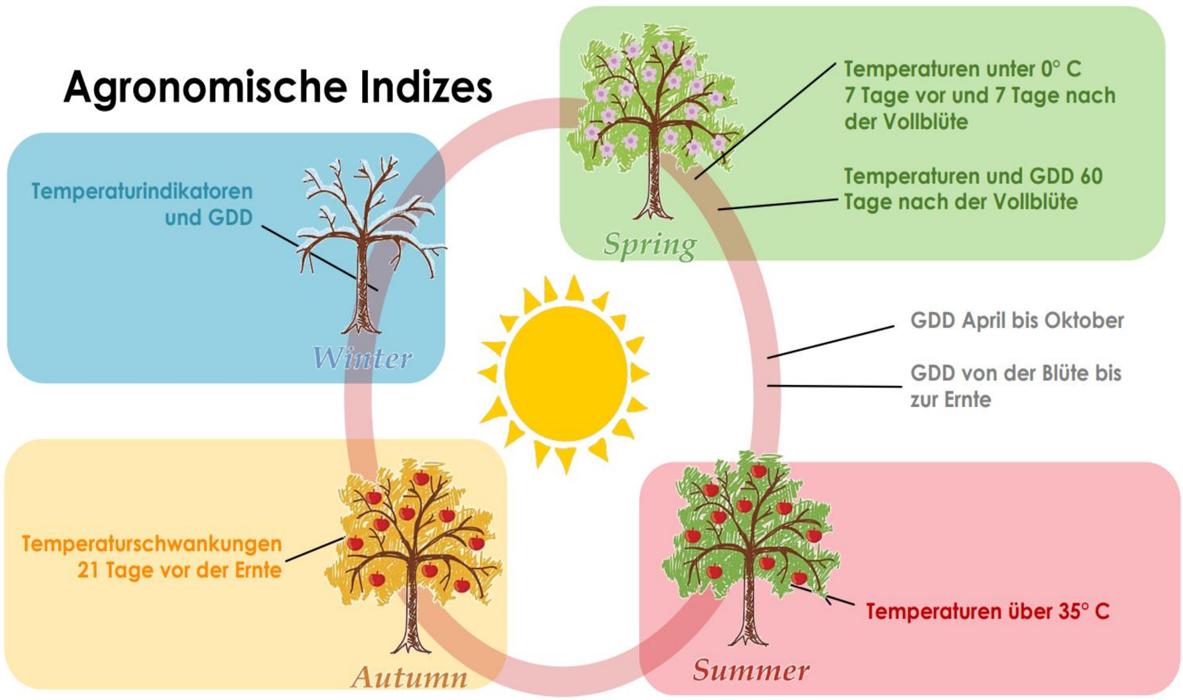
Showing 1 to 299 of 299 entries (in 0.569 seconds)

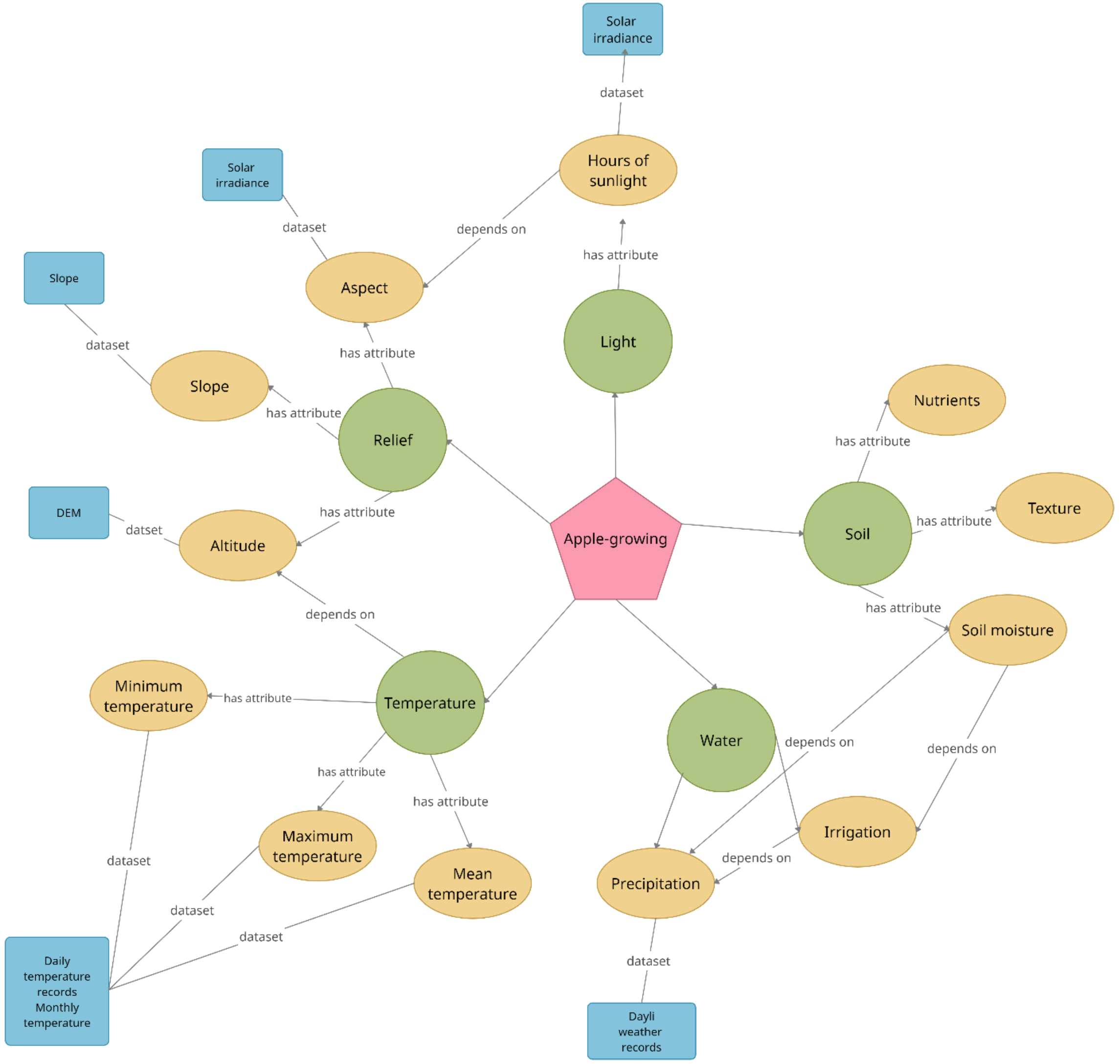
Search: Show All entries

	station	year	NOBS
1	Rovereto	"2021"^^xsd:integer	"6"^^xsd:integer
2	Rovereto	"2020"^^xsd:integer	"4"^^xsd:integer
3	Levico	"2020"^^xsd:integer	"3"^^xsd:integer
4	Aldeno	"2020"^^xsd:integer	"2"^^xsd:integer
5	Santa_Massenza	"2020"^^xsd:integer	"2"^^xsd:integer
6	Dro_Marocche	"2020"^^xsd:integer	"2"^^xsd:integer
7	Telve	"2020"^^xsd:integer	"2"^^xsd:integer
8	Trento_Laste	"2020"^^xsd:integer	"2"^^xsd:integer
9	Bozen	"2020"^^xsd:integer	"2"^^xsd:integer
10	Auer	"2020"^^xsd:integer	"2"^^xsd:integer
11	Zambana	"2020"^^xsd:integer	"2"^^xsd:integer
12	Arco	"2020"^^xsd:integer	"2"^^xsd:integer

Case Study

Study Area | Data Collection and Processing | Ontology-based Data Integration | **OBDI-enabled Visual Analytics** | Evaluation

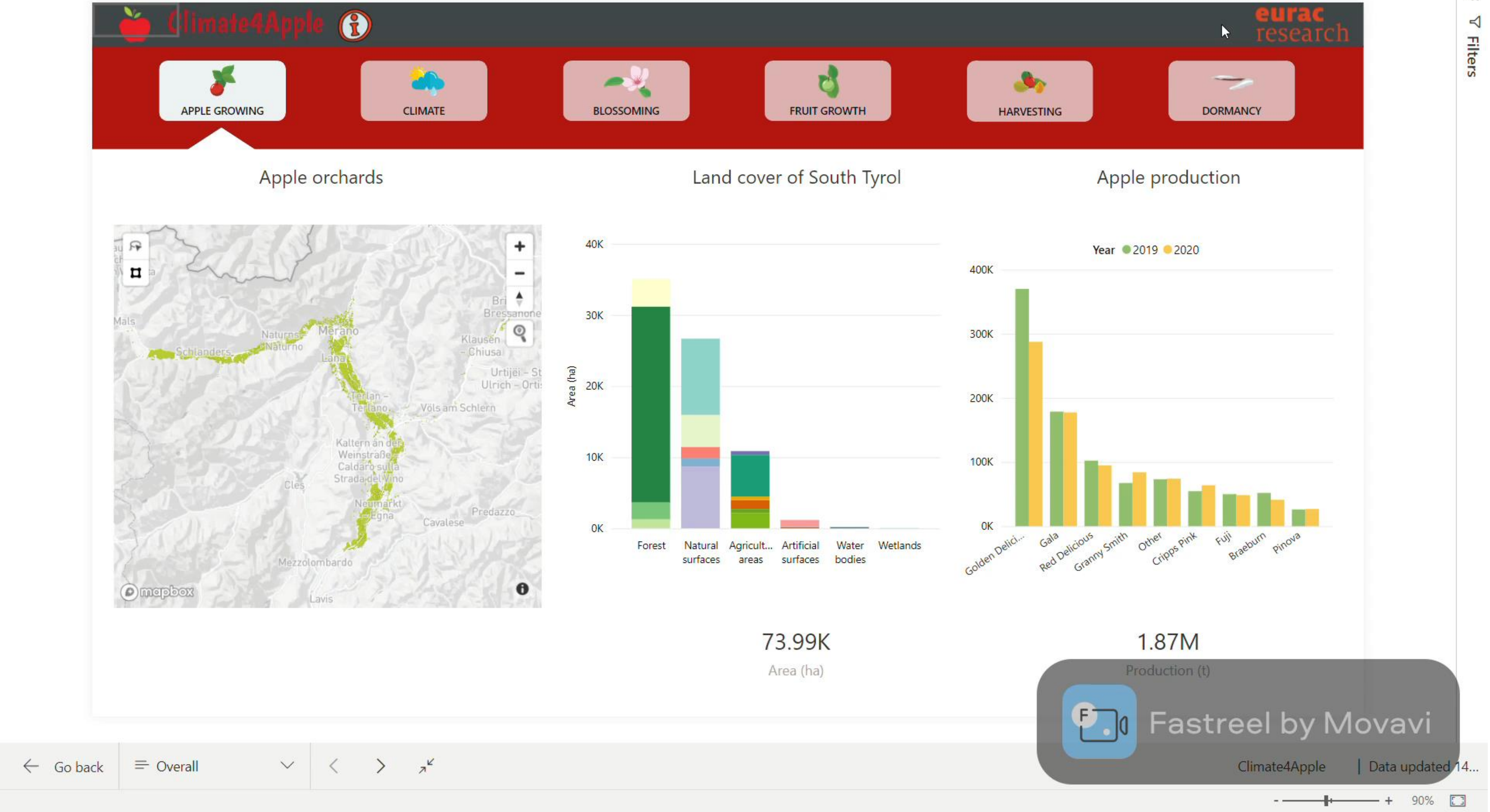




Abschlussveranstaltung Projekt „Kultivas“, Bozen, 16.03.2022

Case Study

Study Area | Data Collection and Processing | Ontology-based Data Integration | **OBDI-enabled Visual Analytics** | Evaluation



Case Study

- Talking aloud and semi-structured interviews;

- Two user groups:

Software developers

- general feedback about the interface
- bugs fixing

Domain experts in agriculture

evaluation of how

- clear
- understandable
- useful

Results

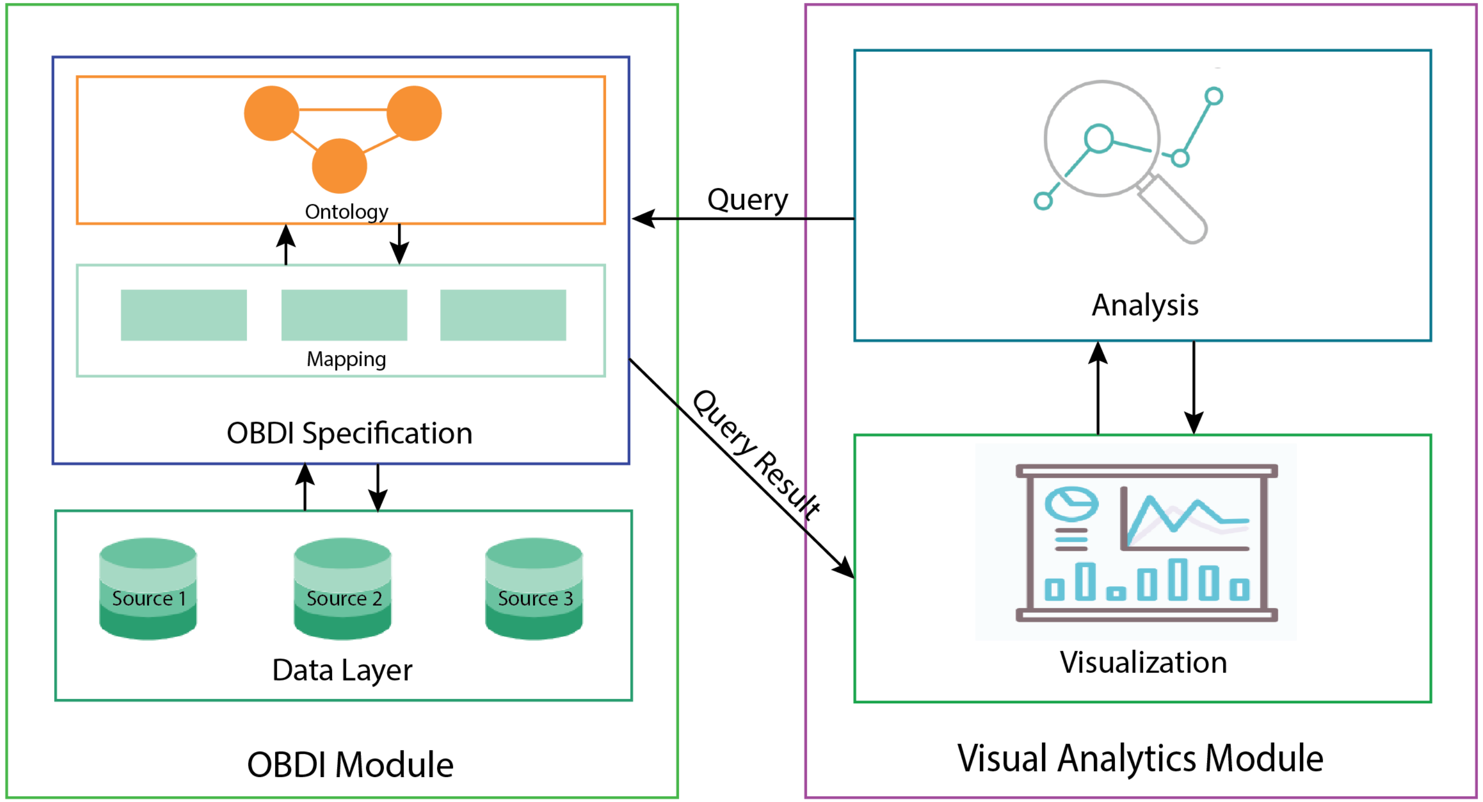
RO 1 | RO 2 | RO 3

Big	Simple	Fast	Efficient
Big geospatial data	Data access via OBDA/OBDI regardless how and where data are stored	Virtualization approach requires no ETL processes or a new custom database	Time and cost of decision-making is reduced

To review the current requirements and methods of semantic integration of geospatial data as well as the visualization of domain knowledge using a semantic-driven approach

Results

RO 1 | RO 2 | RO 3



To propose a semantic-driven geospatial data visualization approach to agriculture, particularly in the apple-growing domain

Results

RO 1 | RO 2 | RO 3

- clear
- easy to understand
- useful
- nice to present to the broader audience

- focus only on orchards
- Include more variables to analysis

To implement and explore the effectiveness of the developed semantic-driven geospatial data integration and visualization framework for the use cases of apple growing in South Tyrol, Italy

Limitations and Outlook

Limitations | Outlook

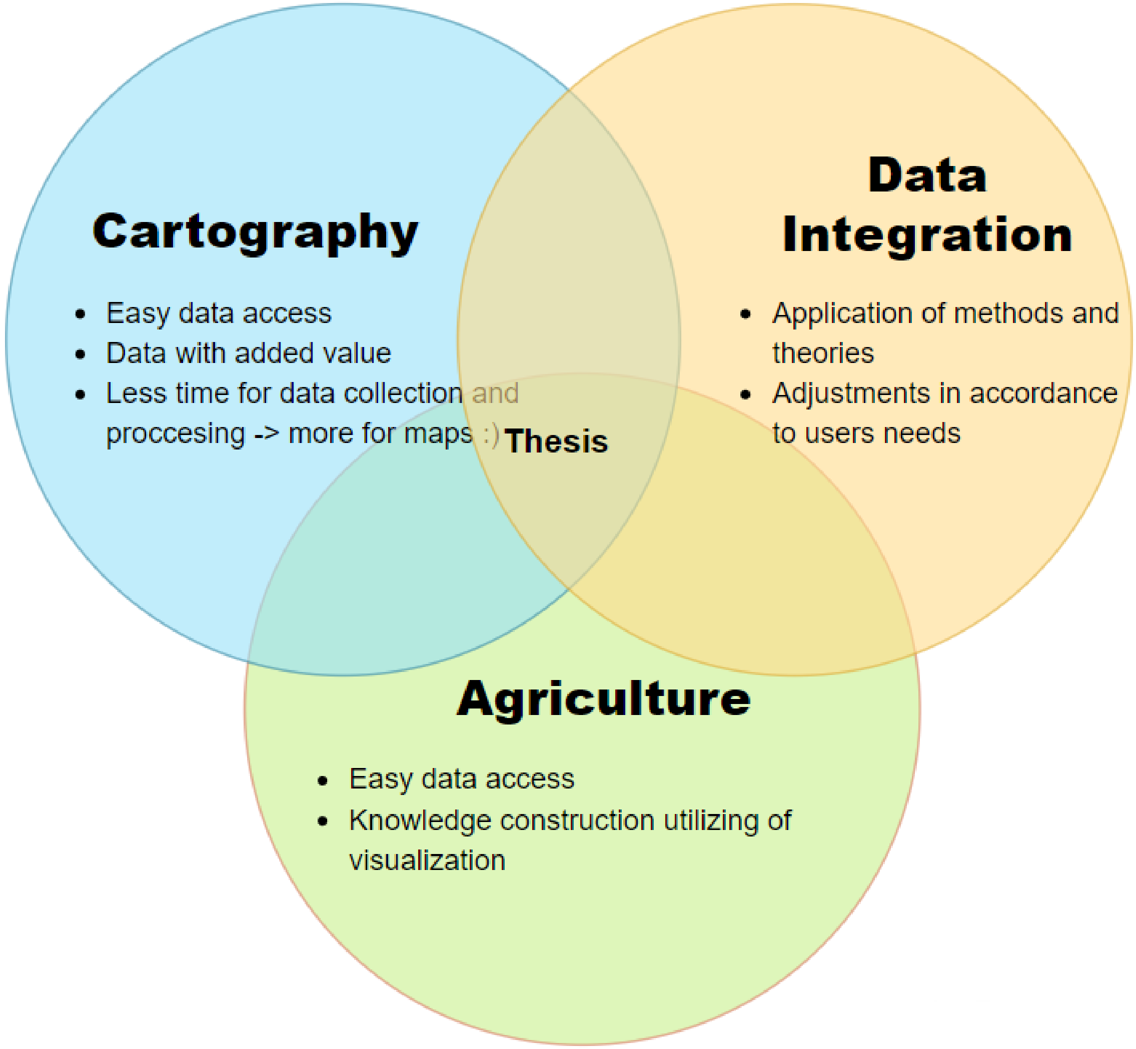
- Incomplete and inconsistent data
- Data privacy concerns
- Ontop & GeoSPARQL

Limitations and Outlook

Limitations | Outlook

- Reusing existing ontologies
- Semantic-driven geospatial data processing and analysis
- Integration and visualization of more data
- A proper user study set up

Conclusion


Cartography
Easy data access
Data with added value
Less time for data collection and proccesing -> more for maps :)
Thesis
Data Integration
Application of methods and theories
Adjustments in accordance to users needs
Agriculture
Easy data access
Knowledge construction utilizing of visualization

Acknowledgments

This research is partially funded by EURAC Research, Center for Sensing Solutions, Cartography MSc ERASMUS+ program and supported by the European Union's Horizon 2020 research and innovation program under the Marie Skłodowska-Curie grant agreement No 894215.

Thank you for your attention.

Happy to hear your questions, ideas and suggestions. ☺ ✨

Results

RO 1 | RO 2 | RO 3

What are the latest standards, methods, and best practices for semantic integration of geospatial data?	How to formalize and visualize domain knowledge using a semantic-driven approach in cartography?	What are examples of successful implementation of semantic technologies in the agricultural domain to support effective decision-making?
• Ontologies • OBDA/OBDI • Virtualization approach ⬇ • No new DB or ETL • Easier and faster data access • More effective decision making	• Cartographic ontologies • Map as a knowledge base	• AGROVOC • FoodOn • Crop Ontology

Results

RO 1 | RO 2 | RO 3

What are the elements of the semantic-driven geospatial data integration and visualization framework?	How can geospatial data be enhanced by using semantic technologies for achieving better integration and interoperability?	Which cartographic techniques are the most suitable for visualizing environmental and agricultural variables?
• OBDA/OBDI Module • Visualization Module • Communication via SPARQL endpoint	• Bridge a semantic gap between raw data and terms used within domain • FAIR principles	• Isarithmic map • Hot spot map • Graphs and charts • Color hue, color saturation, size

Results

RO 1 | RO 2 | RO 3

Which apple-growing use cases should be implemented to illustrate the effectiveness of the proposed framework?	How can users benefit from the proposed semantic-driven geospatial data integration and visualization framework?
• Temperature as the main driver of the environmental processes	Very positive feedback Useful Clear