

Open Data Integration Using SPARQL and SPIN

A Case Study for the Tourism Domain

Antonino Lo Bue, **Alberto Machì**
ICAR-CNR Sezione di Palermo, Italy



Research funded by Italian PON SmartCities Dicit-InMoto-Orchestra project

EU Digital Agenda programs ISA & ISA2 and italian AGID encourage interoperability solutions for public administrations

At present:

- Heterogeneous processes and data structures
- Only islands of informations not communicating to each other
- No strategic and common vision

**Autonomy in
Public Open Data directives**



- *Dataset selection*
- *Cleaning*
- *Analysis & model creation*
- *Enrichment*
- *Interlinking*
- *Validation*
- *Publication*



Source: Agenzia per l'Italia Digitale

AgID Guidelines for Linked Open Data publication



- *Define the semantic model*
as a vocabulary of OWL classes, properties and axioms
- *Map data using the semantic model*
with a conversion of into RDF triples
- *Formalize the interlinking methods*
of each class of the model
using federated queries in SPARQL inference notation (SPIN)
rules
- *Execute SPIN rules*
on the RDF triple store or intermediate service

inMoto implementation



TopBraid Resource

Classes

- ▼ rdfs:Resource (0 + 1054)
 - ▼ geo:Point (0 + 4)
 - ▶ etCore:CulturalResource (4)
 - ▼ owl:Thing
 - etCore:Municipality
 - ▼ foaf:Agent
 - foaf:Group
 - foaf:Document
 - foaf:Image
 - foaf:Organization
 - foaf:Person
 - geo:SpatialThing
 - owl:Nothing
 - ▶ skos:Collection
 - skos:Concept
 - skos:ConceptScheme
 - ▶ rdf:Property (181 + 87)
 - rdf:Statement
 - ▶ rdfs:Class (104 + 195)
 - ▶ sp:SystemClass (0 + 155)
 - spin:Column
 - spin:ConstraintViolation
 - spin:ConstraintViolationLevel (4)
 - ▼ spin:Modules (0 + 294)
 - ▶ spin:Functions (0 + 127)
 - ▶ spin:Templates (0 + 167)
 - spl:TestCase (11)

Navigator

eTourismLite.rdf spin-mibac-test_2.ttl

Class Form

Name: etCore:CulturalResource

Annotations

Class Axioms

rdfs:subClassOf ▼

- geo:Point

spin:constraint ▼

spin:constructor ▼

spin:rule ▼

★ CONSTRUCT {

```

?this owl:sameAs ?lodm .
?this skos:subject ?lodm_sub .
?this skos:subject ?dbpedia_broader .
}
WHERE {
?this geo:lat ?lat1 .
?this geo:long ?long1 .
?this etCore:localityName ?mibac_loc_name .
SERVICE <http://194.119.214.82:8891/sparql> {
GRAPH <http://localhost:8891/linkedopendata-musei> {
?lodm rdfs:label ?lodm_label .
?lodm <http://www.w3.org/2006/vcard/ns#adr> ?adr .
?adr <http://www.w3.org/2006/vcard/ns#locality> ?loc .
?loc rdfs:seeAlso ?lodm_geo .
?adr <http://www.w3.org/2006/vcard/ns#street-address> ?lodm_address .
?lodm <http://www.w3.org/2006/vcard/ns#geo> ?lodm_latlong .
?lodm_latlong <http://www.w3.org/2006/vcard/ns#latitude> ?lat2 .
?lodm_latlong <http://www.w3.org/2006/vcard/ns#longitude> ?long2 .
?loc rdfs:label ?lodm_loc_label .
?lodm skos:subject ?lodm_sub .
} .
FILTER (<bif:haversine_deg_km>(xsd:float(?lat1), xsd:float(?long1), xsd:float(?lat2), xsd:float(?long2)) < 0.1) .
FILTER CONTAINS(str(?mibac_loc_name), str(?lodm_loc_label)) .
} .
SERVICE <http://dbpedia.org/sparql> {
?lodm_sub skos:broader ?dbpedia_broader .
} .
}

```

SPIN rule in an ontology editor



- Domain ontology model as unique reference for the data integration process
- On demand rule execution
mapping-interlinking workflow automated and triggered by specific events (queries or new instances)
- Re-use of interlinking patterns or user-defined functions and constraints
 - Consistency checking
 - Interlinking rules for specific LOD datasets
- Priority management and chain of rules

Benefits



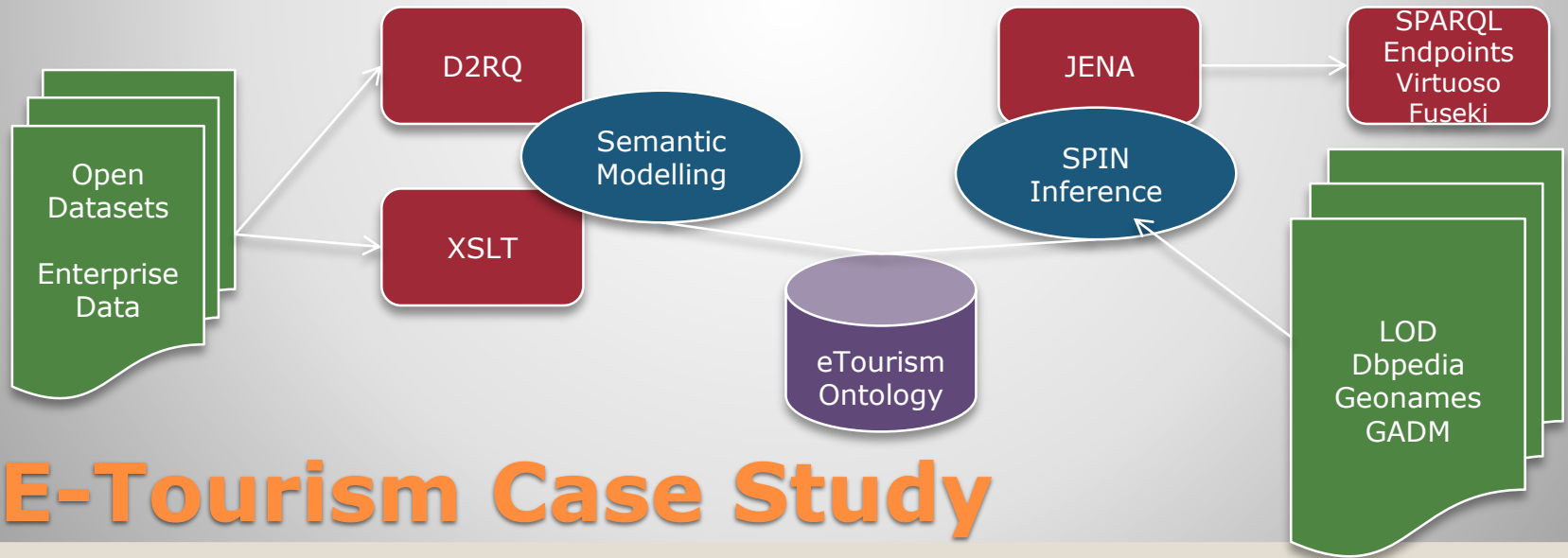
Re-use of existing open data about tourism

- Government/Public agencies (MIBAC, Regione, Camera di Commercio, SIT), Private/Enterprise data (Tripadvisor, Venere.com)

Conversion of tabular, XML and DB data into RDF triples

Semantic inference using federated SPARQL queries for interlinking

- DBpedia, Geonames, LinkedGeoData, GADM



E-Tourism Case Study





NAVIGATORE SEMANTICO DI RISORSE TURISTICHE IN CALABRIA

Criteri per la selezione

CHI:	0
COSA:	5
COME:	0
DOVE:	4

Mappa le risorse che appartengono ad almeno una delle categorie selezionate, e si trovano nel territorio di uno dei comuni scelti.

Cerca



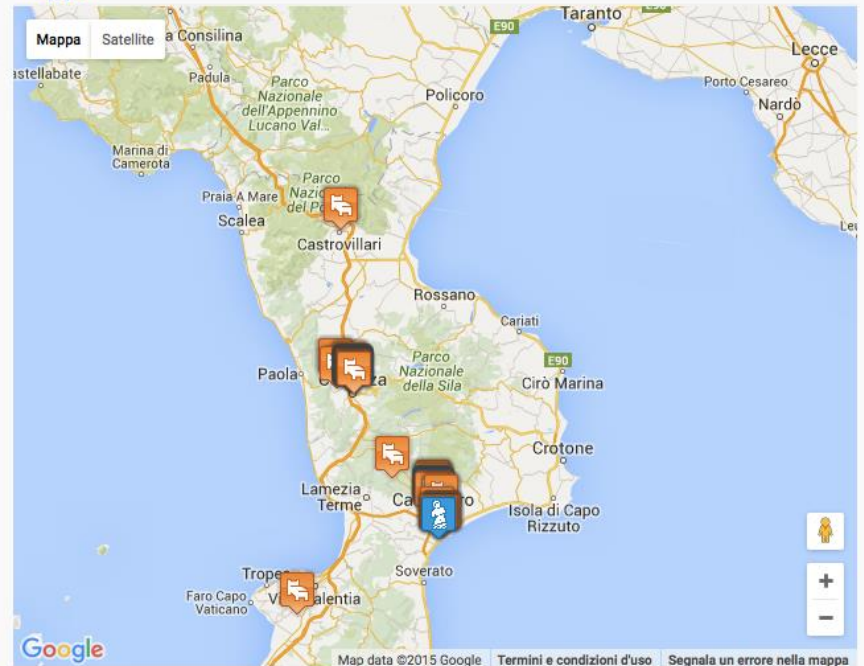
Ricomincia



Risorse individuate

Eventi	0
Punti di Interesse	19
Strutture	94

Mappa



Semantic Browser



AIxIA 2015 Ferrara Sept 24 2015

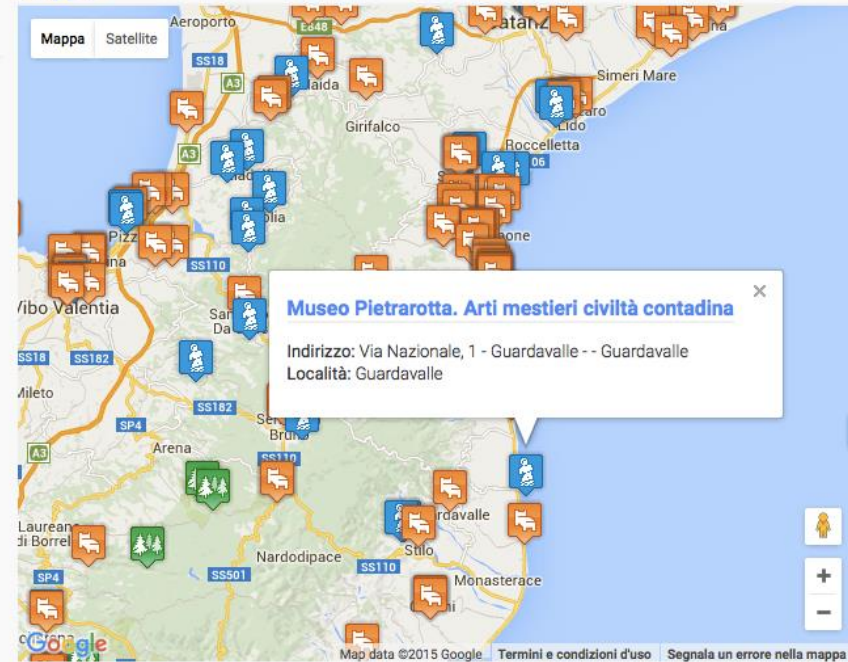
CALA GRECA BEACH at inMoto - ICAR-CNR Pa

http://slab.icar.cnr.it:8080/CalabriaKB/Struttura_80

EXPEDIA

Property	Value
<code>rdfs:comment</code>	<ul style="list-style-type: none">EXPEDIA ()
<code>etLite:deliversServiceOfType</code>	<ul style="list-style-type: none"><code>etLite:_airportShuttle</code><code>etLite:_babysittingService</code><code>etLite:_complimentaryBreakfast</code><code>etLite:_freeParking</code><code>etLite:_hairdressingService</code><code>etLite:_miniClub</code><code>etLite:_nightReception</code><code>etLite:_roomCleaning</code><code>etLite:_roomService</code><code>etLite:_safeService</code><code>etLite:_touristDesk</code><code>etLite:_unclassifiedType</code><code>etLite:_weddingService</code>
<code>etLite:hasEquipmentOfType</code>	<ul style="list-style-type: none"><code>etLite:</code><code>etLite:etFAS:_hairdryer</code>
<code>etLite:hasFacilityOfType</code>	<ul style="list-style-type: none"><code>etLite:_balcony</code><code>etLite:_bar</code><code>etLite:_courtyard</code><code>etLite:_garden</code><code>etLite:_loungeBar</code><code>etLite:_market</code><code>etLite:_privateBathroom</code><code>etLite:_publicParkingArea</code><code>etLite:_restaurant</code>

Mappa



Semantic Browser



AIXIA 2015 Ferrara Sept 24 2015

- eTourismLite Ontology:
<http://slab.icar.cnr.it/inmoto/eTourismLite/index.html>
- Open Data integration services:
<http://kossyra.pa.icar.cnr.it/open-data/index.html>

Technical details



Thanks for your kind attention



AixIA 2015 Ferrara Sept 24 2015

Technical details



AixIA 2015 Ferrara Sept 24 2015

Enrich enterprise tourist Accomodation Facilities descriptions With Points Of Interst, Places,Events and georeferences

Enterprise data

- [Tripadvisor](#), [Venere.com](#), [Booking.com](#), [Expedia](#) services descriptions (html)

Open data sources

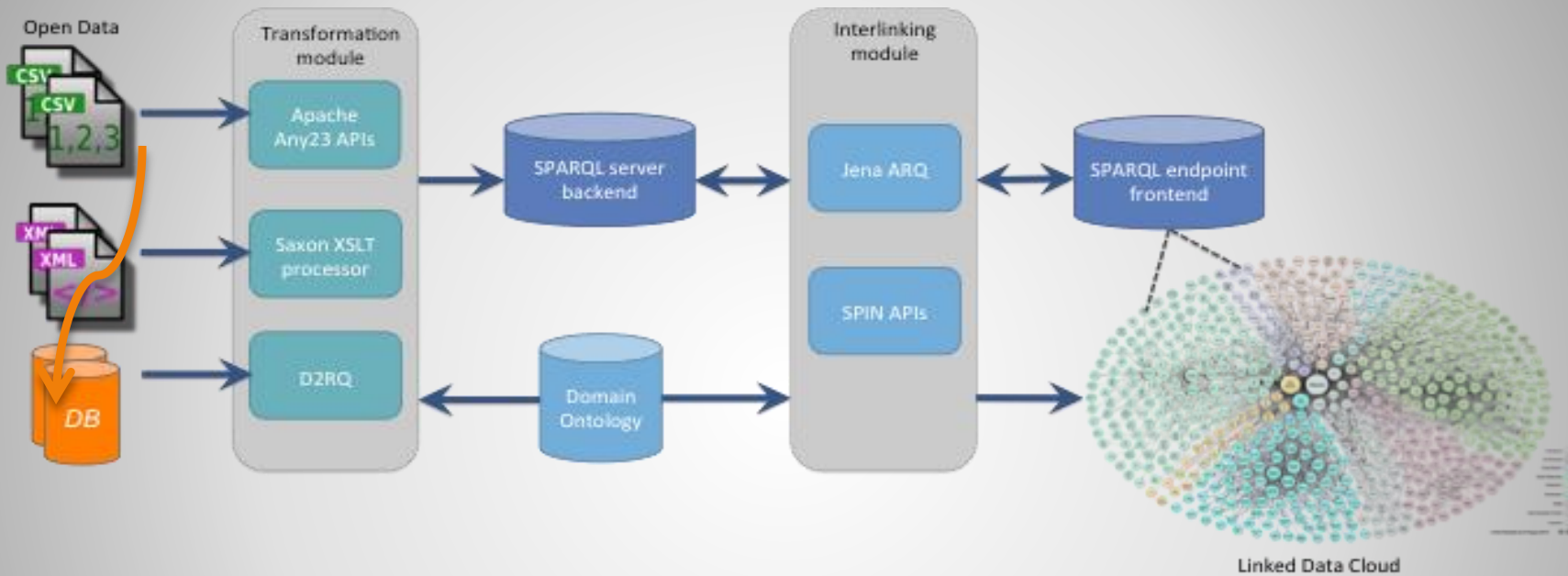
- [Trade Chambers](#) : administrative descriptions of Accomodation services (csv)
- [UNESCO Sites of EU Community interest](#) (csv)
- [MIBAC Italian Ministry of Cultural Resources DB Places of Culture](#) (xml)
- [Regional Government Territorial Informative System](#)

LOD sources:

- [DBPedia](#) RDF representation of Wikipedia subjects and contents
- [GADM-RDF](#) RDF spatial representation of all the administrative regions in the world
- [CulturaItalia Linked Data](#) data about museums, public archives and libraries.

E-Tourism Case Study





Web-Services Platform



AixIA 2015 Ferrara Sept 24 2015

SPARQL

- Horn clauses as RDF queries
- Federated queries :different endpoints in the same expression

SPIN

- SPARQL queries expressed as RDF Rules
- Rules sattached to ontology classes
- Expresses priority and chaining
- Expresses and evaluate constraints
- Support incremental reasoning
- Templates and ad hoc properties (also using Javascript)

SPARQL Inference Notation



Equivalence of a locality with a resource on the Linked Open Data Geonames dataset

```
1  INSERT {
2      ?municipality a etLite:Municipality .
3      ?municipality owl:sameAs ?geonames .
4      ?this etLite:inMunicipality ?municipality .
5  }
6  }
7  WHERE {
8      ?this geo:lat ?lat .
9      ?this geo:long ?long .
10     BIND (IRI(CONCAT("http://gadm.geovocab.org/services/withinRegion?",
11                    "lat=", ?lat, "&long=", ?long, "#point"))) AS ?GADMservice) .
12     SERVICE <http://slab.icar.cnr.it:8891/sparql> {
13         OPTIONAL {
14             ?GADMservice <http://gadm.geovocab.org/spatial#PP> ?GADMplace.
15             BIND (IRI(REPLACE(str(?GADMplace), "_", "/"))) AS ?GADMplace_loc).
16         } .
17     GRAPH <http://slab.icar.cnr.it/graph/GADM> {
18         ?GADMplace_loc rdfs:label ?GADMlabel .
19         ?GADMplace_loc a <http://gadm.geovocab.org/ontology#Level3> .
20         ?GADMplace_loc <http://gadm.geovocab.org/ontology#in_country> ?country .
21     }
22 } .
23 SERVICE <http://dbpedia.org/sparql> {
24     ?dbpedia owl:sameAs ?GADMplace .
25     ?dbpedia owl:sameAs ?geonames .
26     FILTER STRSTARTS(str(?geonames), "http://sws.geonames.org").
27 } .
28 BIND (IRI(CONCAT("http://slab.icar.cnr.it/testKB/",REPLACE(?GADMlabel, " ", "_")))
29       AS ?municipality).
30 }
```

Interlinking example 1



Equivalence of a museum instance with a resource on the *Linked Open Data Museums* dataset and inference of its cultural topics .

```
1  INSERT {
2    ?this owl:sameAs ?lodm .
3    ?this skos:subject ?lodm_sub .
4    ?this skos:subject ?dbpedia_broader
5  }
6  WHERE {
7    ?this geo:lat ?lat1 .
8    ?this geo:long ?long1 .
9    ?this etCore:localityName ?mibac_loc_name .
10
11   SERVICE <http://slab.icar.cnr.it:8891/sparql> {
12     GRAPH <http://slab.icar.cnr.it/graph/linkedopendata-musei> {
13       ?lodm rdfs:label ?lodm_loc_label .
14       ?lodm <http://www.w3.org/2006/vcard/ns#latitude> ?lat2 .
15       ?lodm <http://www.w3.org/2006/vcard/ns#longitude> ?long2 .
16       ?lodm skos:subject ?lodm_sub
17     }.
18   FILTER (<bif:haversine_deg_km>(xsd:float(?lat1), xsd:float(?long1), xsd:float(?lat2),
19     xsd:float(?long2)) < 0.1) .
20   FILTER ilo:levenshtein(str(?mibac_loc_name), str(?lodm_loc_label), 0.5).
21   }.
22   SERVICE <http://dbpedia.org/sparql> {
23     ?lodm_sub skos:broader ?dbpedia_broader
24   }.
25 }
26 }
```

Interlinking example 2

