OnGIS: Ontology Driven Geospatial Search and Integration

Marek Šmíd, Zdeněk Kouba
{smidmare,kouba}@fel.cvut.cz

Department of Cybernetics
Faculty of Electrical Engineering
Czech Technical University in Prague

Boston, MA, 12.11.2012

Introduction

- There is a lot of geographical data.
  - Government institutions collect and maintain them – cadastral, municipal, postal, meteorological, and other maps.
  - Open international mapping projects, e.g. `http://www.openstreetmap.org/`.
- Have rigid structure.
- Suitable for semantic description with ontologies and mutual interlinking, which allows for easy querying.

OnGIS

- Provides non-expert users with simple search over complex heterogeneous geographical data.
- Still allows the queries to have non-trivial structure.
- Access to the data is mediated via semantic layer, which also makes the integration easy.
- It is possible to access data with different structure and variously technically available via OnGIS plugins.
- Uses OWL 2 QL, a profile of semantic languages OWL 2, allowing polynomial querying.
- Is a web application.

Motivation – ÚRM

- Developing cooperation with the department of urban planning of Prague (the capital of the Czech Republic), being a part of City Development Authority of Prague (ÚRM).
- Responsible for collecting many spatial data, e.g. pollution, noise, flood risks, land prices, etc.
- For a general user, looking e.g. for info where to build a house, it is not easy to find places according to his criteria.
- Metadata of its geoportal based on ArcGIS server were extracted into an ontology.

OWL 2 QL

- Supports e.g. sub-class, sub-property, domain and range axioms.
- Goes beyond RDFS expressivity.
- Possesses the open-world assumption.
- Due to its query answering being tractable, it allows performing them directly in relational databases using SQL.
- Its semantics is based on description logics $DL\!-\!Lite^{\mathcal{H}}_{core}$.

OWL 2 QL

Basic constructs:

- $B ::= A \mid \exists R$
- $C ::= B \mid \neg B$
- $R ::= P \mid P^{-}$

- A – concept name
- B – basic concept
- C – general concept
- P – role name
- R – complex role

- TBox axioms: $B \sqsubseteq C$ and $R_1 \sqsubseteq R_2$.
- ABox axioms: $A(a)$ and $P(a, b)$.
  - a, b – individuals .
- With the usual semantics.
- Extended with various features not affecting its tractability, e.g. data roles.

OnGIS Form Annotations

- Annotations, which ensure independence of the generic querying system on domain specific ontologies and data structures.

searchable annotates, what should be searched for a user's query string.

geometry annotates objects representing spatial geometries, useful for spatial queries.

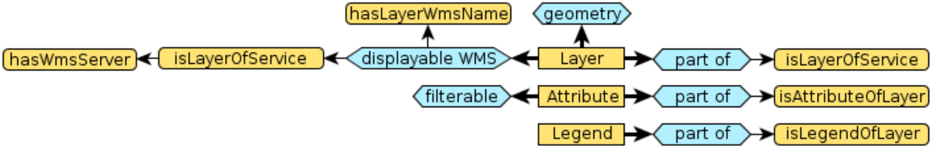
filterable annotates what could be filtered by a string (e.g. an attribute).

partof specifies part-of relations between objects (integral components).

displayable annotates what could be displayed on a map.

OnGIS Form Annotations on ÚRM Domain

- Part of OnGIS annotations (blue diamonds) on ÚRM domain terms (yellow rectangles).
- Thick arrows denote annotating.
- Thin arrows denotes annotation values (which are represented with round rectangles).
- E.g. “part of” annotations link to object properties, which relate instances and its integral parts.



Plugin for Connecting to Databases

- One of OnGIS plugins.
- Uses our OWL 2 QL reasoner (OwlgresMM, based on owlgres by Clark&Parsia), which answers semantic queries directly from relational databases.
- It is necessary to map an ontology to database tables.
  - Again done with annotations.
- Supports spatial data in PostGIS (spatial PostgreSQL extension).

OwlgresMM

- Used database schema: class (resp. object/data property) assertions in separate tables per named class (resp. property).
- Allows using multiple databases – query distribution.
- Supports some basic spatial operations:
  - spatial filters (within, within distance, bounding box),
  - geometry accessors (geometry, centroid, area), and
  - aggregation functions (count, min, max, total length and area).
- Being developed to fully support spatial query language GeoSPARQL (an OGC standard), currently only inspired by it.

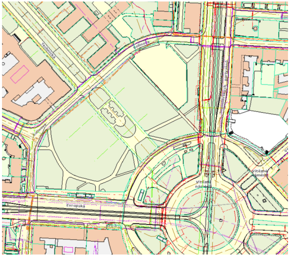
Used Data

The prototype uses:

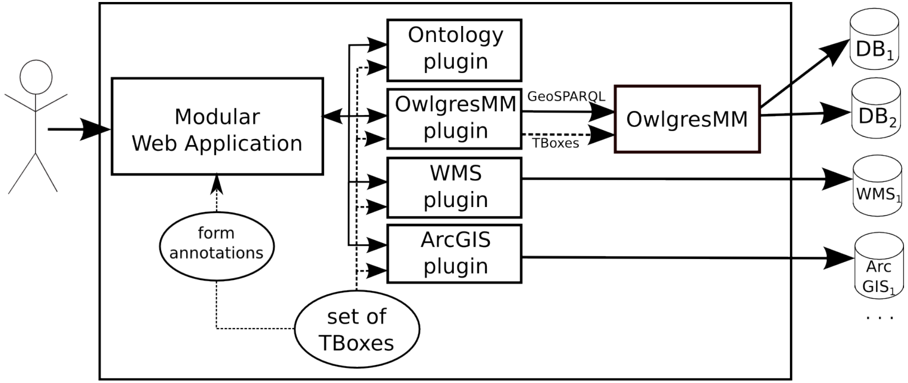
- OpenStreetMap – publicly available extensive data of the World.
- GeoNames – point data of the World with labels translated to many languages and hierarchically categorized.
    - These two sources are imported into our own relational database (PostgreSQL+PostGIS).
    - They have their ontologies: LinkedGeoData and GeoNames.
- Geoportal of the department of urban planning of Prague (ÚRM).
    - ArcGIS server, used remotely.

http://www.openstreetmap.org

http://www.geoportalpraha.cz



System Architecture

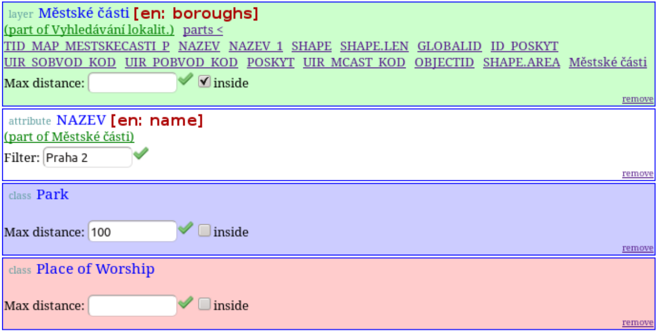


Simple Query Form

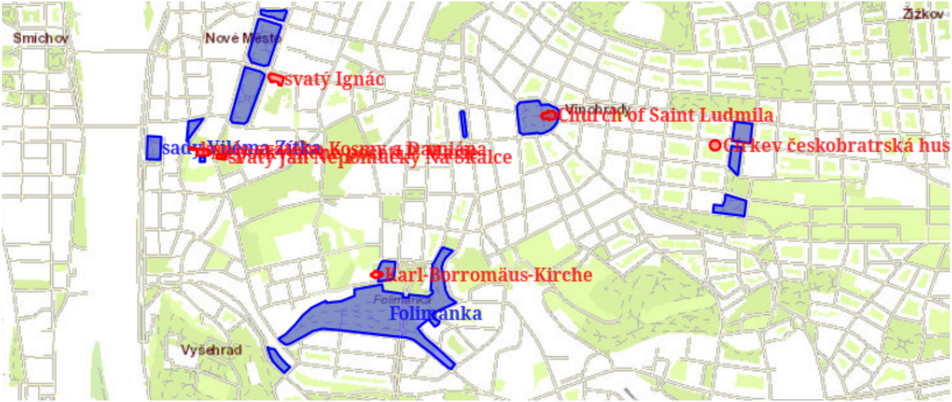
- Using text search (over data properties annotated as *searchable*), the system shows objects from different sources, and the user selects the relevant ones.
- These are added to the list of displayed items (the colored rectangles on the next slide).
- Various restrictions can be entered into the list:
  - spatial: max. distance, "inside" (both for objects annotated as *geometry*), and
    - Their semantics is that a restriction is applied to all other search results (with recursion).
  - text filtering (for objects annotated as *filterable*).
- Also linking the results pair-wise with spatial restrictions by links is possible, but it is not used in the following example.

Query Example

- Query: find places of worship, which are:
  - close to a park (within 100 m),
  - inside a specific part of Prague (borough “Praha 2”).
- Searching by keywords finds:
  - “Park” and “Place of worship” in OpenStreetMap ontology,
  - boroughs (“Městské části”) in ÚRM geoportal.
- Appropriate filters are applied.



Result of the Example



Displayed with OpenLayers.

Conclusion

- OnGIS is capable to distribute a query to multiple different sources.
- Queries support spatial restrictions.
- Based on OWL ontologies
  - for data source description,
  - for their integration, and
  - for making them available for querying.
- Independence on data source structure and technology.
- Support for spatial data from PostgreSQL+PostGIS, WMS and ArcGIS servers.
- Querying by a simple query form.

Ongoing and Future Work

- Ongoing development:
  - Structured query by a set theory-like expression with restricted, rigid structure (not a free text), with the help of autocompletion.
  - Access to RDF data via SPARQL endpoints.
    - Data in Linked Data initiative, e.g. DBpedia.
    - Some of them contain spatial data using W3C Basic Geo Vocabulary.
  - Using object properties (relations) in queries.
- Future work:
  - Support for other GIS servers (e.g. WFS).
  - Fully support GeoSPARQL.
  - ...