Semantic Visual Query Answering on Heterogeneous Territorial Data

Loris Bozzato Giulia Bruscagin Gaetano Calabrese
Luciano Serafini Cesare Furlanello

Fondazione Bruno Kessler
Email: <surname>@fbk.eu

12th International Conference on Information, Process, and Knowledge Management (eKNOW 2020)
November 21-25, 2020 – Valencia, Spain

Presenters

Presenter

Loris Bozzato bozzato[at]fbk.eu
Researcher @ DKM - Data and Knowledge Management Research Unit
FBK - Fondazione Bruno Kessler, Trento, Italy

Authors

- DKM - Data and Knowledge Management @ FBK
  Loris Bozzato, Gaetano Calabrese, Luciano Serafini
- MPBA - Predictive Models for Biomedicine & Environment @ FBK
  Giulia Bruscagin, Cesare Furlanello

Project

This work reports some of the outcomes of the Geo@Reporter project, developed as a collaboration with GeoPartner (Trento, Italy)
https://www.geopartner.it/servizi/georeporter/
https://dkm.fbk.eu/projects/georeporter

Objectives and contributions

- **Challenges:** local municipalities need to integrate information coming from different sources

Objectives and contributions

- **Challenges:** local municipalities need to integrate information coming from different sources

- **Contributions:**
  1. semantic integration of real estate and associated tributes data
  2. support for semantic visual analytic, and query formulation and geo-visualization

Objectives and contributions

- **Challenges:** local municipalities need to integrate information coming from different sources

- **Contributions:**
  1. semantic integration of real estate and associated tributes data
  2. support for semantic visual analytic, and query formulation and geo-visualization
- **Delivered assets:**
  1. **Ontology** for the domain or real estate and taxation
  2. Scripts for **ontology population** with data from cadastre, municipal registry, and household utilities
  3. System for the **visual query answering** over the integrated knowledge base

Ontology: Structure and Development

Ontology specification

Aim

To represent the key objects types of the domain, i.e., buildings, their cadastral subdivisions together with their owners and their associated features, the activated services, and the taxation relations.

Specification language

The ontology is specified in Web Ontology Language (OWL 2), which allows us to take advantage of the available tools for modelling and reasoning in the standard ontology language.

Ontology domain

Cadastral data:

Cadastral data contains the principal objects of the domain. They represent the central objects of the model. Principal objects from this data source are land and real estate units, owner subjects and their onwership relation with cadastral units.

Ontology domain

Cadastral data:

Cadastral data contains the principal objects of the domain. They represent the central objects of the model. Principal objects from this data source are land and real estate units, owner subjects and their onwership relation with cadastral units.

Civil registry data:

Information about family units from the municipalities' registry data. Main objects of this sources are resident individuals and family units.

Ontology domain

Cadastral data:

Cadastral data contains the principal objects of the domain. They represent the central objects of the model. Principal objects from this data source are land and real estate units, owner subjects and their onwership relation with cadastral units.

Civil registry data:

Information about family units from the municipalities' registry data. Main objects of this sources are resident individuals and family units.

Real estate tributes:

Information from the municipality about real estate tributes for the specific year. The main objects are municipalities taxes to be linked to local real estate units.

Ontology domain

Cadastral data:

Cadastral data contains the principal objects of the domain. They represent the central objects of the model. Principal objects from this data source are land and real estate units, owner subjects and their onwership relation with cadastral units.

Civil registry data:

Information about family units from the municipalities' registry data. Main objects of this sources are resident individuals and family units.

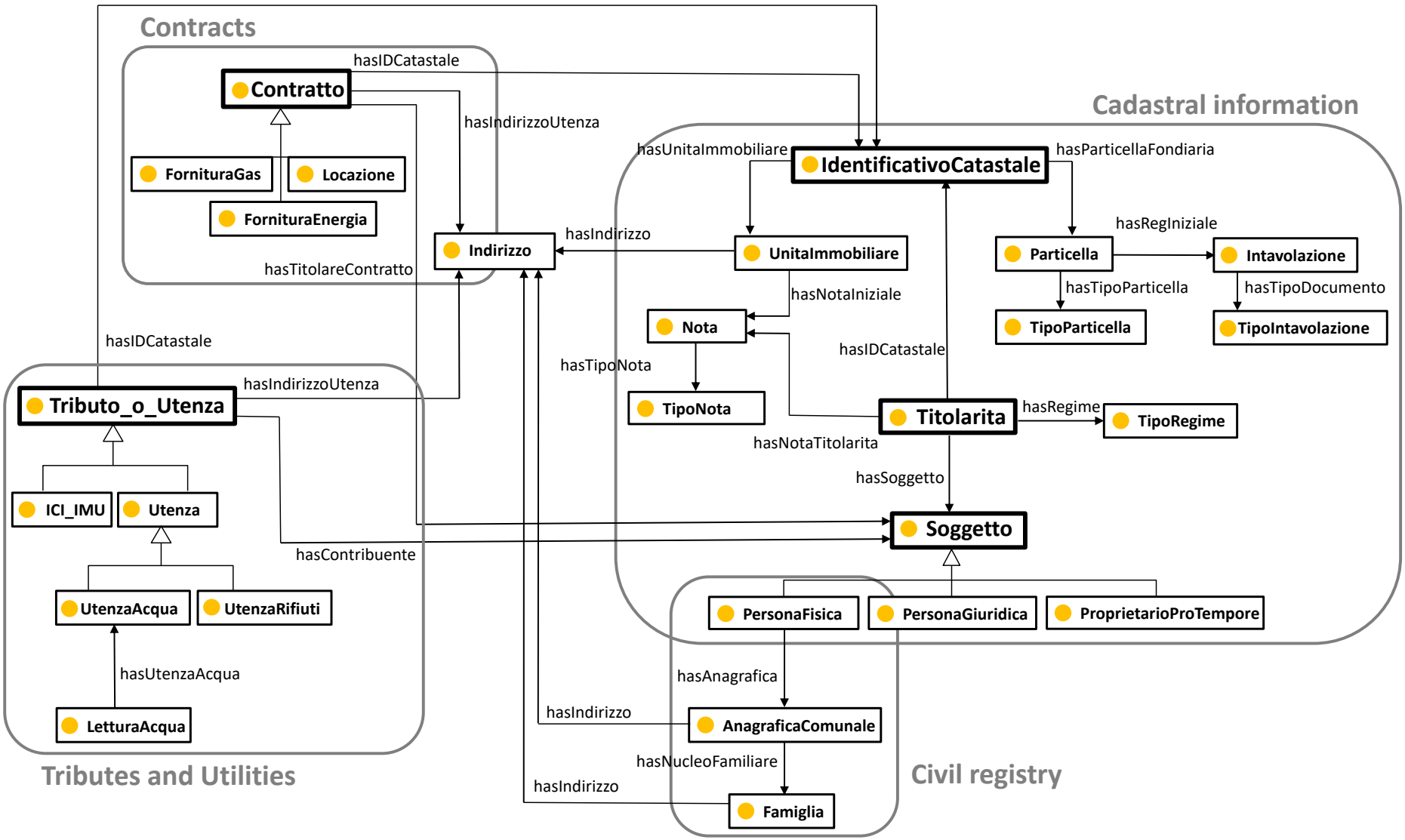
Real estate tributes:

Information from the municipality about real estate tributes for the specific year. The main objects are municipalities taxes to be linked to local real estate units.

Utility and rent contracts:

Data about rents and utility contracts for waste collection, gas, electricity and water were provided by the municipality. The main objects of interest are contracts (and their related information) for the different utilities linked to specific real estate units.

General schema of the Geo@Reporter ontology



Ontology population with real data

Data sources

- Cadastral data: provided in a proprietary format in form of pipe-separated text files (together with documentation for interpretation of fields).
- Civil registry data: provided as comma-separated values (CSV) format files for individuals and families units.
- Real estate tributes: provided as a spreadsheet export of the administrative software of the municipality.
- Utility and rent contracts: extracted as spreadsheets from the Sistema Interscambio Anagrafe Tributarie Enti Locali (SIATEL) system for exchange of tributary data in the Italian public administration.

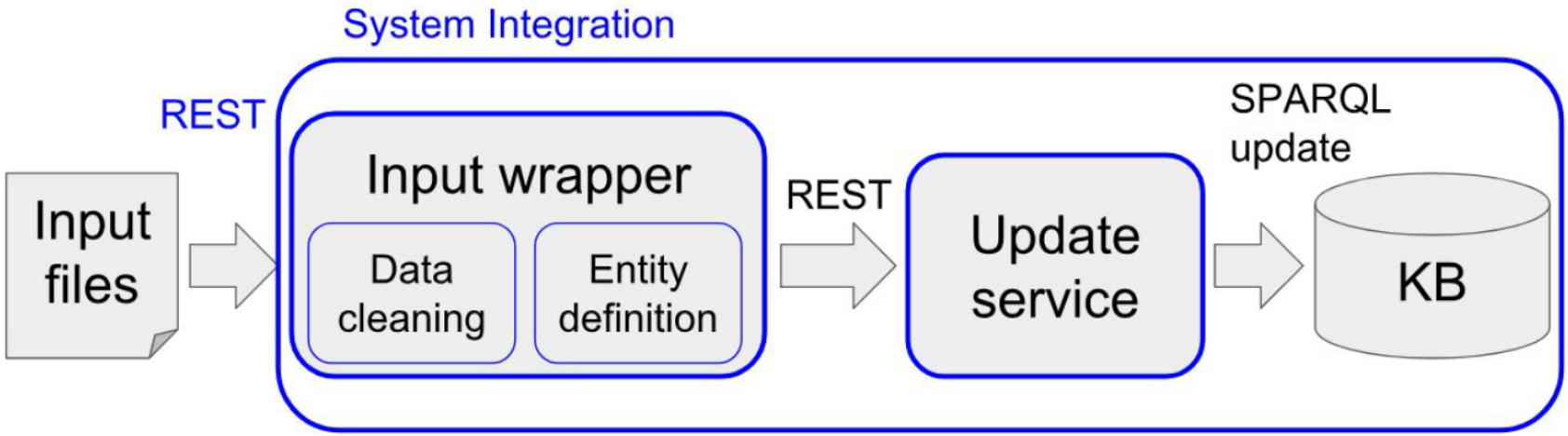
Data sources

- Cadastral data: provided in a proprietary format in form of pipe-separated text files (together with documentation for interpretation of fields).
- Civil registry data: provided as comma-separated values (CSV) format files for individuals and families units.
- Real estate tributes: provided as a spreadsheet export of the administrative software of the municipality.
- Utility and rent contracts: extracted as spreadsheets from the Sistema Interscambio Anagrafe Tributarie Enti Locali (SIATEL) system for exchange of tributary data in the Italian public administration.

Note

Sources are heterogeneous for format and structure of the data, while also overlapping on some of the represented objects.

Ontology population system



Import process

1. Input wrapper component works on the format of the input files to apply data cleaning operations and to recognize relations across entities.
2. Update services, implemented as *Representational State Transfer (REST)* interfaces, allow to interact with the KB in order to add to the KB (or update) the information derived from input files in terms of the structure of the ontology.

Visual Query Answering System

Visual Query Answering System

Aim:

provide a Web-based instrument that can make affordable the analysis of heterogeneous data through one visual interface. The visual query answering system can guide the user to explore and combine different data sources with flexibility, consistency, and intuitivity.

Visual Query Answering System

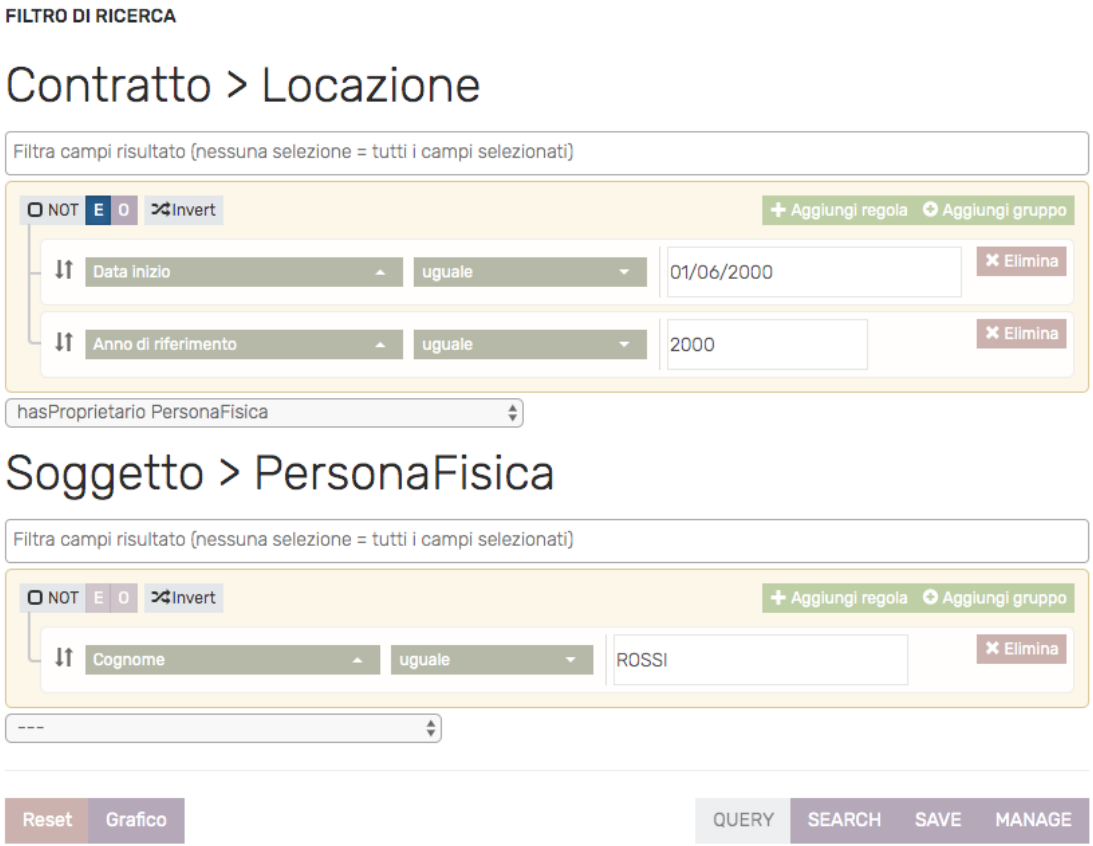
Aim:

provide a Web-based instrument that can make affordable the analysis of heterogeneous data through one visual interface. The visual query answering system can guide the user to explore and combine different data sources with flexibility, consistency, and intuitivity.

Components

- Smart Query Builder: a visual-based system for complex ontology based query building
- Logic Tree: a visual representation of the query results, highlighting the relationship between the items
- Hybrid 2D/3D Geographic Information System (GIS): a multi-level visualization that combines 2D and 3D environments to inspect geo-referenced entities

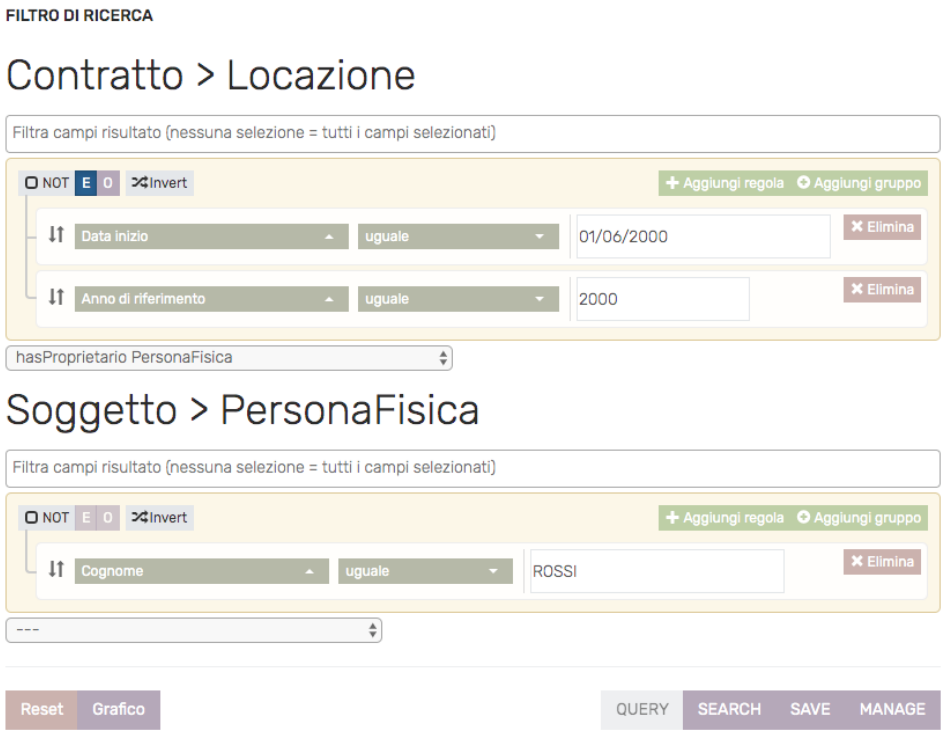
Smart Query Builder

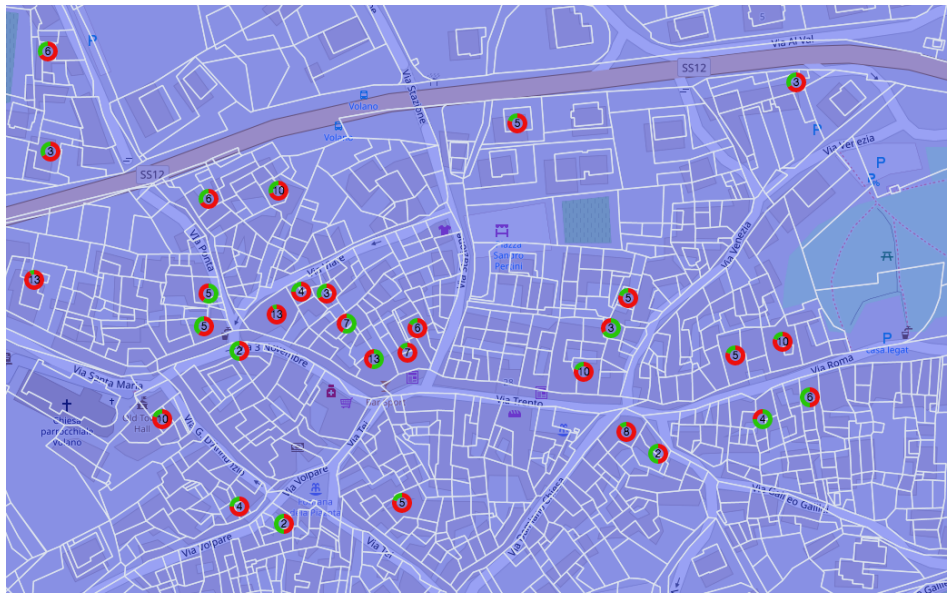


After choosing the first entity (i.e., *Contratto Locazione*, rent contract), the list of connected entities is available in a drop-down menu. The user can choose mathematical and logical operators in a dedicated drop-down menu and insert free text in the empty box to complete the query with custom parameters. As long as the Smart Query Builder finds relations with other entities, the user can keep adding new boxes, creating a more complex query.

Smart Query Builder

Example of graphical and geo-referenced representation of the results of the query:

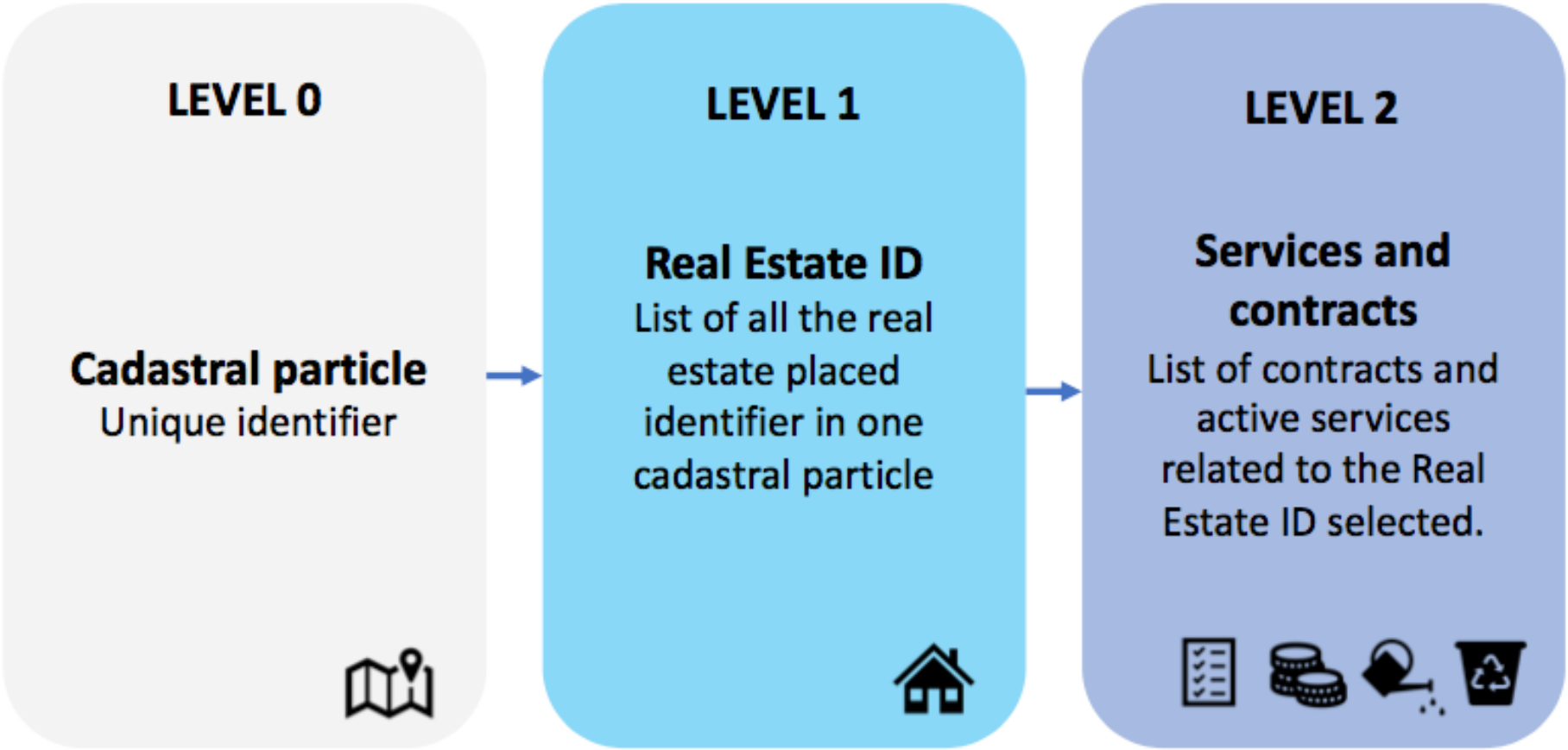




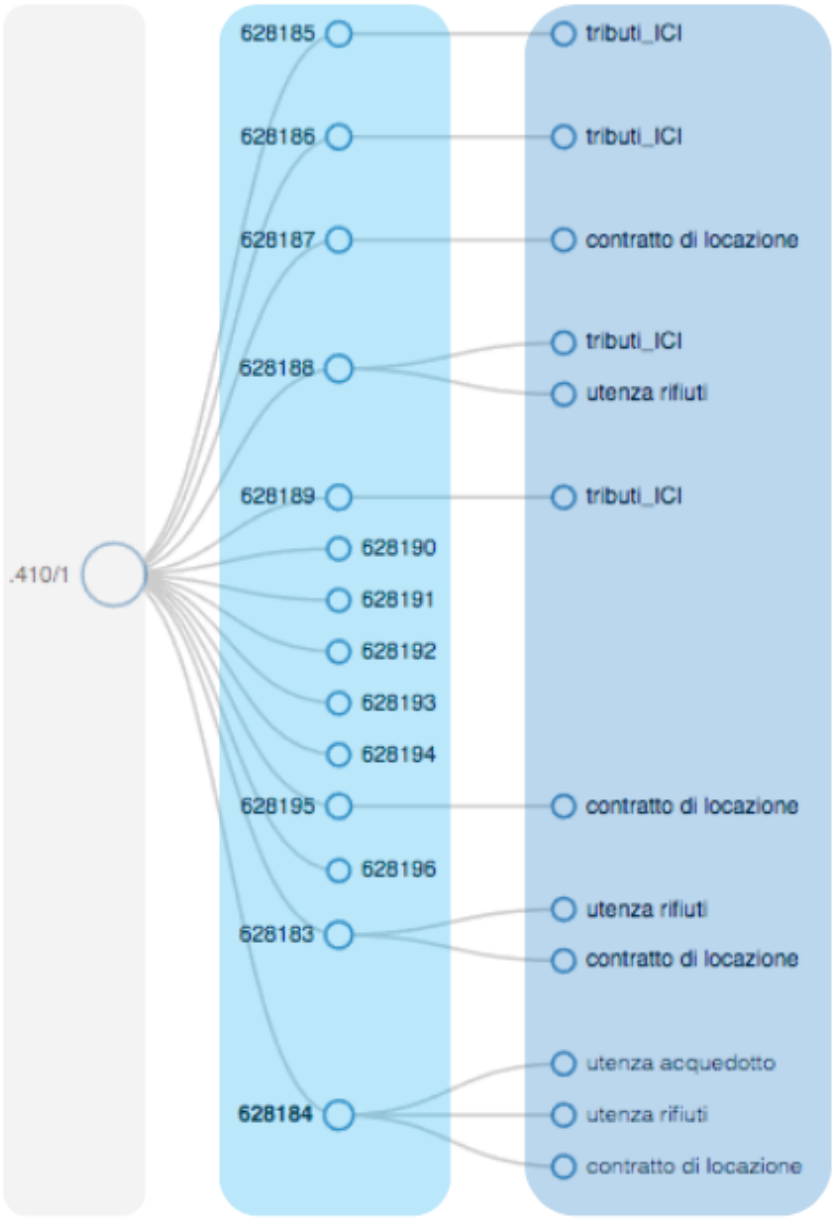
(Map data © OpenStreetMap contributors, CC BY-SA)

Logic Tree

In order to highlight the association between cadastral parcel, real estate identifier (*Unità Immobiliare*) and personal/services/tax-related data in the ontology, we defined a graphical logic structure to represent the relations between the various entities.

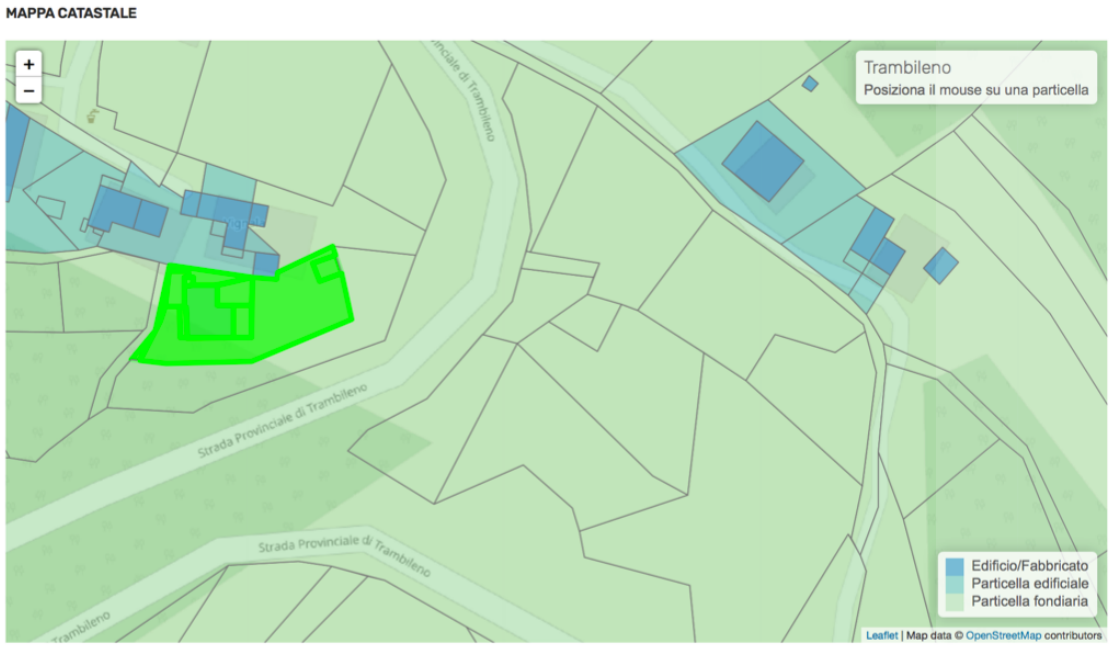


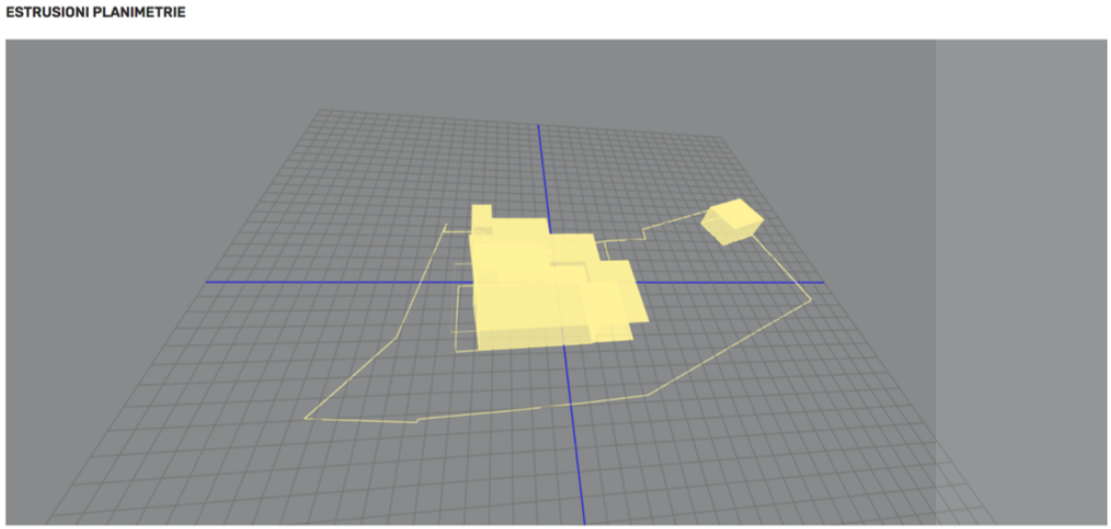
Logic Tree


.410/1
628185
628186
628187
628188
628189
628190
628191
628192
628193
628194
628195
628196
628183
628184
tributi_ICI
tributi_ICI
contratto di locazione
tributi_ICI
utenza rifiuti
tributi_ICI
contratto di locazione
utenza rifiuti
contratto di locazione
utenza acquedotto
utenza rifiuti
contratto di locazione

Hybrid 2D/3D GIS System

After a usability test (performed with the domain experts) we choose a hybrid representation to explore provided by a tool called Three.js where the point of view of the visualization is always fixed towards the center of the building and the library allows the reconstruction of those parts that are under the road surface, allowing a complete exploration of the structure.





Conclusion

Contributions of our work can be summerized in three points:

1. We first detailed the structure and modelling of the project ontology, by showing how the different data sources guided the schema definition.
2. Then, we detailed the integration system of the platform, which allows to clean and map different data formats to the semantic based representation of the project KB.
3. Finally, we presented our solutions for data visualization, providing different tools for data analysis and 2D/3D map navigation.