



POLITECNICO DI MILANO

Corso di Laurea in Ingegneria Informatica

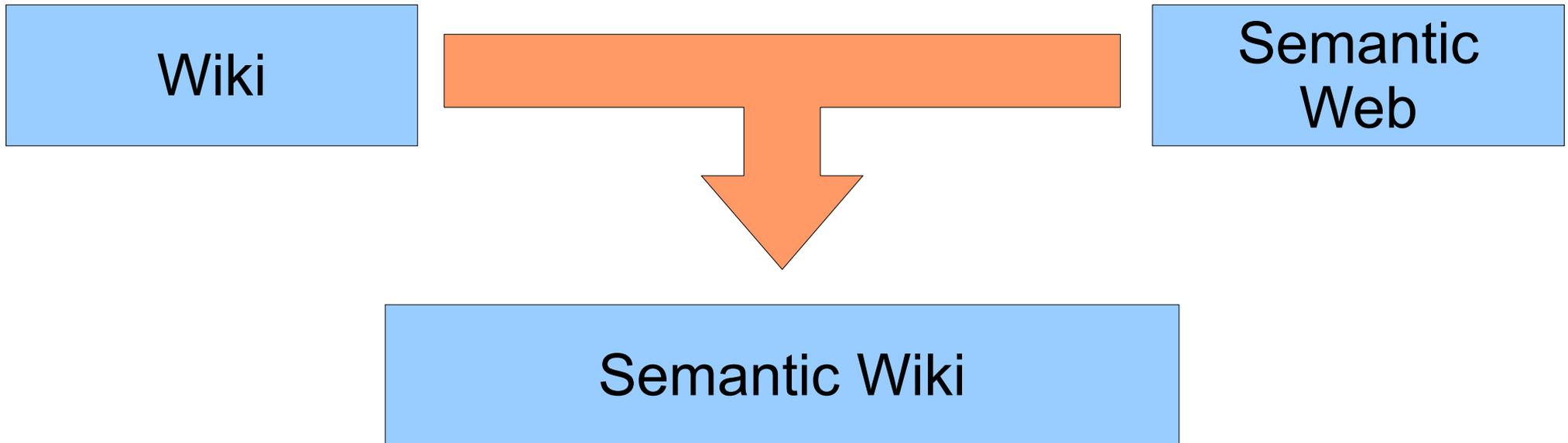
Facoltà di Ingegneria dell'Informazione

SVILUPPO DI UN'ONTOLOGIA GEOGRAFICA PER UN WIKI SEMANTICO

Tommaso Soru – matricola 662103

Introduzione

- **Web 2.0** → Content Management
- **Web 3.0** → Semantic Web

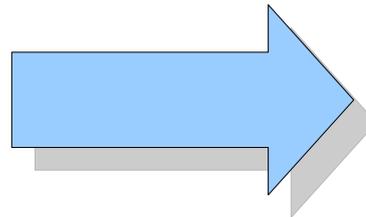


Scopo della tesi

- Sviluppo di un'ontologia geografica
- Quantità abbondante di dati semantici

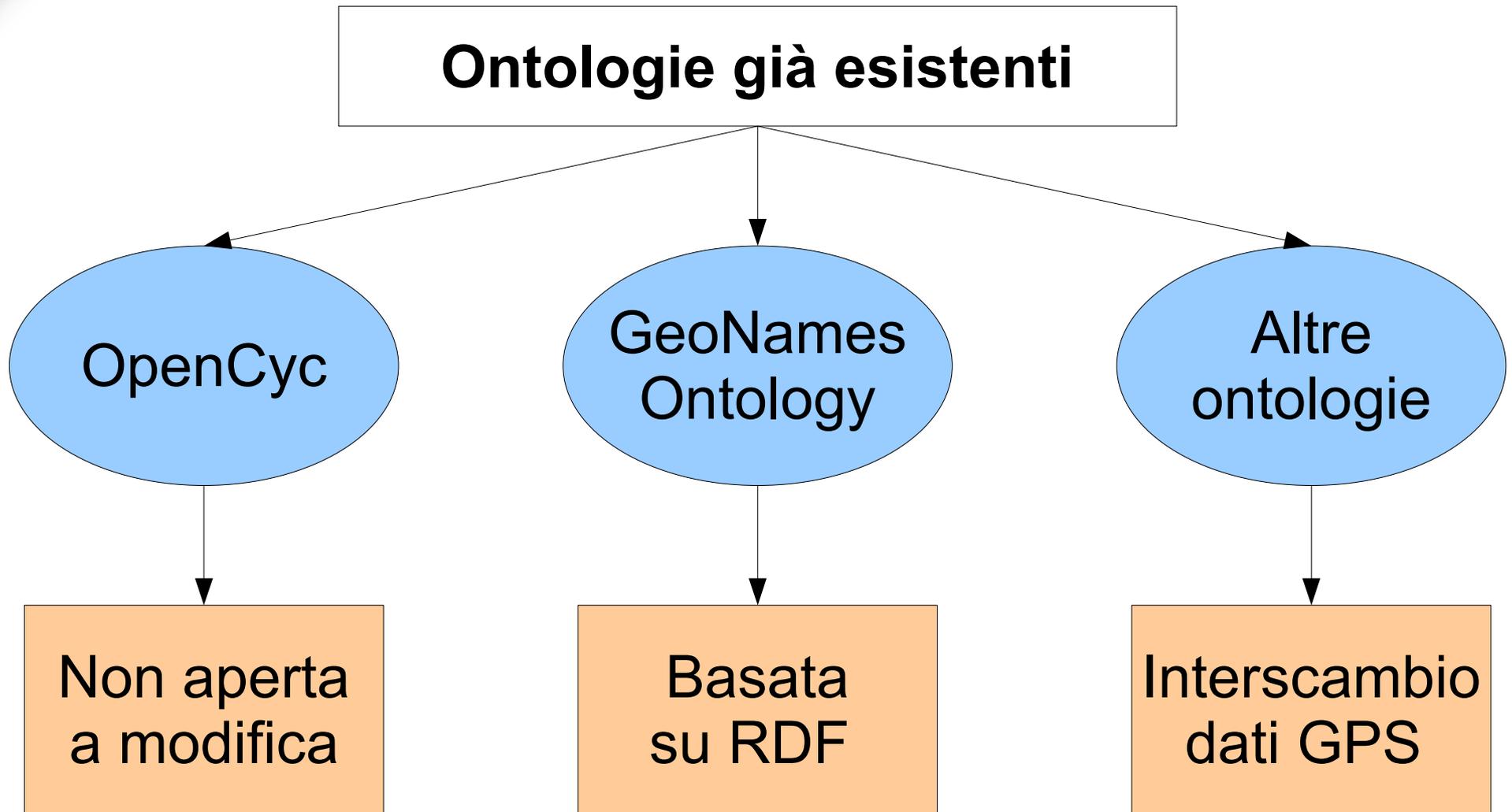
Programmazione **scraper** Java
per il recupero delle informazioni

Comune di Milano	
Location of the city of Milan	
Coordinates:  45°28'N 09°10'E	
Sovereign state	Italy
Region	Lombardy
Province	Province of Milan
Insubric settlement	c. 600 BC
Roman foundation	222 BC
Government	
 - Mayor	Letizia Moratti
Area	
 - City	182 km ² (70.3 sq mi)
 - Urban	1,982 km ² (765.3 sq mi)
Elevation	+120 m (394 ft)
Population (December 2006) ^[1]	
 - City	1,303,437 (2 nd)
 - Density	7,159/km ² (18,541.7/sq mi)
 - Metro	7.4 million
 - Called	Milanesi or Meneghini
Time zone	CET (UTC+1)
 - Summer (DST)	CEST (UTC+2)
Postal codes	20100, 20121-20162
Area code(s)	02
Patron saints	Ambrose (7 December)
Website:	www.comune.milano.it 



Milano	
<input type="checkbox"/> hasPopulation	"1306661" http://www.w3.org/2001/XMLSchema#int
<input type="checkbox"/> hasPopulationUptodate	"2003-01-01" http://www.w3.org/2001/XMLSchema#date
<input type="checkbox"/> hasTimezone	CentralEuropeanTime
<input type="checkbox"/> isContainedIn	Earth
<input type="checkbox"/> isContainedIn	Europe
<input type="checkbox"/> isContainedIn	Lombardy
<input type="checkbox"/> isContainedIn	ProvinciaDiMilano
<input type="checkbox"/> isContainedIn	Italy
<input type="checkbox"/> isDirectlyContainedIn	ProvinciaDiMilano
<input type="checkbox"/> isPopulatedPlaceOf	Italy
<input type="checkbox"/> isPopulatedPlaceOf	Earth
<input type="checkbox"/> isPopulatedPlaceOf	Europe
<input type="checkbox"/> isPopulatedPlaceOf	Lombardy
<input type="checkbox"/> isPopulatedPlaceOf	ProvinciaDiMilano
<input type="checkbox"/> type	PopulatedPlace
<input type="checkbox"/> type	AdministrativeDivision3

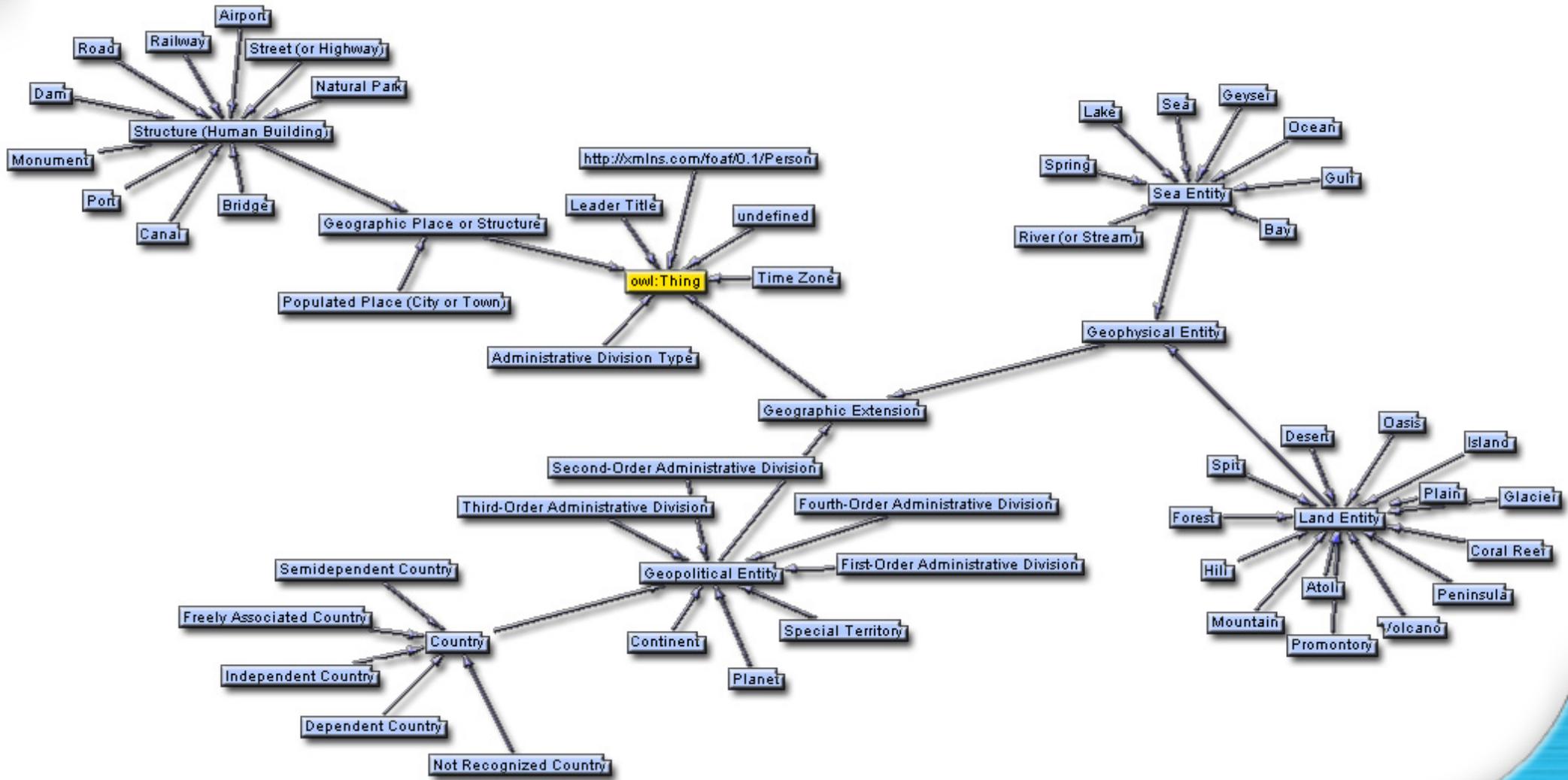
L'ontologia geografica



1. PROGETTAZIONE
2. POPOLAMENTO
3. CONSULTAZIONE

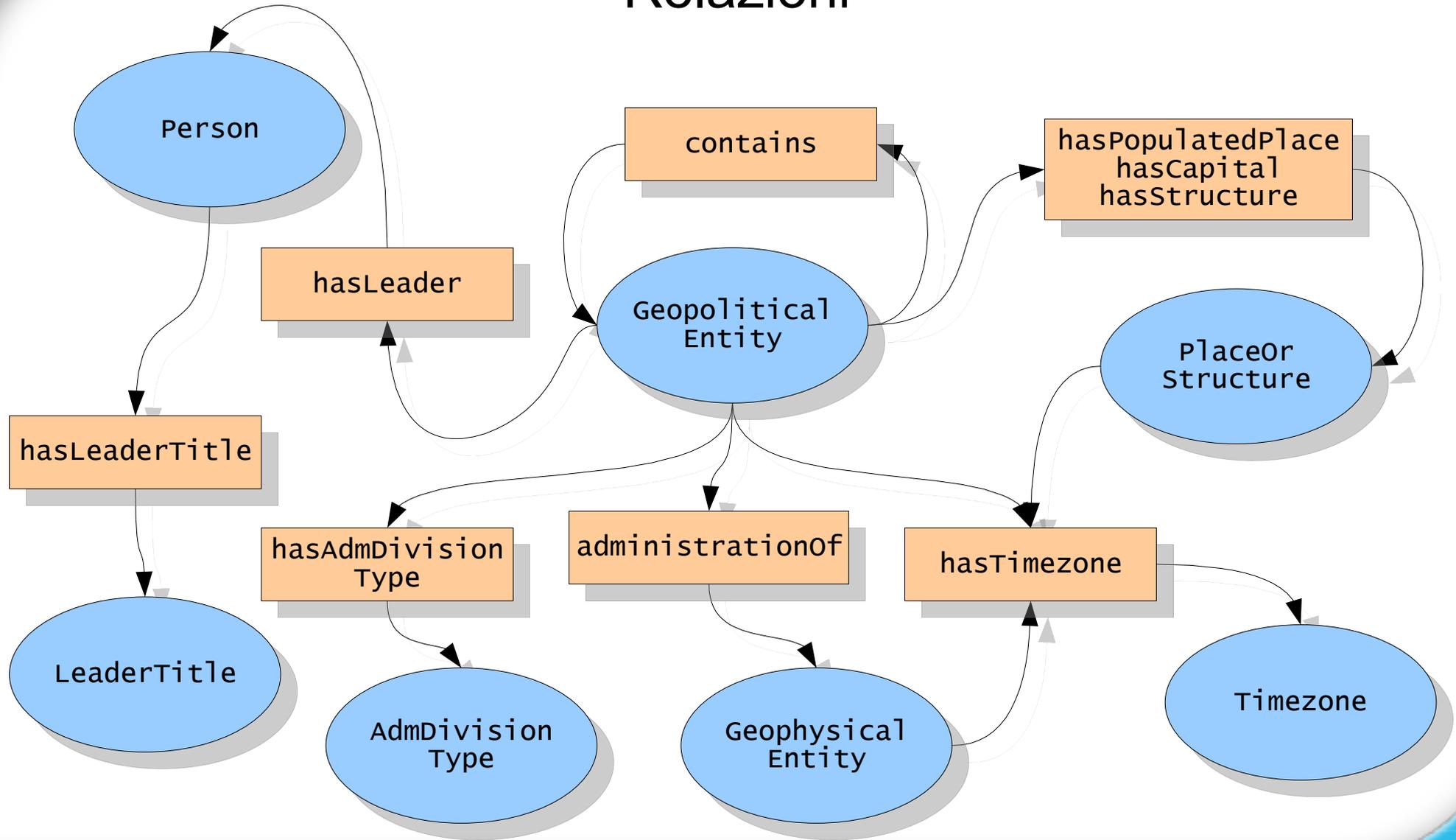
1. Progettazione

Le classi dell'ontologia geografica.



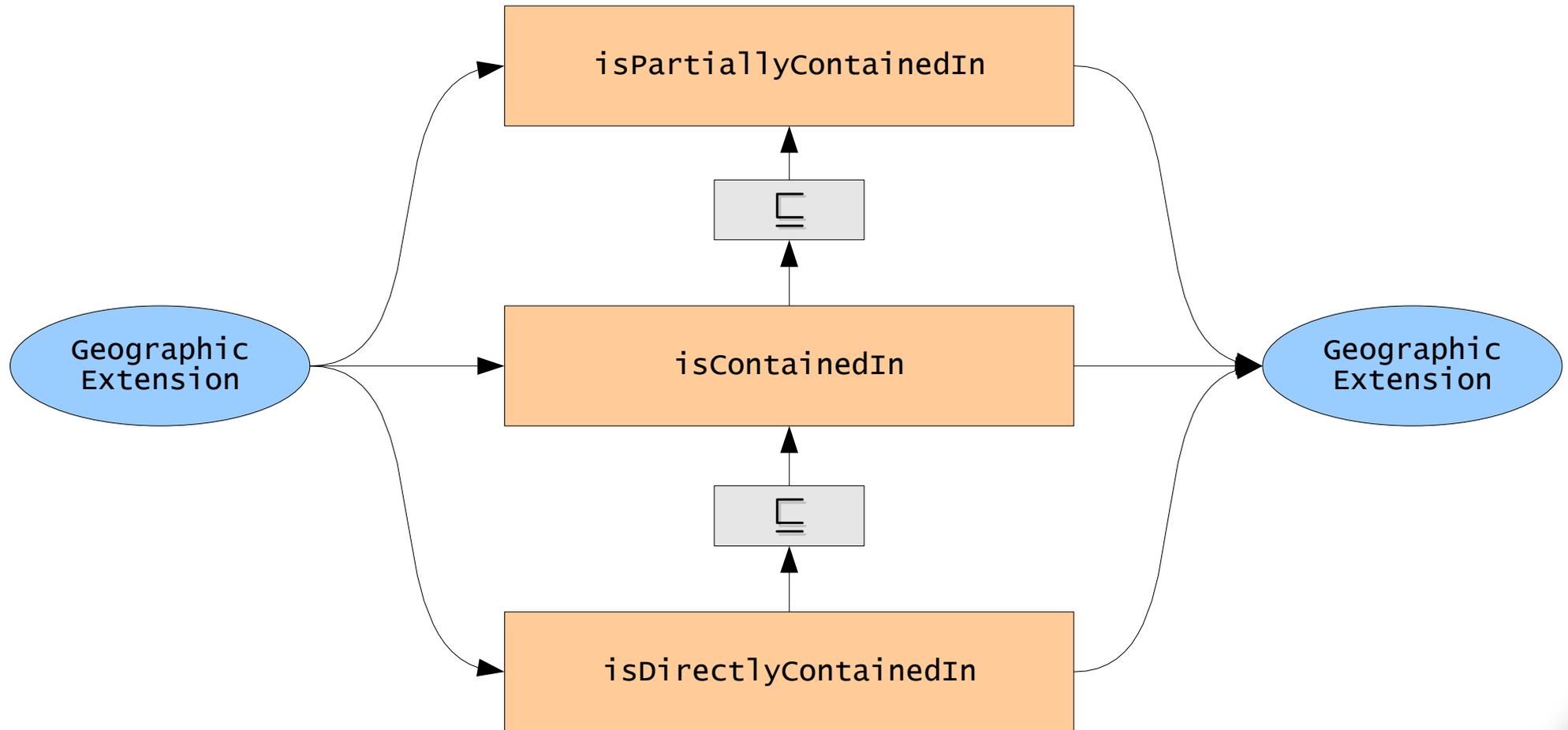
1. Progettazione

Relazioni



1. Progettazione

Le proprietà **contains** \leftrightarrow **isContainedIn**



1. Progettazione

Le proprietà **contains** \leftrightarrow **isContainedIn**. Problema.

Turkey \rightarrow isDirectlyContainedIn \rightarrow Europe
Turkey \rightarrow isDirectlyContainedIn \rightarrow Asia



Turkey \rightarrow isContainedIn \rightarrow Europe
Turkey \rightarrow isContainedIn \rightarrow Asia

Marmara \rightarrow isDirectlyContainedIn \rightarrow Turkey



Marmara \rightarrow isContainedIn \rightarrow Turkey

Turkey \rightarrow isContainedIn \rightarrow Europe
Turkey \rightarrow isContainedIn \rightarrow Asia
Marmara \rightarrow isContainedIn \rightarrow Turkey



Marmara \rightarrow isContainedIn \rightarrow Europe
Marmara \rightarrow isContainedIn \rightarrow Asia



1. Progettazione

Le proprietà **contains** ↔ **isContainedIn**. Soluzione.

Turkey → **isPartiallyContainedIn** → Europe
Turkey → **isPartiallyContainedIn** → Asia

Marmara → **isDirectlyContainedIn** → Turkey
Anatolia → **isDirectlyContainedIn** → Turkey

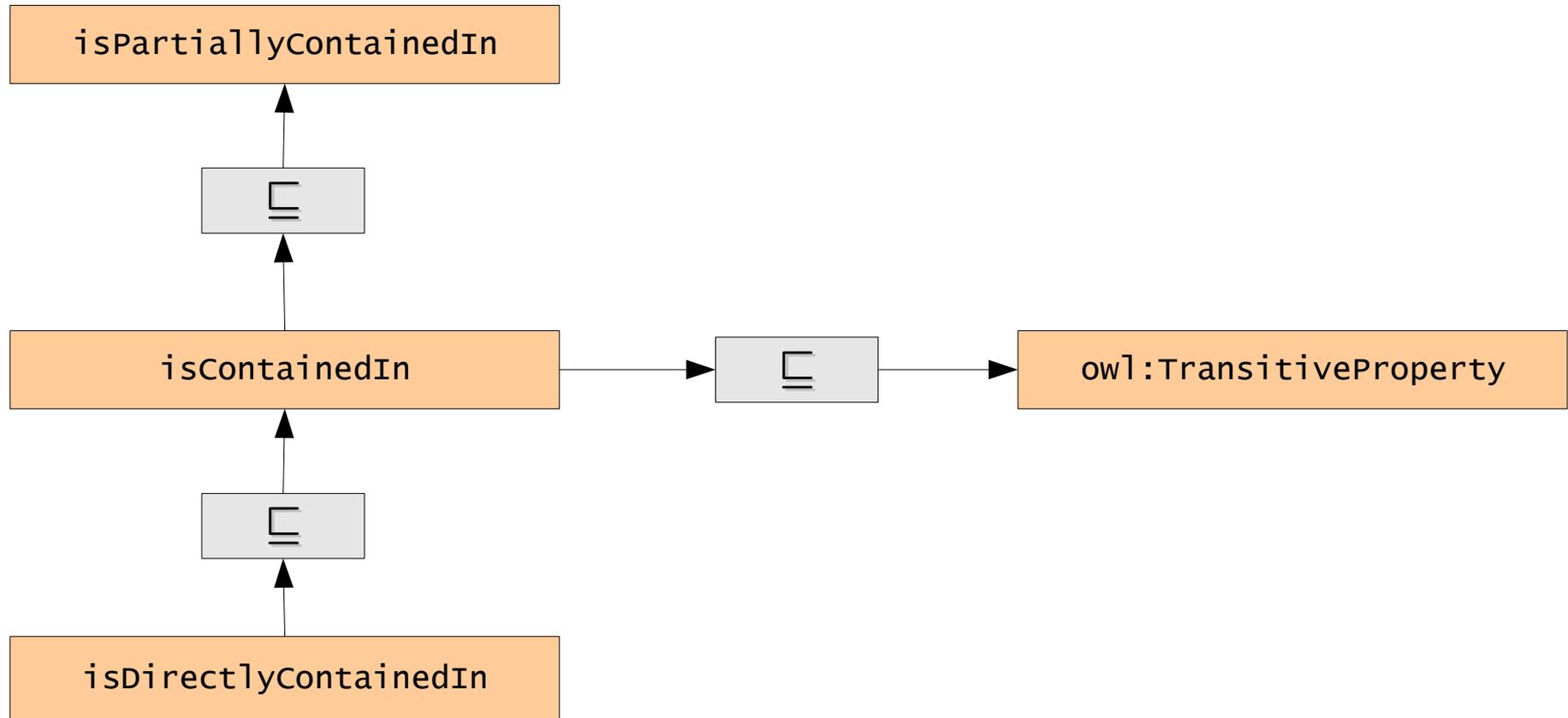
Utilizzando le coordinate geografiche...
(proprietà **geo:long** e **geo:lat**)

Marmara → **isContainedIn** → Europe
Anatolia → **isContainedIn** → Asia



1. Progettazione

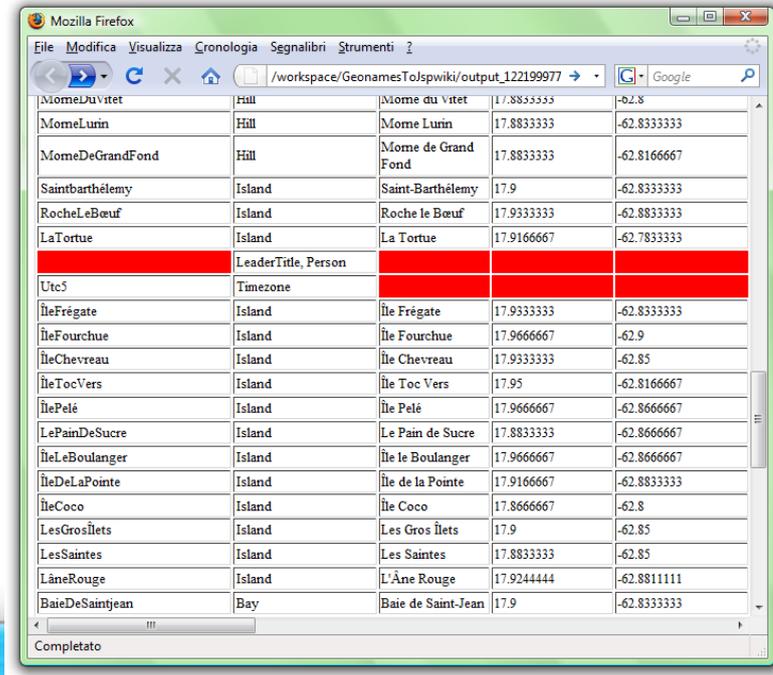
Le proprietà **contains** \leftrightarrow **isContainedIn**. Transitività.



2. Popolamento

Lo **scraper** java:

- recupera informazioni semantiche (Wikipedia)
- esegue query a webservice (GeoNames)
- effettua un controllo dei dati importati e segnala le anomalie in un file HTML
- traduce i dati in triple RDF
- inserisce le triple in ontologia

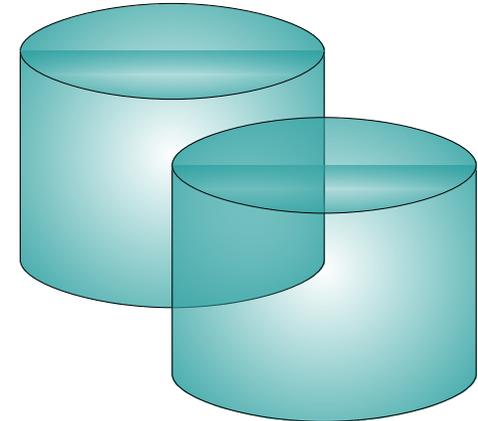
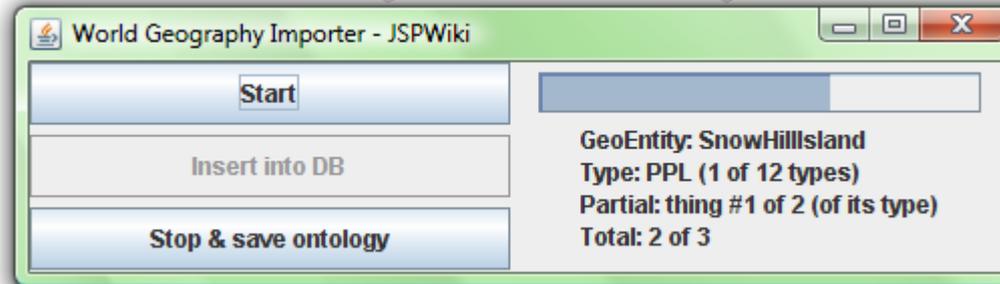


Name	Type	Name	Longitude	Latitude
MomeDuVitet	Hill	Mome du Vitet	17.8833333	-62.8
MomeLurin	Hill	Mome Lurin	17.8833333	-62.8333333
MomeDeGrandFond	Hill	Mome de Grand Fond	17.8833333	-62.8166667
Saintbarthelemy	Island	Saint-Barthélemy	17.9	-62.8333333
RocheLeBœuf	Island	Roche le Bœuf	17.9333333	-62.8833333
LaTortue	Island	La Tortue	17.9166667	-62.7833333
Uto5	Timezone			
ÎleFrégate	Island	Île Frégate	17.9333333	-62.8333333
ÎleFourchue	Island	Île Fourchue	17.9666667	-62.9
ÎleChevreau	Island	Île Chevreau	17.9333333	-62.85
ÎleTocVers	Island	Île Toc Vers	17.95	-62.8166667
ÎlePelé	Island	Île Pelé	17.9666667	-62.8666667
LePainDeSucre	Island	Le Pain de Sucre	17.8833333	-62.8666667
ÎleLeBoulangier	Island	Île le Boulangier	17.9666667	-62.8666667
ÎleDeLaPointe	Island	Île de la Pointe	17.9166667	-62.8833333
ÎleCoco	Island	Île Coco	17.8666667	-62.8
LesGrosÎlets	Island	Les Gros Îlets	17.9	-62.85
LesSaintes	Island	Les Saintes	17.8833333	-62.85
LâneRouge	Island	L'Âne Rouge	17.9244444	-62.8811111
BaieDeSaintjean	Bay	Baie de Saint-Jean	17.9	-62.8333333

2. Popolamento

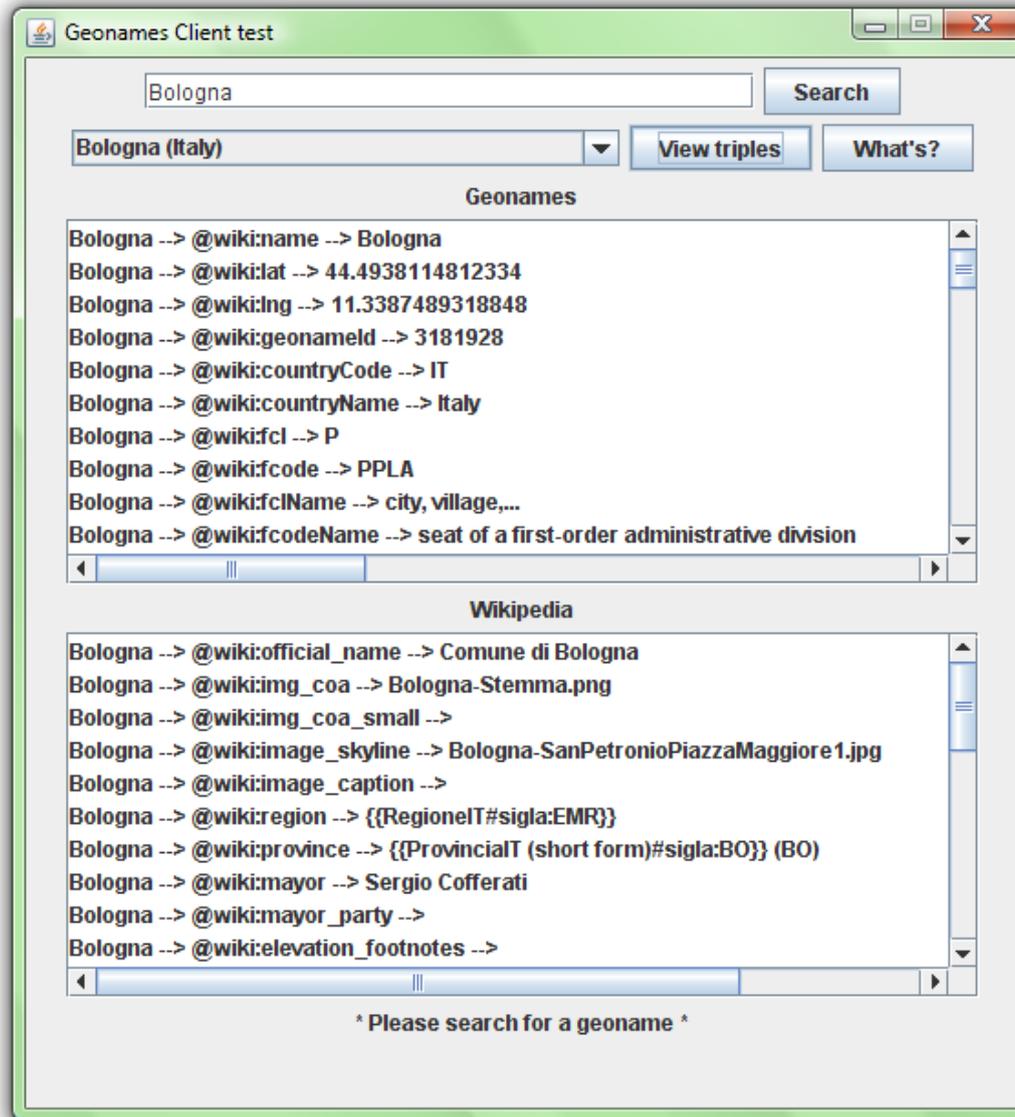
Wikipedia

GeoNames



Semantic JSPWiki

2. Popolamento



The screenshot shows a window titled "Geonames Client test" with a search bar containing "Bologna" and a "Search" button. Below the search bar, a dropdown menu shows "Bologna (Italy)" and buttons for "View triples" and "What's?". The main content is divided into two sections: "Geonames" and "Wikipedia".

Geonames

- Bologna --> @wiki:name --> Bologna
- Bologna --> @wiki:lat --> 44.4938114812334
- Bologna --> @wiki:lng --> 11.3387489318848
- Bologna --> @wiki:geonameid --> 3181928
- Bologna --> @wiki:countryCode --> IT
- Bologna --> @wiki:countryName --> Italy
- Bologna --> @wiki:fcl --> P
- Bologna --> @wiki:fcode --> PPLA
- Bologna --> @wiki:fclName --> city, village,...
- Bologna --> @wiki:fcodeName --> seat of a first-order administrative division

Wikipedia

- Bologna --> @wiki:official_name --> Comune di Bologna
- Bologna --> @wiki:img_coa --> Bologna-Stemma.png
- Bologna --> @wiki:img_coa_small -->
- Bologna --> @wiki:image_skyline --> Bologna-SanPetronioPiazzaMaggiore1.jpg
- Bologna --> @wiki:image_caption -->
- Bologna --> @wiki:region --> {{RegioneIT#sigla:EMR}}
- Bologna --> @wiki:province --> {{ProvinciaIT (short form)#sigla:BO}} (BO)
- Bologna --> @wiki:mayor --> Sergio Cofferati
- Bologna --> @wiki:mayor_party -->
- Bologna --> @wiki:elevation_footnotes -->

* Please search for a geoname *

3. Consultazione

Milano

Your trail: [Main](#)

[View](#) [Attach](#) [Info](#) [Edit](#) [More...](#)

JSPWiki v2.6.1-cvs-7

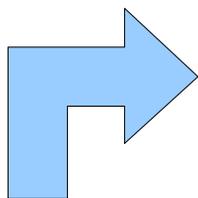
Milan (Italian: Milano; Western Lombard: Milan (listen)) is one of the largest cities in Italy, located in the plains of Lombardy. It is the capital in the Province of Milano. The municipality (Comune di Milano) has a population of 1.3 million. The Milan metropolitan area, depending on the specific definition, has a population ranging from 2.9 to 9.8 million. The municipal border covers a relatively small area (about one-eighth that of Rome) because of the historical development of high density centres in agriculturally rich Lombardy.

Milan is renowned as one of the world capitals of design and fashion. The English word milliner is derived from the name of the city. The Lombard metropolis is famous for its fashion houses and shops (such as along via Montenapoleone) and the Galleria Vittorio Emanuele in the Piazza Duomo (reputed to be the world's oldest shopping mall). The city hosted the World Exposition in 1906 and will host the Universal Expo in 2015. Inhabitants of Milan are referred to as "Milanese" (Italian: Milanesi or informally Meneghini or Ambrosiani).

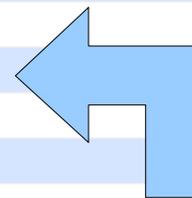
The Olona river, the Lambro river and the Seveso creek run through Milan. Olona and Seveso run mostly underground.

- hasPopulation	"1306661" http://www.w3.org/2001/XMLSchema#int "
- hasPopulationUptodate	"2003-01-01" http://www.w3.org/2001/XMLSchema#date "
- hasTimezone	CentralEuropeanTime
- isContainedIn	Earth
- isContainedIn	Europe
- isContainedIn	Lombardy
- isContainedIn	ProvinciaDiMilano
- isContainedIn	Italy
- isDirectlyContainedIn	ProvinciaDiMilano
- isPopulatedPlaceOf	Italy
- isPopulatedPlaceOf	Earth
- isPopulatedPlaceOf	Europe
- isPopulatedPlaceOf	Lombardy
- isPopulatedPlaceOf	ProvinciaDiMilano
- type	PopulatedPlace
- type	AdministrativeDivision3

Completo



human readable



machine readable

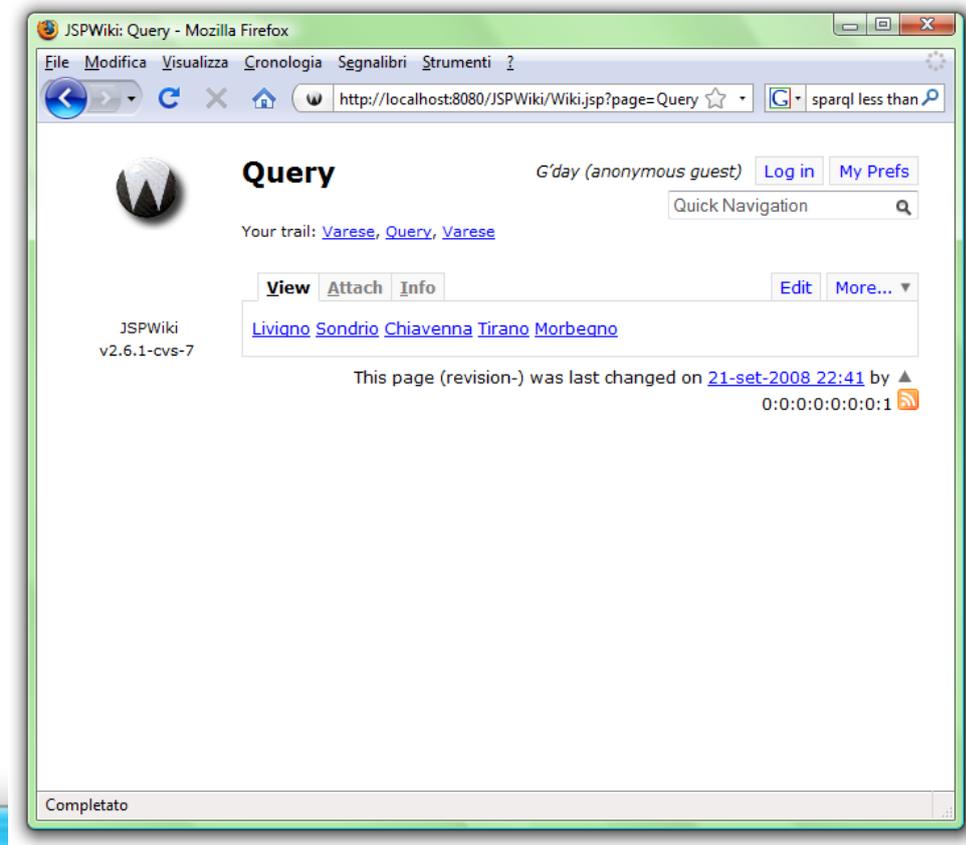


3. Consultazione

Esempio di query SPARQL.

Città in provincia di Sondrio con almeno 5,000 abitanti.

```
SELECT ?x WHERE {?x <http://www.w3.org/1999/02/22-rdf-syntax-ns#type>  
<http://geoontology.altervista.org/geoontology.owl#PopulatedPlace> . ?x  
<http://geoontology.altervista.org/geoontology.owl#isPopulatedPlaceOf>  
<http://geoontology.altervista.org/geoontology.owl#ProvinciaDiSondrio> . ?x  
<http://geoontology.altervista.org/geoontology.owl#hasPopulation> ?y . FILTER (?y > 5000)}
```



Conclusioni

Pro

- Quantità di informazioni raccolte
- Rispetto delle specifiche W3C
- Ontologia dotata di buona espressività

Contro

- Valutazione dell'aggiornamento dei dati

Sviluppi futuri

- Primo passo verso un'ontologia completa (Cyc)
- Oggetto di studio come mezzo tra i linguaggi naturale e semantico