SPOI

SDI4apps: Points of Interest

Otakar Čerba

Last update: 2015-10-10

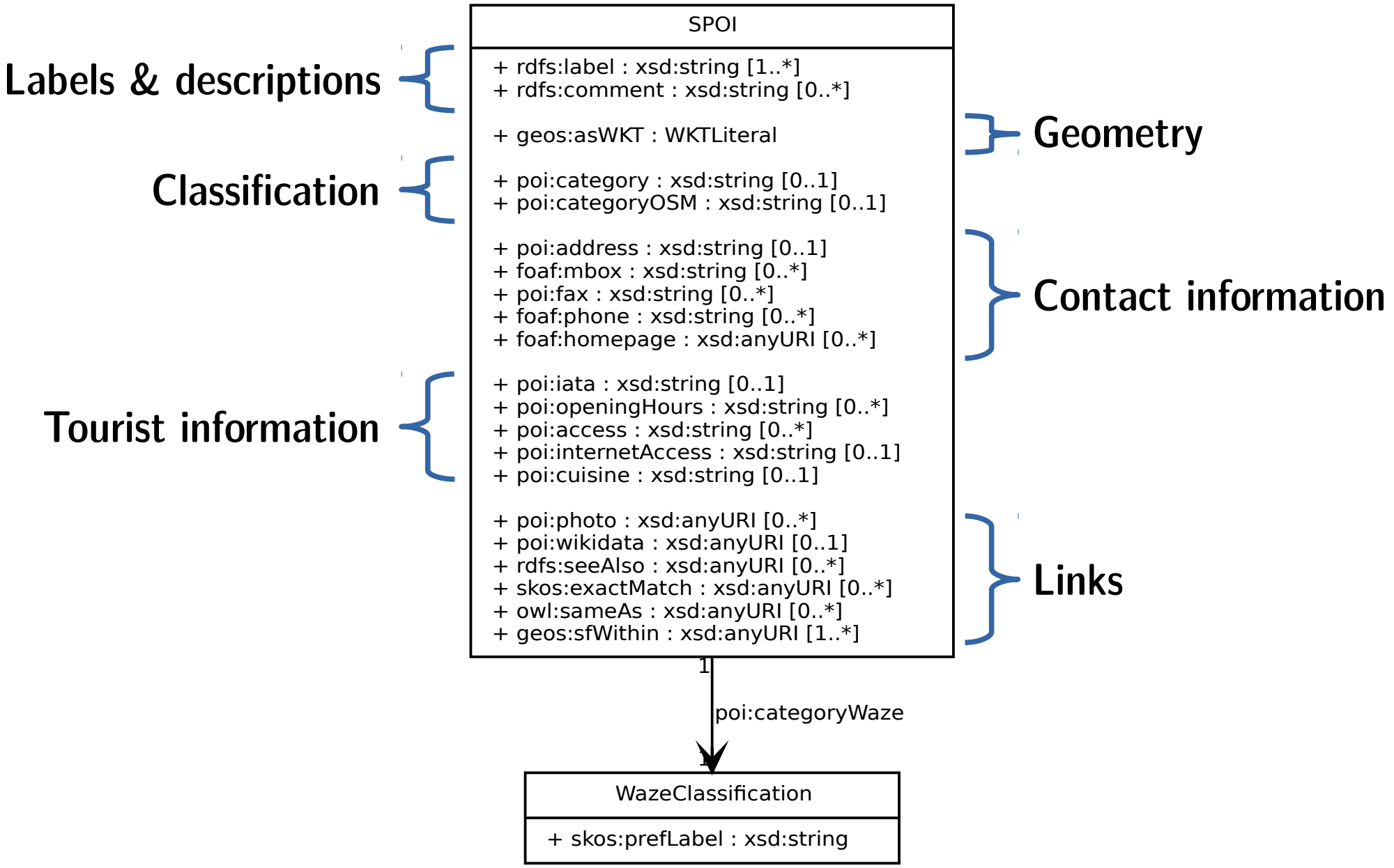
SPOI data set

- Open data for tourism and travelling
- Many heterogeneous input data
- Data harmonization process
- Based on standards, semantic description and Linked data
- Published on map portal and SPARQL endpoint
- Open Database License (ODbL)

Source data

- OpenStreetMap
- GeoNames.org (dumps)
- Local data – documents from Posumavi region, Sicily and travel agency
- Semantic data – experimental ontologies (OWL) of UWB (ski resort, sight in Rome)

Data model


SPOI
+ rdfs:label : xsd:string [1..*]
+ rdfs:comment : xsd:string [0..*]
+ geos:asWKT : WKTLiteral
+ poi:category : xsd:string [0..1]
+ poi:categoryOSM : xsd:string [0..1]
+ poi:address : xsd:string [0..1]
+ foaf:mbox : xsd:string [0..*]
+ poi:fax : xsd:string [0..*]
+ foaf:phone : xsd:string [0..*]
+ foaf:homepage : xsd:anyURI [0..*]
+ poi:iata : xsd:string [0..1]
+ poi:openingHours : xsd:string [0..*]
+ poi:access : xsd:string [0..*]
+ poi:internetAccess : xsd:string [0..1]
+ poi:cuisine : xsd:string [0..1]
+ poi:photo : xsd:anyURI [0..*]
+ poi:wikidata : xsd:anyURI [0..1]
+ rdfs:seeAlso : xsd:anyURI [0..*]
+ skos:exactMatch : xsd:anyURI [0..*]
+ owl:sameAs : xsd:anyURI [0..*]
+ geos:sfWithin : xsd:anyURI [1..*]
Labels & descriptions
Classification
Tourist information
Geometry
Contact information
Links
1
poi:categoryWaze
1
WazeClassification
+ skos:prefLabel : xsd:string

Data model

- **Identifier** – persistent URI
- Description – **label**(s) & description / comment (rdfs:) – multilinguality
- **Geometry** (WGS 84 W3C Geo Positioning → GeoSPARQL WKT as a link)
- Classification (**Waze**, OSM)
- Contact information (foaf:) – address, email...
- Tourism information – opening hours, cuisine...
- Links – **country**...

Identifier

- URI
- Combination
  - ISO 3166-1 alpha-2 country code
  - Acronym of category of POI according Waze navigation data
  - Coordinates (long_lat)

```
<rdf:Description rdf:about="http://www.sdi4apps.eu/poi/ML_NAT_0.8712_14.9746">
```

Links

- Classification – vocabulary to re-use
- Photos and pictures
- Same features
  - Web pages (Wikipedia, Wolfram|Alpha)
  - Linked data resources (DBpedia, GeoNames.org)
- Topological relation – to countries (GeoSPARQL)

SPOI & 5-star rating system

- * Accessibility: SPARQL endpoint & Open Database License (OdbL)
- ** Structured data: JSON, XML, CSV or RDF
- *** Non-proprietary format: RDF and other exports
- **** URIs: Identifiers of each POI
- ***** Links: see the previous slide

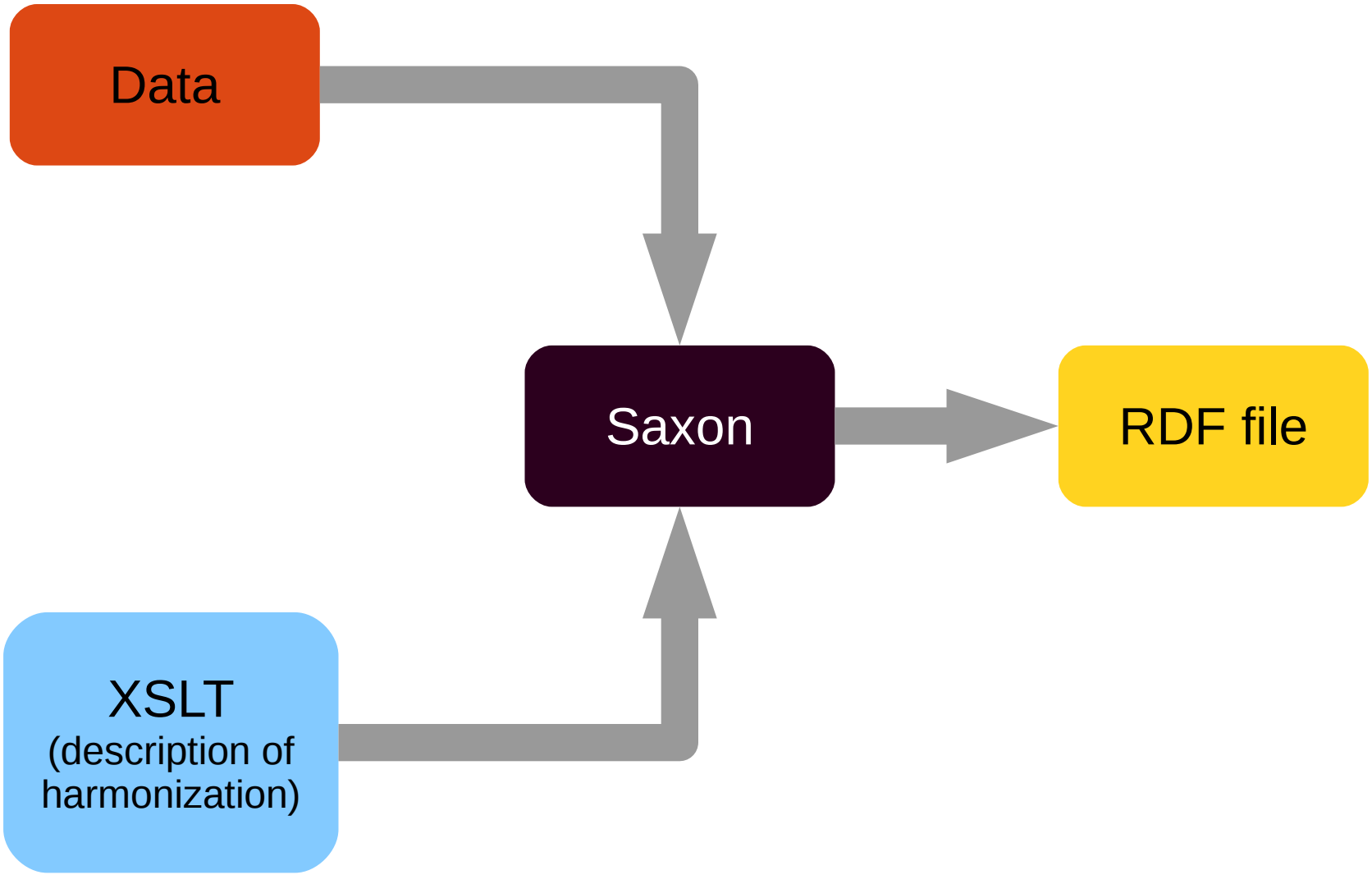
Data harmonization – steps

- Transcription to structured data (table, text → XML)
- Preparation of common vocabularies (Waze) and mappings (OSM, GeoNames.org → Waze)
- Filtering
- Adding information (links to countries)
- Transformation to common data model
- Export to common data format (RDF)

Data harmonization – technology

- XSLT 2.0 templates
  - Transformation language based on XML
  - Process XML based files and non-structured files
- Saxon processor
  - Java-based
  - XSLT + input data

Data harmonization - scheme



RDF example

<rdf:Description rdf:about="http://www.sdi4apps.eu/poi/IT_CAE_15.023497_37.613992 ">

<rdfs:label xml:lang="it">Casa Museo della civiltà contadina</rdfs:label>

<rdfs:label xml:lang="en">Museum of civilization</rdfs:label>

<geos:asWKT POINT(15.023497 37.613992)</geos:asWKT>

<poi:categoryWaze rdf:resource="
http://www.openvoc.eu/waze_classification#Culture_&_entertainment "/>

<poi:category>tourism</poi:category>

<poi:categoryOSM>tourism.museum</poi:categoryOSM>

<poi:address>Via Garibaldi, 58/58a – 95030 Nicolosi (CT)</poi:address>

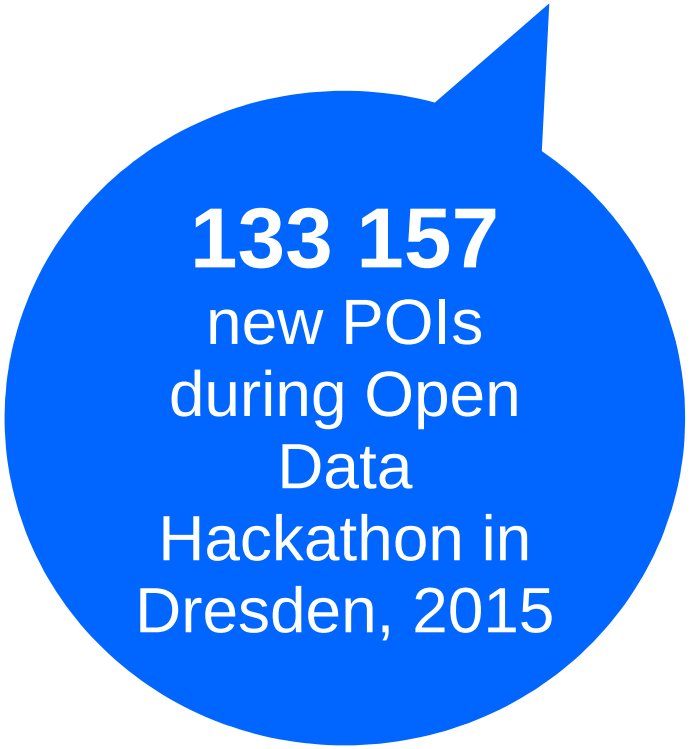
<foaf:phone>(+39) 095 – 910 980</foaf:phone>

<geos:sfWithin rdf:resource="http://www.geonames.org/3175395 "/>

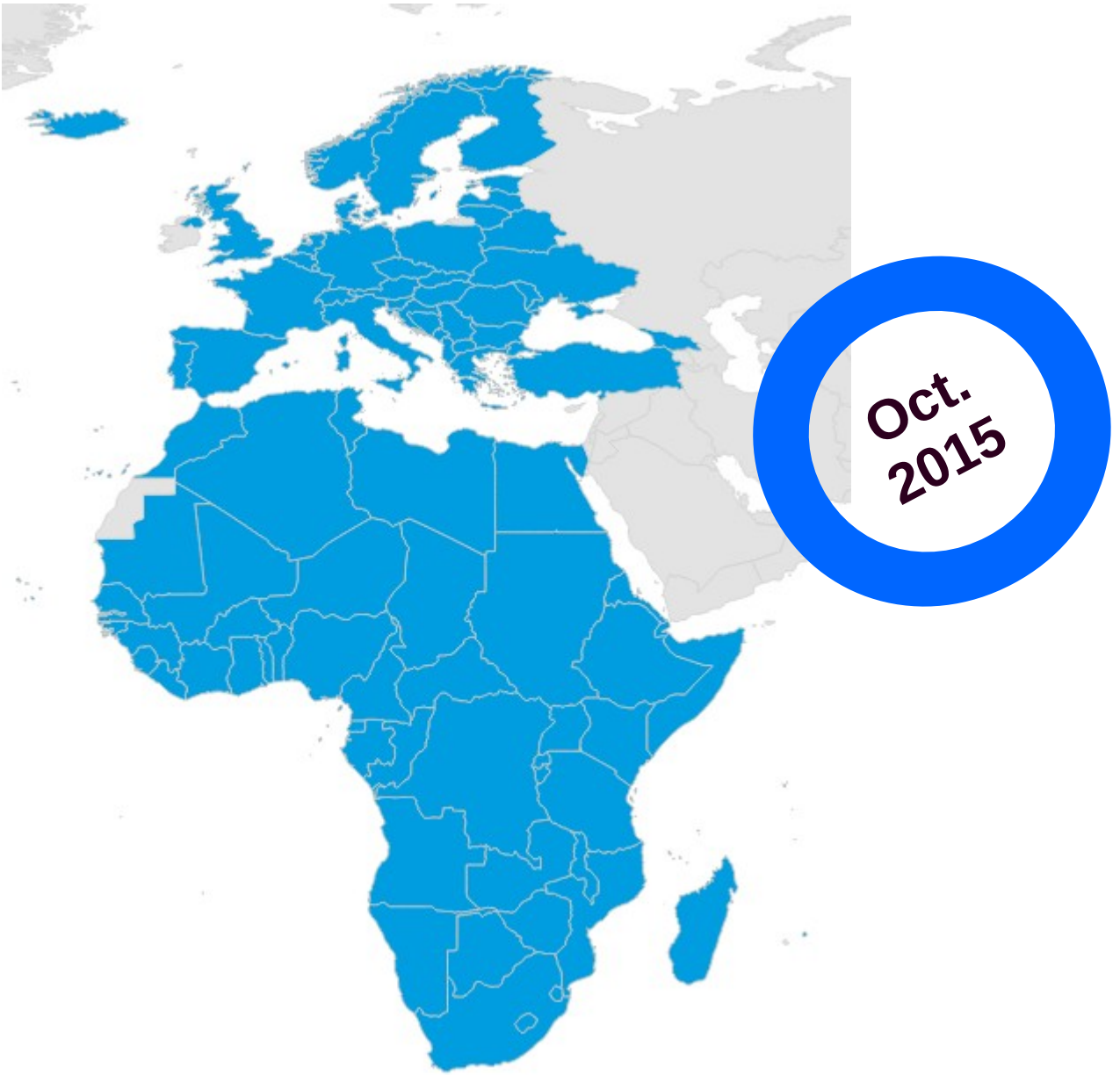
<geos:sfWithin rdf:resource="http://dbpedia.org/resource/Italy "/>

</rdf:Description>

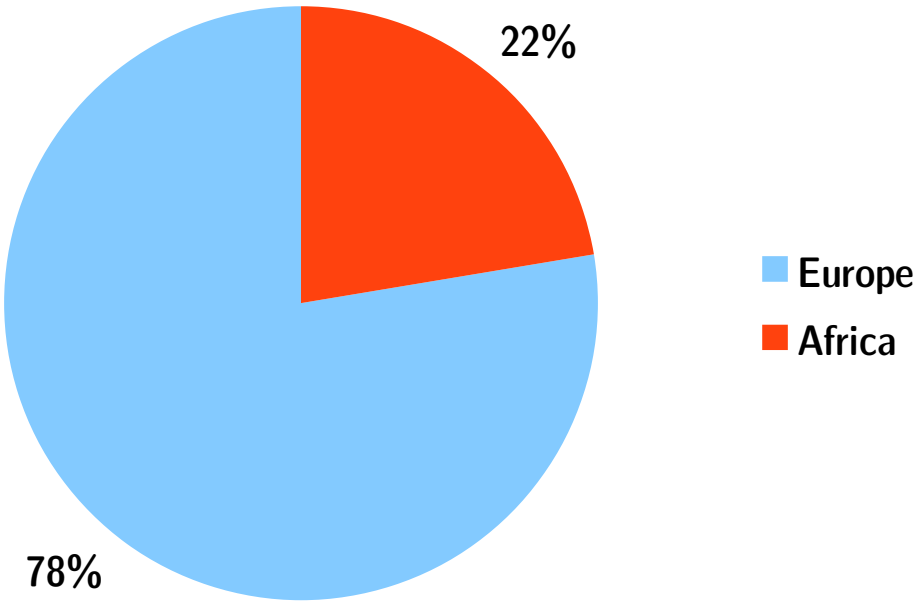
4 247 371 POIs

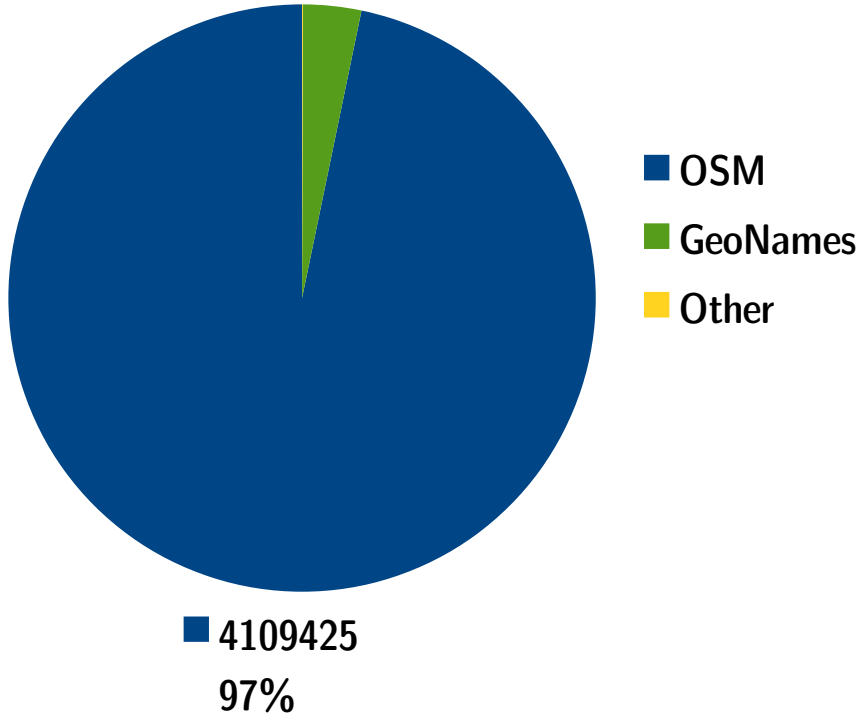


Coverage



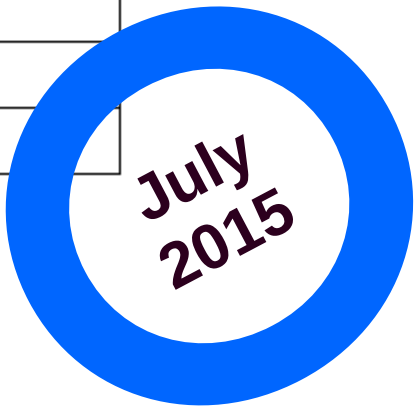
Statistics



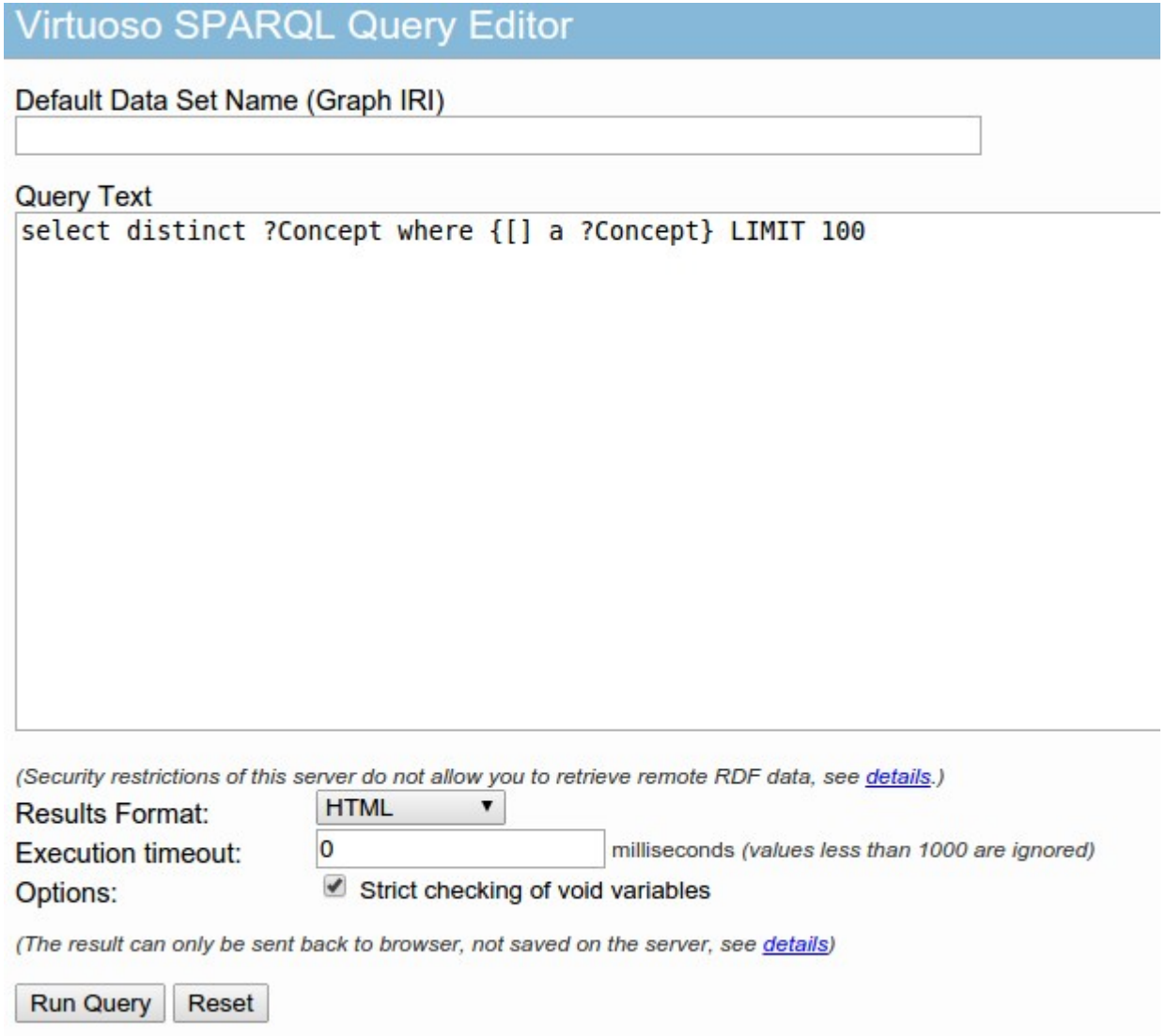


Comparison

Area	SPOI	OpenPOIs
Seaside resort (Croatia)	7	4
Submontane area (Czech republic)	1	0
Mountains (France)	1	1
Rural area (Germany)	28	28
Historical site (Greece)	9	10
Large city (Italy)	57	60
Coast (Latvia)	0	0
Small towns and villages (Netherlands)	6	8
Sport center (Norway)	46	41
Industrial area (Poland)	54	57



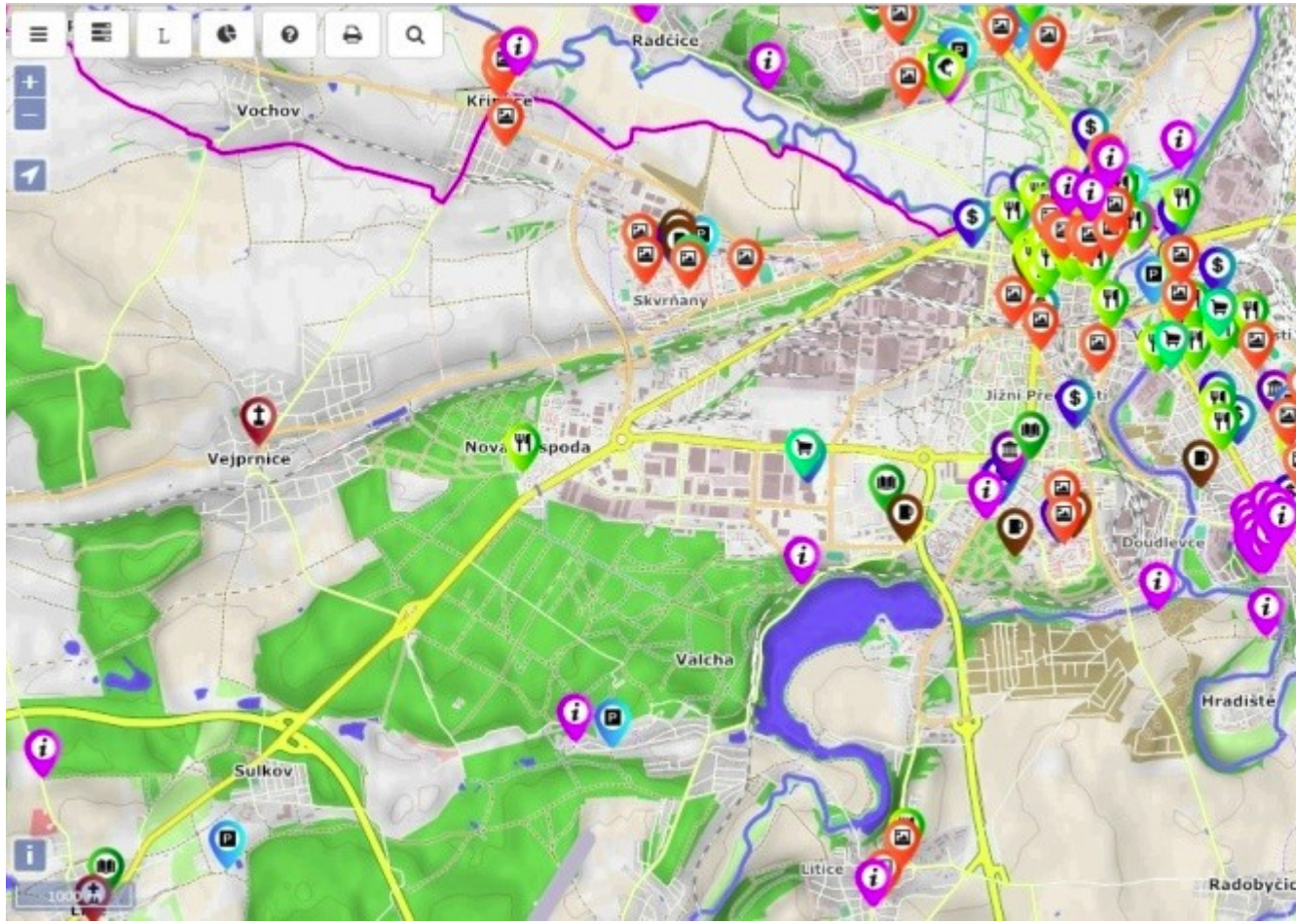
SPARQL endpoint



http://ha.isaf2014.info:8890/sparql

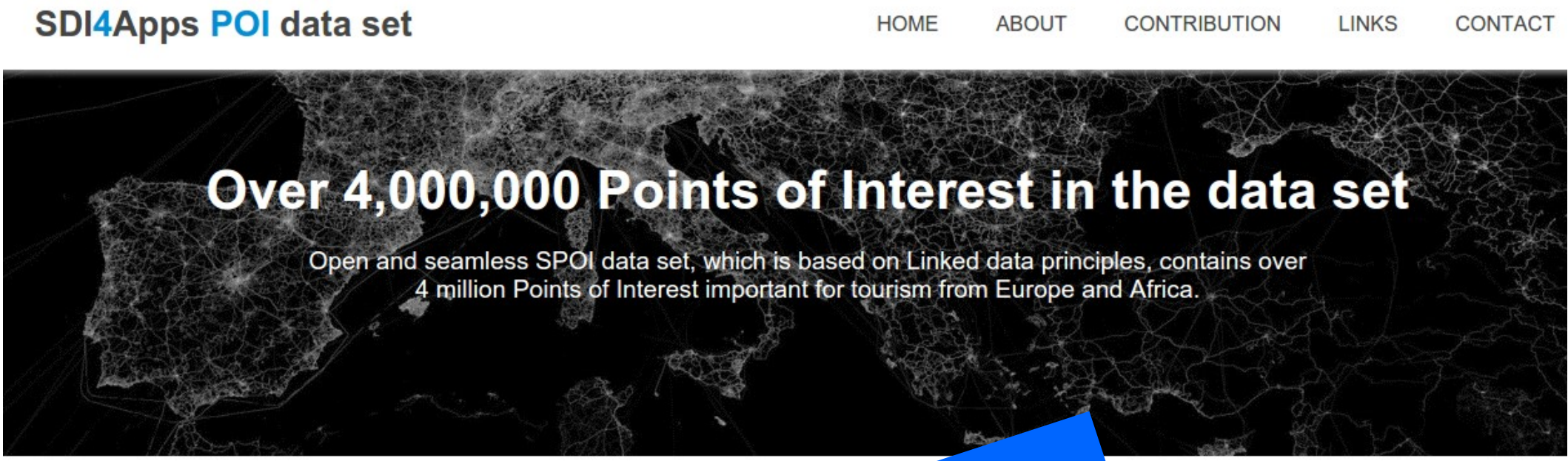
Map client

- HS layers
- Specific tiles for cycling and biking
- OpenCycleMap, MTB map, OSM...
- Panoramio, OpenWeatherMap



http://ng.hslayers.org/examples/geosparql/

Web page



sdi4apps.eu/spoi

gis.zcu.cz/spoi

Next steps

- Extension of information resources (imported data, links, APIs)
- Optimization of data model, data storage, data harmonization & data refining
- Cartographic challenges (clustering)
- Context-based application (user will get only information related to concrete needs)
- Analyses & itineraries & routing
- Advertising

Innovations

- Common data model
- Re-using existing standards
- Linked data
- SPARQL endpoint (http://ha.isaf2014.info:8890/sparql)