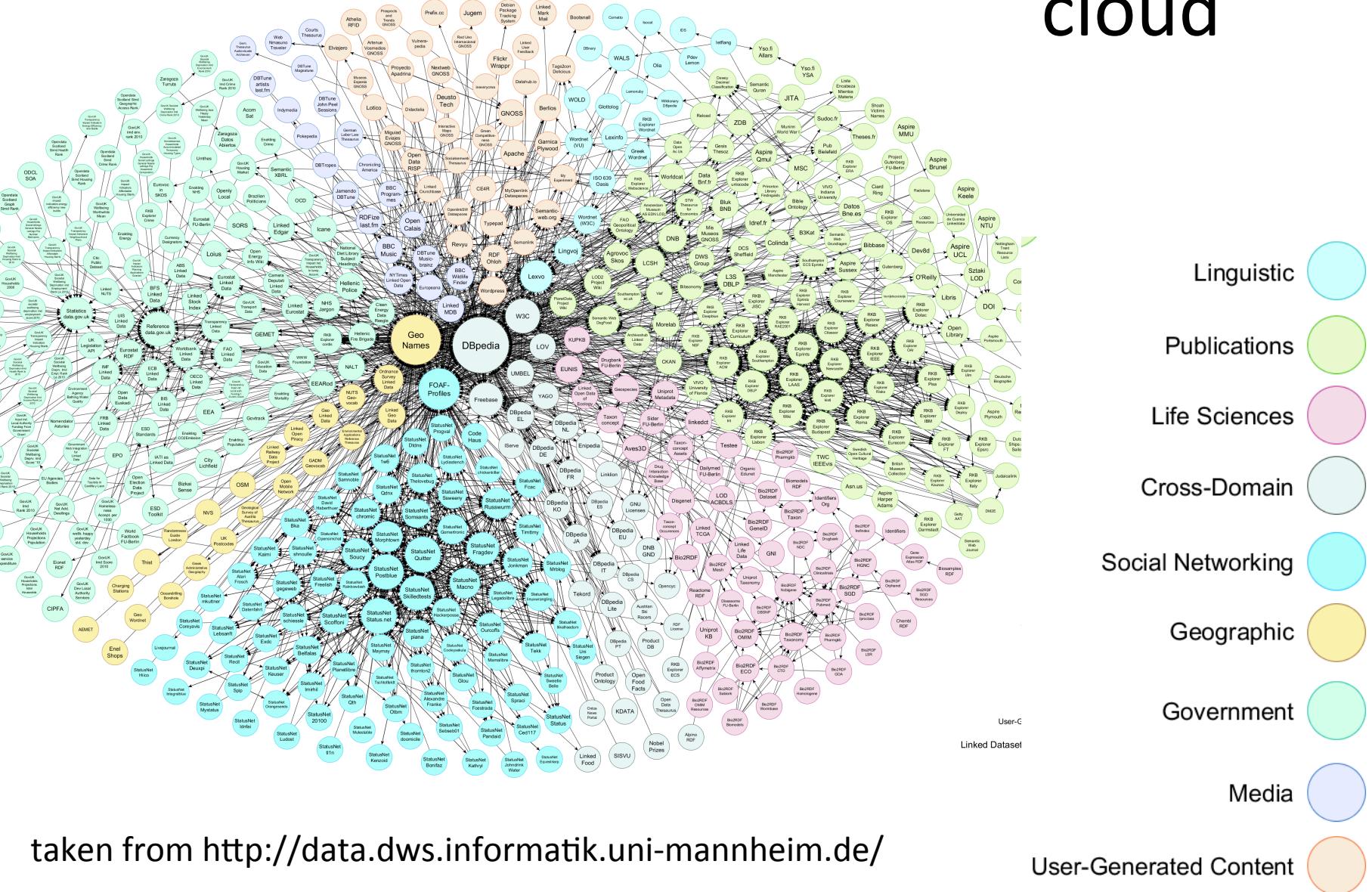


# **QwwwQ: Querying Wikipedia without writing queries**

Massimiliano Battan and Marco Ronchetti  
DISI, Università degli Studi di Trento  
38123 Povo di Trento  
Italy

# Linked Open Data more and more public admins join!

# The LOD cloud



Sometimes, too much data = no data

Christian Wolmar, railway historian:

*"The huge **corruption** rate in the 19<sup>th</sup> century in USA railway administration **was facilitated by an excessive level of accountability**, where expense documentation for every single spike were collected".*

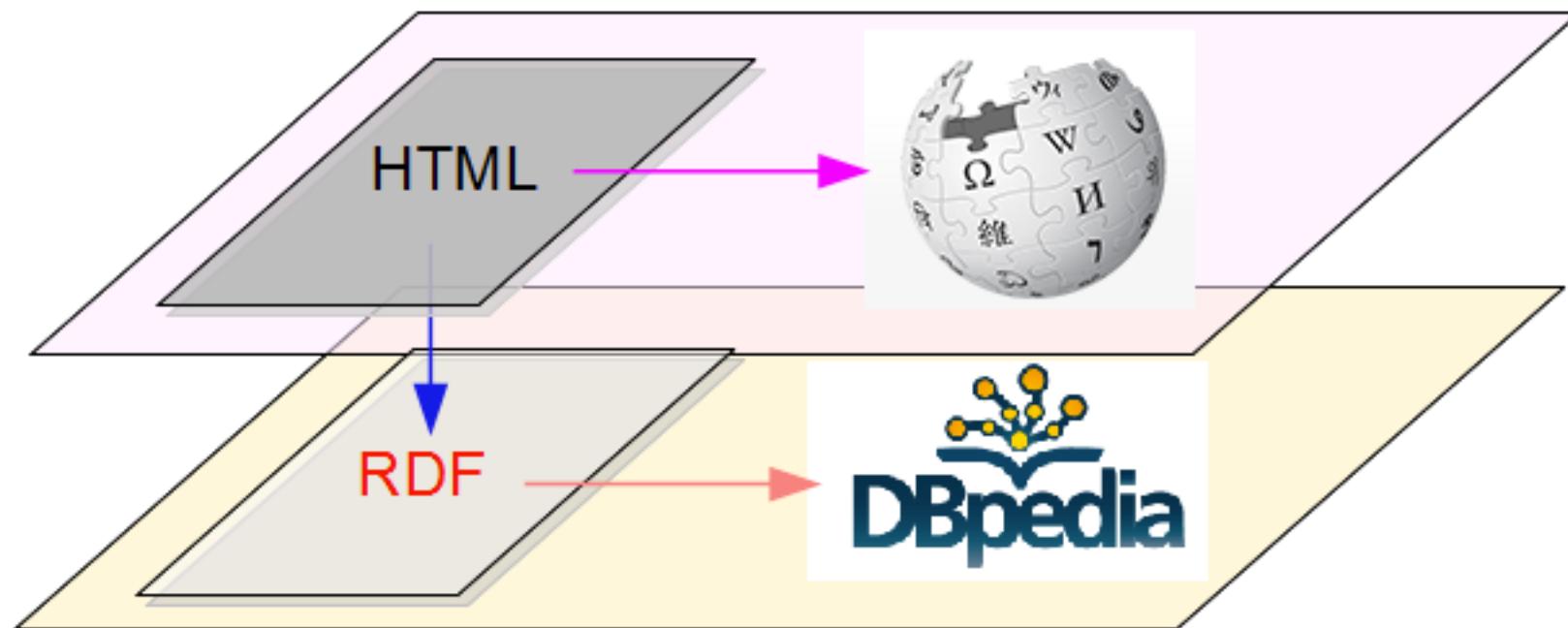
# Data is only the first step...

- Data -> Information -> Knowledge
  - Data Mining, Search Engines, Recommendation...

Make sense of the data,  
extract meaningful information  
**in simple and accessible way**

**user + data**

# DBpedia



S. Auer, and J. Lehmann, "What have Innsbruck and Leipzig in common?  
Extracting semantics from wiki content." 2007

S. Auer et al. "Dbpedia: A nucleus for a web of open data." 2007

# A wikipedia page...

 **WIKIPEDIA**  
The Free Encyclopedia

Main page  
Contents  
Featured content  
Current events  
Random article  
Donate to Wikipedia  
Wikimedia Shop

Interaction  
Help  
About Wikipedia  
Community portal  
Recent changes  
Contact page

Tools  
What links here  
Related changes  
Upload file  
Special pages  
Permanent link  
Page information  
Wikidata item  
Cite this page

Print/export  
Create a book  
Download as PDF  
Printable version

Languages   
Deutsch  
Français  
Italiano  
Magyar  
Edit links

## FS Class E.444

From Wikipedia, the free encyclopedia

The **FS E.444** is a class of [Italian railways](#) electric locomotives. They were introduced in the course of the 1960 until 1975. Starting from 1995, all E.444s were upgraded as E.444R.

The locomotives are nicknamed *Tartaruga* (turtle).

**Contents** [hide]

- 1 History
  - 1.1 E.444 standard
  - 1.2 E.444R
- 2 Footnotes
- 3 External links

### History [edit]

#### E.444 standard [edit]

The E.444 locomotive was designed in the 1960s as the first Italian electric locomotive capable to reach 200 km/h (120 mph) (in that period first high-speed trains like the Japanese [Shinkansen](#) and the French [TGV](#) were appearing). Italian railways could boast fast trains like the [ETR 200](#), but they were getting old and the [Pendolino](#) project was just moving its first phases.

The first 4 prototypes, built at [Savigliano](#), made their debut in 1967-1968: their power output 3,000 kW (4,000 hp) was respectable for the time, but they proved unable to hold the fast international services required for the new locomotive, and the bogies were limited to 180 km/h (110 mph). They proved anyway that Italian industry could produce locomotive capable of more than 200 km/h (120 mph)<sup>[3]</sup> was registered on November 8, 1967 in the maiden trip Rome-Milan.

The series production saw the introduction of the more powerful T750 motors, which boosted the power to 4,200 kW (5,600 hp), while the bogies were upgraded for 200 km/h (120 mph). The frontal part was improved and made more aerodynamic. A characteristic livery with two blue stripes on a pale grey background was adopted. The first 50 units proved successful, and the [Ferrovie dello Stato](#) (FS) ordered 60 more to be built starting from 1972. 16 units were adapted for feeding at 1,5 kV DC. During 1974 two locomotives (units 056 and 057) were provided with a "shunt chopper" system (later the modifications applied



FS class E.444

An E.444 in Bari Centrale station in 1995.

Specifications	
Power type	Electric
Build date	1965-1975
UIC classification	Bo-Bo
Wheel diameter	1,250 m (49.21 in)
Wheelbase	9,000 m (29 ft 6.3 in) between bogies 3,020 m (9 ft 10.9 in) between axles on each bogie
Length	16,840 m (55 ft 3.0 in)
Width	3,020 m (9 ft 10.9 in)
Height	4,300 m (14 ft 1.3 in)
Locomotive weight	83 t (82 long tons; 91 short tons)
Electric system(s)	3,000 V DC Catenary
Current collection method	Pantograph
Traction motors	DC series
Transmission	41/77 gear ratio
Top speed	200 km/h (120 mph)
Power output	4,272 kW (5,729 hp)
Tractive effort	201 kN (45,000 lb <sub>f</sub> )
Safety systems	RSC4; SCMT
Career	
Railroad(s)	FS Trenitalia

# DBpedia

From [dbpedia.org/about](http://dbpedia.org/about):

"DBpedia allows you to ask sophisticated queries against Wikipedia, and to link the different data sets on the Web to Wikipedia data. We hope that this work will make it easier for the huge amount of information in Wikipedia to be used in some new interesting ways.

Furthermore, it might inspire new mechanisms for navigating, linking, and improving the encyclopedia itself."

# Querying DBpedia

Using RDF, you can formulate powerful queries

Which European countries have a capital with  
more than 3 million people in which flows a  
river longer than 300 km?

## The problem:

How can a generic user express a query on DBpedia?

S(he) needs to know the schema...

S(he) needs to know a query language...

# Proposed solutions



[DBpedia Blog](#) | [Get Involved](#) | [Get Help](#)

OpenLink Virtuoso  
RelFinder  
Lodlive  
gFacet  
DBpedia Mobile  
Faceted Wikipedia Search (FWS)

## DBpedia Applications

This page lists a number of applications (in no particular order) to get you started using DBpedia:

### Contents

- [Faceted Browsers](#)
- [User Applications](#)
- [Query Results Visualization](#)
- [URI Lookup Services](#)
- [Query Builders](#)
- [SPARQL query interfaces](#)
- [Browser enhancements](#)
- [Annotation and/or Information Extraction](#)
- [Natural Language Processing \(NLP\) Services](#)

[About / News](#)  
[Applications](#)  
[Use Cases](#)  
[Datasets](#)  
[Online Access](#)  
[DBpedia Live](#)  
[Downloads](#)  
[Interlinking](#)  
[Development](#)

# iSPARQL

running  
queries

iSPARQL File Help Login

QBE Advanced Results

Graph Named Graphs (0) http://dbpedia.org Clear

SPARQL Query -- Recent Queries -- -- Prefixes -- -- Template -- -- Statement Help --

```
select distinct ?Concept where {[] a ?Concept} LIMIT 100
```

Query options

Result size limit: 50 rows Leave empty for server maximum setting.

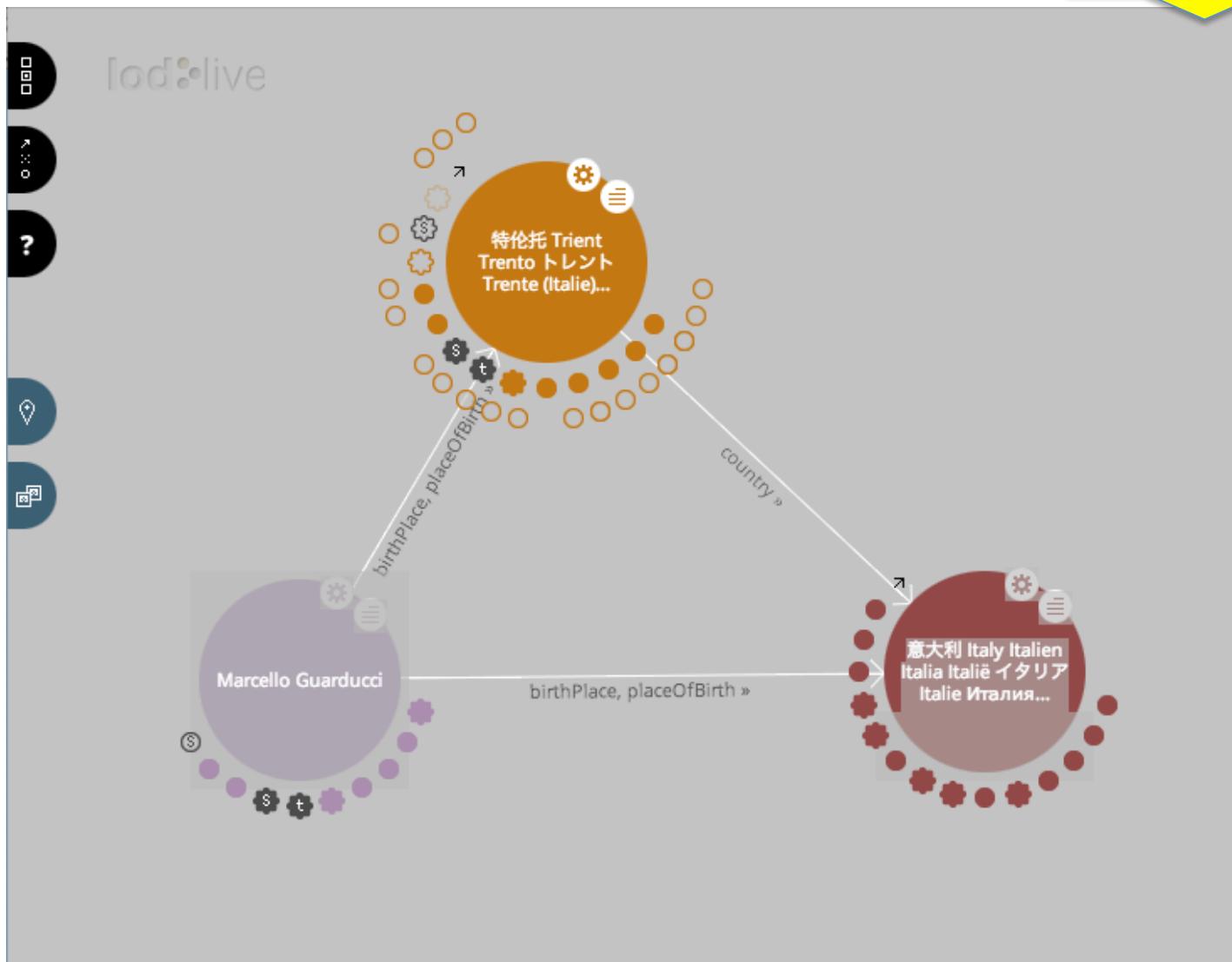
Query timeout: msec - leave blank for server default, or when not querying a Virtuoso endpoint.

▶ Sponger

▶ Query Metadata

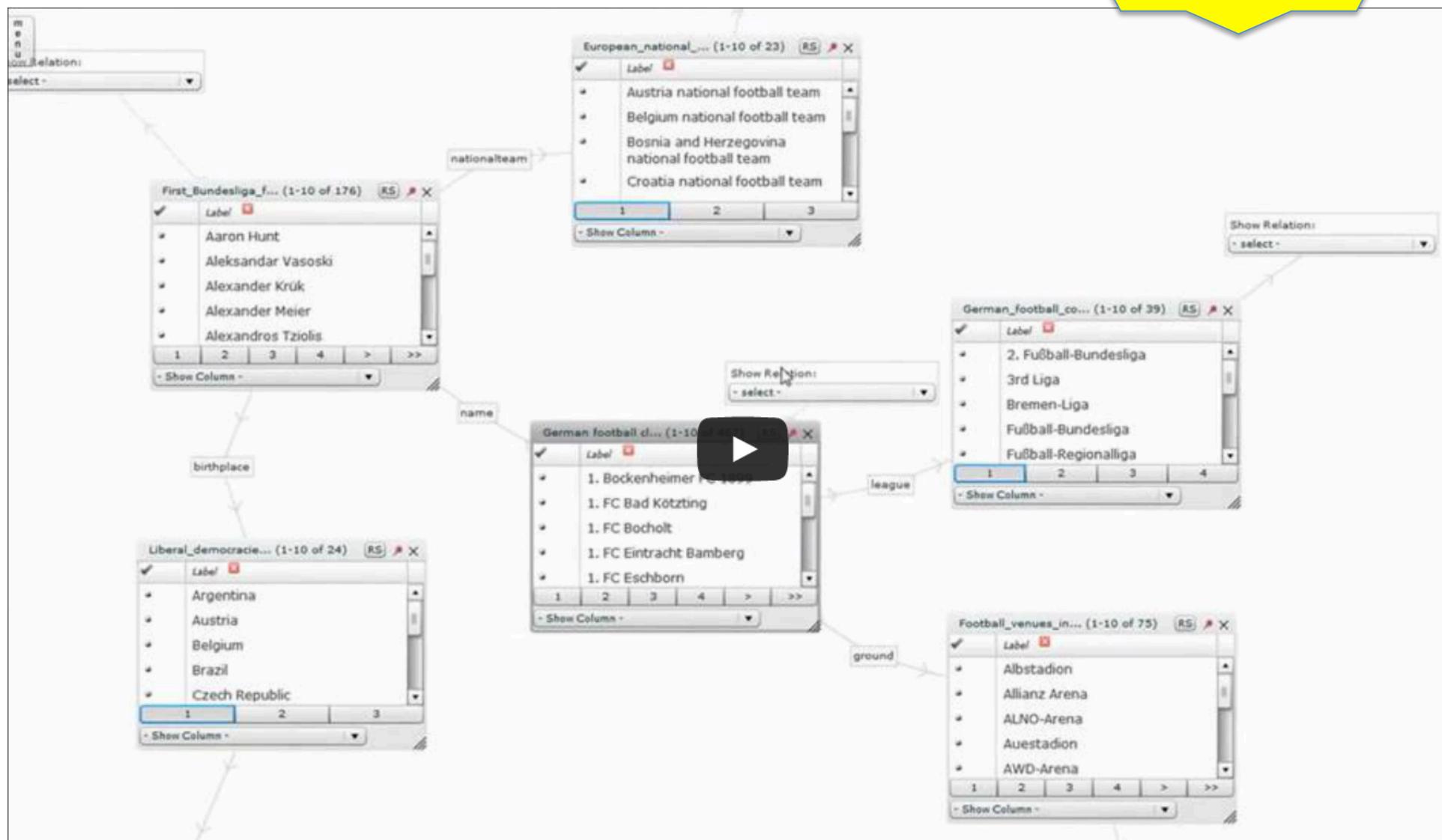
# lodlive

linked data  
browsers



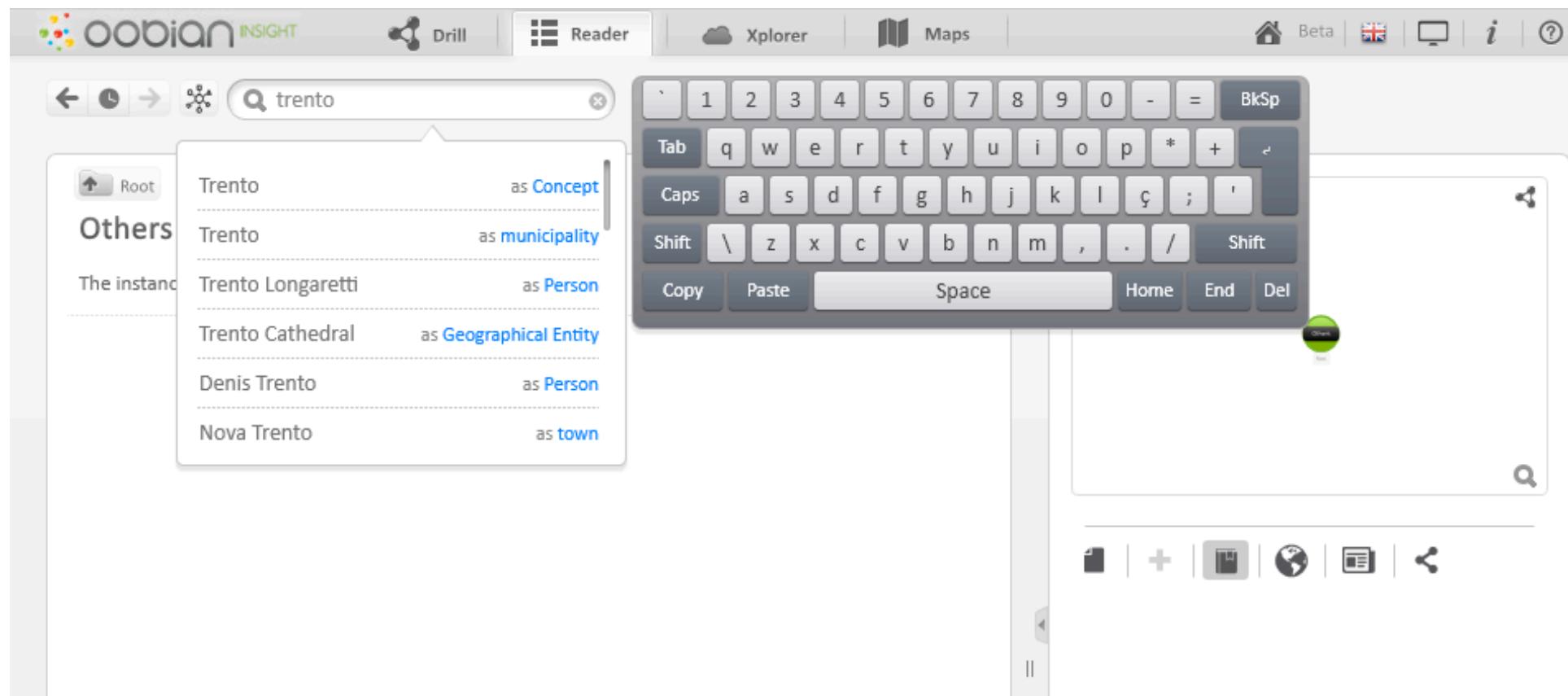
# gfacet

exploring data  
and relations



# oobian

focusing on  
instances



The screenshot shows the oobian INSIGHT web application interface. At the top, there is a navigation bar with tabs: Drill, Reader, Xplorer, and Maps. Below the navigation bar is a search bar containing the text 'trento'. A floating keyboard is overlaid on the interface, centered over the search bar. The main content area displays a list of search results for 'trento', categorized under 'Others'. The results are as follows:

- Trento as Concept
- Trento as municipality
- Trento Longaretti as Person
- Trento Cathedral as Geographical Entity
- Denis Trento as Person
- Nova Trento as town

At the bottom of the interface, there is a toolbar with various icons for file operations, including a file icon, a plus sign, a save icon, a globe, a map, and a share icon.

# oobian

◀ ▶ ⚙ 🔍

municipality ▾ ▴ municipality

## Trento

Trento is an Italian city located in the Adige River valley in Trentino-Alto Adige/Südtirol. It is the capital of Trentino. In the 16th century the city was the location of the Council of Trent. Trento is a major educational, scientific, financial and political centre in Trentino-Alto Adige/Südtirol and Northern Italy in general.

### URI

☞ <http://dbpedia.org/resource/Trento>

### area code

0461

### area total (m2)

1.579E8

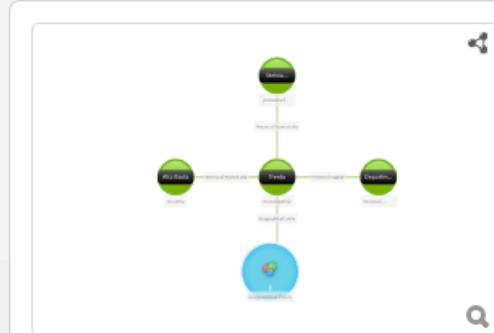
### elevation (μ)

190.0

### grss:point

46.06666666666667 11.116666666666667

Ver Relacionados



File + Print 🌐 More

### More Like This

 Italy as country

 Trentino as municipality

 Carè Alto as mountain

 University of Trento as university

 Gianni Caproni Museum of Aeronautics as museum

# What do we really need?

Get closer to the final user!

# Our proposal: QwwwQ

*Get me something like this!*

- Query by example
- Zloof 1975, relational data model

TYPE	ITEM	COLOR	SIZE
	Pen	Green	
→	Lipstick	Red	
	Pen	Blue	
→	Pen	Red	

The diagram illustrates a query example using a relational database table. The table has four columns: TYPE, ITEM, COLOR, and SIZE. The rows represent items in the database. A red arrow points to the second row (Lipstick). Three blue arrows point to the third row (Pen, Blue) and the fourth row (Pen, Red).

ITEM=P.PEN COLOR=RED -> {Lipstick, Pen}

ITEM=PEN COLOR=P.RED -> {Green, Red, Blue }

# *Get me something like this!*

User starts from a wikipedia page, and invokes a Chrome plug-in

- Infobox provides the schema
- User selects the variables of interest
- Add filtering and based on wikipedia categories, add ordering on parameters

# Query on Berlin

Infobox: Infobox German state

Ask me!

population	> 5000000
pop_date	= 2013-12-01
elevation	= 34
population_demonym	= Berliner
GDP	= 109.2
GDP_year	= 2013
Website	= <a href="http://www.berlin.de/internat">http://www.berlin.de/internat</a>
leader_title	= Governing Mayor
leader	= Klaus Wowereit
leader_party	= SPD
ruling_party1	= SPD

Something found! Your results (2):

Name	DBpedia page	Wikipedia page	population
Lower Saxony	<a href="#">DBpedia</a>	<a href="#">Wikipedia</a>	7977000
North Rhine-Westphalia	<a href="#">DBpedia</a>	<a href="#">Wikipedia</a>	17920000



Location within European Union and Germany  
Coordinates: 52°31'N 13°23'E

Country	Germany
Government	<ul style="list-style-type: none"><li>Governing Mayor Michael Müller (SPD)</li><li>Governing parties SPD / CDU</li><li>Votes in Bundesrat 4 (of 69)</li></ul>
Area	<ul style="list-style-type: none"><li>City 891.85 km<sup>2</sup> (344.35 sq mi)</li></ul>
Elevation	34 m (112 ft)
Population (December 2014) <sup>[1]</sup>	<ul style="list-style-type: none"><li>City 3,562,166</li><li>Density 4,000/km<sup>2</sup> (10,000/sq mi)</li></ul>
Demonym(s)	Berliner
Time zone	CET (UTC+1)
Summer (DST)	CEST (UTC+2)
Postal code(s)	10115–14199
Area code(s)	030
ISO 3166 code	DE-BE
Vehicle registration	B <sup>[2]</sup>
GDP/ Nominal	€117.2 billion (2014) <sup>[3]</sup>
GDP per capita	€31,500 (2014)
NUTS Region	DE3
Website	<a href="http://berlin.de">berlin.de</a>

Output: DBpedia page, Wikipedia page

# *Get me something like this!*

User starts from a wikipedia page, and invokes a Chrome plug-in

- Infobox provides the schema
- User selects the variables of interest
- Add filtering and based on wikipedia categories, add ordering on parameters

# Queen Elizabeth I, House=House of Tudor, Religion=anything

Something found! Your results (5):

✓	Name	DBpedia page	Wikipedia page	Religion	☰
✓	Arthur, Prince of Wales			Catholic_Church	
✓	Edward VI of England			Church_of_England	
✓	Elizabeth I of England			Anglicanism	
✓	Mary I of England			Catholic_Church	
✓	Mary Tudor, Queen of France			Catholic_Church	

Similarity class <http://dbpedia.org/ontology>

Filter by a Wikipedia category 16th-century women

Order By Religion

Advanced research

Something found! Your results (3):

✓	Name	DBpedia page	Wikipedia page	Religion	☰
✓	Elizabeth I of England			Anglicanism	
✓	Mary Tudor, Queen of France			Catholic_Church	
✓	Mary I of England			Catholic_Church	

exportable to csv, pdf

# How many UK defunct political parties have green among their colors?

The screenshot shows a semantic search interface with the following components:

- Infobox: Infobox political party** (Header): A search bar with the placeholder "Ask me!"
- Settings** (Header): Includes checkboxes for "Include current page in the result" (checked), "Max. results" (set to 9950), "Similarity class" (set to <http://dbpedia.org/ontology>), "Filter by a Wikipedia category" (set to "Defunct political parties in United Kingdom"), and "Order By" (set to "Choose an optional order").
- Advanced research** (Header): A blue bar with a diagonal striped pattern.
- Results:** A table titled "Something found! Your results (1):" showing one result for the "Islamic Party of Britain". The table columns are: Name, DBpedia page, Wikipedia page, and colours. The result row shows: Name (Islamic Party of Britain), DBpedia page (link to DBpedia), Wikipedia page (link to Wikipedia), and colours (Green).

# Implementation

- core library:
  - wikipedia parser module (local copy of wp)
  - sparql query builder
- restful web services
  - callable via Ajax
- client code:
  - AngularJS
- browser embedding
  - Google Chrome extension

# Problems

- mostly wikipedia-related:
  - Wikipedia is **not complete, nor fully correct**
  - **presence** of infoboxes
    - Berlin is a town or a state?
  - Infoboxes have "**lack of standard**" problems:
    - terminology, units, or even comments
    - lists in infoboxes
  - Wikipedia **categorization**
    - incomplete: Berlin is not in the category of "University towns in Germany"
    - not even an acyclic graph!

people are good at collaborative writing,  
not at collaborative creating categorizations

# Get me something like Aprilia...

**Infobox: Azienda**

**Ask me!**

<input type="checkbox"/>	nome	 = Aprilia
<input type="checkbox"/>	logo	 = Aprilialogo.png
<input type="checkbox"/>	data_fondazione	 = 1945
<input type="checkbox"/>	luogo_fondazione	 = Noale
<input type="checkbox"/>	fondatori	 = Alberto Beggio
<input type="checkbox"/>	nazione	 = ITA
<input type="checkbox"/>	sede	 = Noale
<input type="checkbox"/>	filiali	 = Scorzè
<input checked="" type="checkbox"/>	gruppo	 = Piaggio
<input type="checkbox"/>	persone_chiave	 = Roberto Colaninno
<input type="checkbox"/>	industria	 = Motoveicoli
<input type="checkbox"/>	prodotti	 = Scooter
<input type="checkbox"/>	slogan	 = #be a racer
<input type="checkbox"/>	sito	 = www.aprilia.com

**Settings** 

**Include current page in the result**

**Max. results**

**Similarity class**

**Filter by a Wikipedia category**

**Order By**

**Advanced research**

**Infobox: Azienda**

**Ask me!**

<input type="checkbox"/>	nome	 = Aprilia
<input type="checkbox"/>	logo	 = Aprilialogo.png
<input type="checkbox"/>	data_fondazione	 = 1945
<input type="checkbox"/>	luogo_fondazione	 = Noale
<input type="checkbox"/>	fondatori	 = Alberto Beggio
<input type="checkbox"/>	nazione	 = ITA
<input type="checkbox"/>	sede	 = Noale
<input type="checkbox"/>	filiali	 = Scorzè
<input type="checkbox"/>	gruppo	 = Piaggio
<input checked="" type="checkbox"/>	persone_chiave	 = Roberto Colaninno
<input checked="" type="checkbox"/>	industria	 = Motoveicoli
<input type="checkbox"/>	prodotti	 = Scooter
<input type="checkbox"/>	slogan	 = #be a racer
<input type="checkbox"/>	sito	 = www.aprilia.com

**Settings** 

**Include current page in the result**

**Max. results**

**Similarity class**

**Filter by a Wikipedia category**

**Order By**

**Advanced research**

**Something found! Your results (5):**

Name	DBpedia page	Wikipedia page	≡
Aprilia (azienda)			
Derbi			
Gilera			
Moto Guzzi			
Moto Laverda			

**Something found! Your results (5):**

Name	DBpedia page	Wikipedia page	≡
Aprilia (azienda)			
FTR Moto			
Harris Performance Products			
Kymco			
Suter Racing Technology			

# Inconsistencies..

**Gilera & C. SpA**



**GILERA®**

**Stato**  Italia

**Tipo** Società per azioni

**Fondazione** 1909 a Milano in Corso XXII Marzo

**Sede principale** Pontedera

**Gruppo** Piaggio

**Settore** Casa motocistica

**Prodotti** motocicli

**Slogan** «Nata dal desiderio»

**Sito web** [www.it.gilera.com](http://www.it.gilera.com)

Modifica dati su Wikidata · Manuale

**Aprilia**



**Stato**  Italia

**Fondazione** 1945 a Noale

**Fondata da** Alberto Beggio

**Sede principale** Noale

**Gruppo** Piaggio

**Filiali** Scorzè

**Persone chiave** Roberto Colaninno (presidente)

**Settore** Motoveicoli

**Prodotti** Motociclette  
Scooter

**Slogan** «#be a racer»

**Sito web** [www.aprilia.com](http://www.aprilia.com)

Modifica dati su Wikidata · Manuale

# Which Italian F1 drivers became World Champions?

Ask me!

<input type="checkbox"/>	<b>name</b>	<a href="#">Michael Schumacher</a>
<input type="checkbox"/>	<b>image</b>	<a href="#">Michael Schumacher-I'm the man .</a>
<input type="checkbox"/>	<b>caption</b>	<a href="#">Michael Schumacher in</a>
<input type="checkbox"/>	<b>birth_place</b>	<a href="#">Italy</a>
<input checked="" type="checkbox"/>	<b>nationality</b>	<a href="#">Italian</a>
<input type="checkbox"/>	<b>Years</b>	<a href="#">-,-</a>
<input type="checkbox"/>	<b>Team(s)</b>	<a href="#">Ferrari</a>
<input type="checkbox"/>	<b>2012 Team</b>	<a href="#">Mercedes</a>
<input type="checkbox"/>	<b>2012 Car number</b>	<a href="#">7</a>
<input type="checkbox"/>	<b>Races</b>	<a href="#">=</a>
<input checked="" type="checkbox"/>	<b>Championships</b>	<a href="#">&gt; 0</a>
<input type="checkbox"/>	<b>First race</b>	<a href="#">1991 Belgian Grand Prix</a>
<input type="checkbox"/>	<b>First win</b>	<a href="#">1992 Belgian Grand Prix</a>
<input type="checkbox"/>	<b>Last win</b>	<a href="#">2006 Chinese Grand Prix</a>
<input type="checkbox"/>	<b>Last race</b>	<a href="#">2012 Brazilian Grand Prix</a>

Include current page in the result

Max. results  9950

Similarity class <http://dbpedia.org/ontology/FormulaOne>

Filter by a Wikipedia category [Choose a category](#)

Order By [Choose an optional order](#)

Infobox: Navboxes

Infobox: Persondata

Advanced research

Something found! Your results (3):

Name	DBpedia page	Wikipedia page	Championships
Alberto Ascari	<a href="#">DBpedia</a>	<a href="#">Wikipedia</a>	2
Cosimo Aldo Cannone	<a href="#">DBpedia</a>	<a href="#">Wikipedia</a>	2
Giuseppe Farina	<a href="#">DBpedia</a>	<a href="#">Wikipedia</a>	1

Answer present in wikipedia

Spotted a wikipedia error!

<input type="checkbox"/>	name	Cosimo Aldo Cannone
<input type="checkbox"/>	image	Principe_Cosimo.jpg
<input type="checkbox"/>	caption	Prince Cosimo of Macedonia
<input checked="" type="checkbox"/>	nationality	Italian
<input type="checkbox"/>	2010 Team	Cannone
<input type="checkbox"/>	Boat number	45
<input type="checkbox"/>	Races	27
<input type="checkbox"/>	Championships	> 0
<input checked="" type="checkbox"/>	Wins	Rn[1,3]
<input checked="" type="checkbox"/>	Podiums	> -1
<input type="checkbox"/>	Fastest laps	22
<input type="checkbox"/>	First race	2003 Jesolo Grand Prix
<input type="checkbox"/>	First win	2004 Palermo Grand Prix
<input type="checkbox"/>	Last win	2009 Como Grand Prix
<input type="checkbox"/>	Last season	2010
<input type="checkbox"/>	Last position	2

result

Max. results: 9950

Similarity class: <http://dbpedia.org/ontology/FormulaOneRacer>

Filter by a Wikipedia category: Choose a category

Order By: Podiums

more exotic query

Italian F1 drivers with  
1 to 3 victories,  
ordered by podiums

Infobox: Persondata

Advanced research

Something found! Your results (11):

	Name	DBpedia page	Wikipedia page	Podiums
<input checked="" type="checkbox"/>	Ludovico Scarfiotti			1
<input checked="" type="checkbox"/>	Gianni Morbidelli			1
<input checked="" type="checkbox"/>	Giancarlo Baghetti			1
<input checked="" type="checkbox"/>	Vittorio Brambilla			1
<input checked="" type="checkbox"/>	Piero Taruffi			5
<input checked="" type="checkbox"/>	Luigi Fagioli			6
<input checked="" type="checkbox"/>	Luigi Musso			7
<input checked="" type="checkbox"/>	Lorenzo Bandini			8
<input checked="" type="checkbox"/>	Alessandro Nannini			9
<input checked="" type="checkbox"/>	Elio de Angelis			9
<input checked="" type="checkbox"/>	Giancarlo Fisichella			19

# Problems are non only ours...

- Same problems affect DBpedia, and hence LOD...
- Some effort could be done to mitigate them
  - e.g., recognize unusable properties

# **Spacetime: a Two Dimensions Search and Visualisation Engine Based on Linked Data**

Fabio Valsecchi and **Marco Ronchetti**  
DISI, Università degli Studi di Trento  
Povo di Trento, Italy

# Our proposal: Spacetime

- WHAT? Set of DBpedia entities having spatial and temporal data

*(Organisation, Person, Event, Place, Species and Work)*

- WHERE? Attributes having geocoordinates
- WHEN? Attributes having data type *xsd:date*.

# Formulating queries

# spacetime



[Search](#)
[My maps](#)
[Demo](#)

**1** Select a category:

- [A populated place](#)
- + [All the Populated place](#)
- + [Administrative region](#)
- + [Country](#)
- + [Island](#)
- + [Settlement](#)
- + [Protected area](#)
- + [Sport facility](#)

**2** Where:  

filter:

- country
- district
- region
- department
- state
- neighboring municipality
- part
- arrondissement

from  to

**3** When:

filter:

- population as of
- founding date
- dissolution date
- day
- birth date

Options:     

Search 

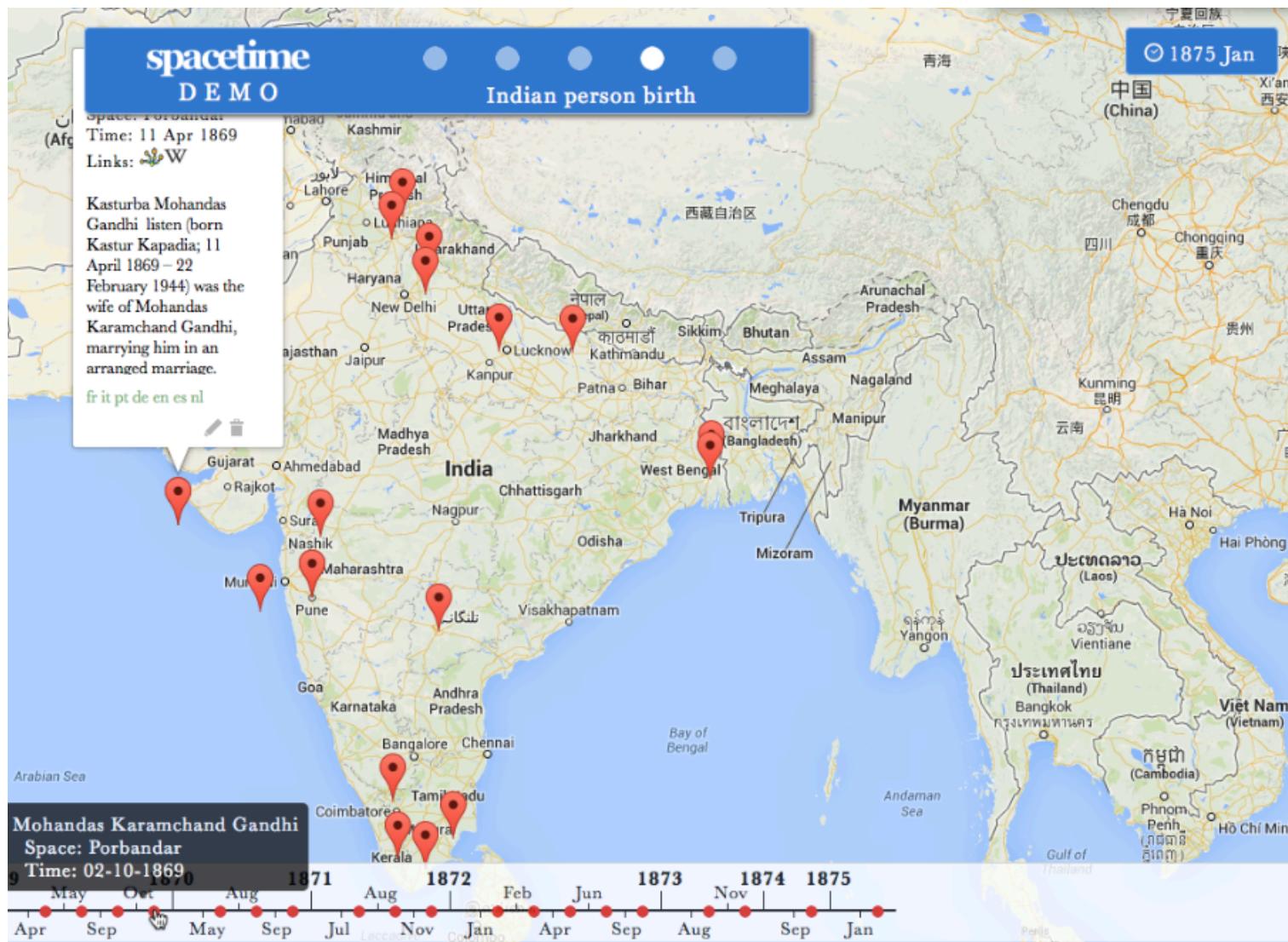


# Geographical approximation

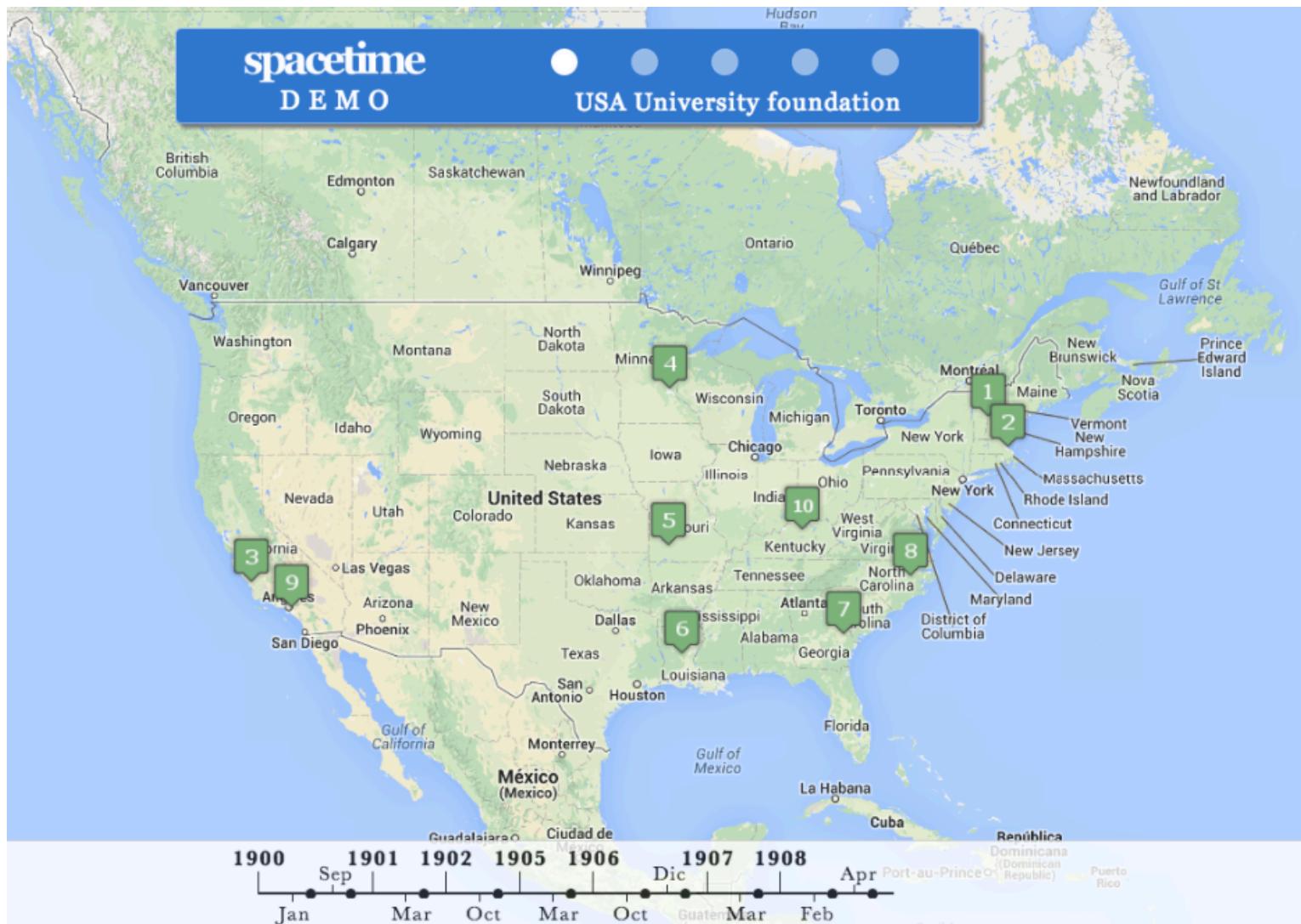
- Bounding box of the geoentity



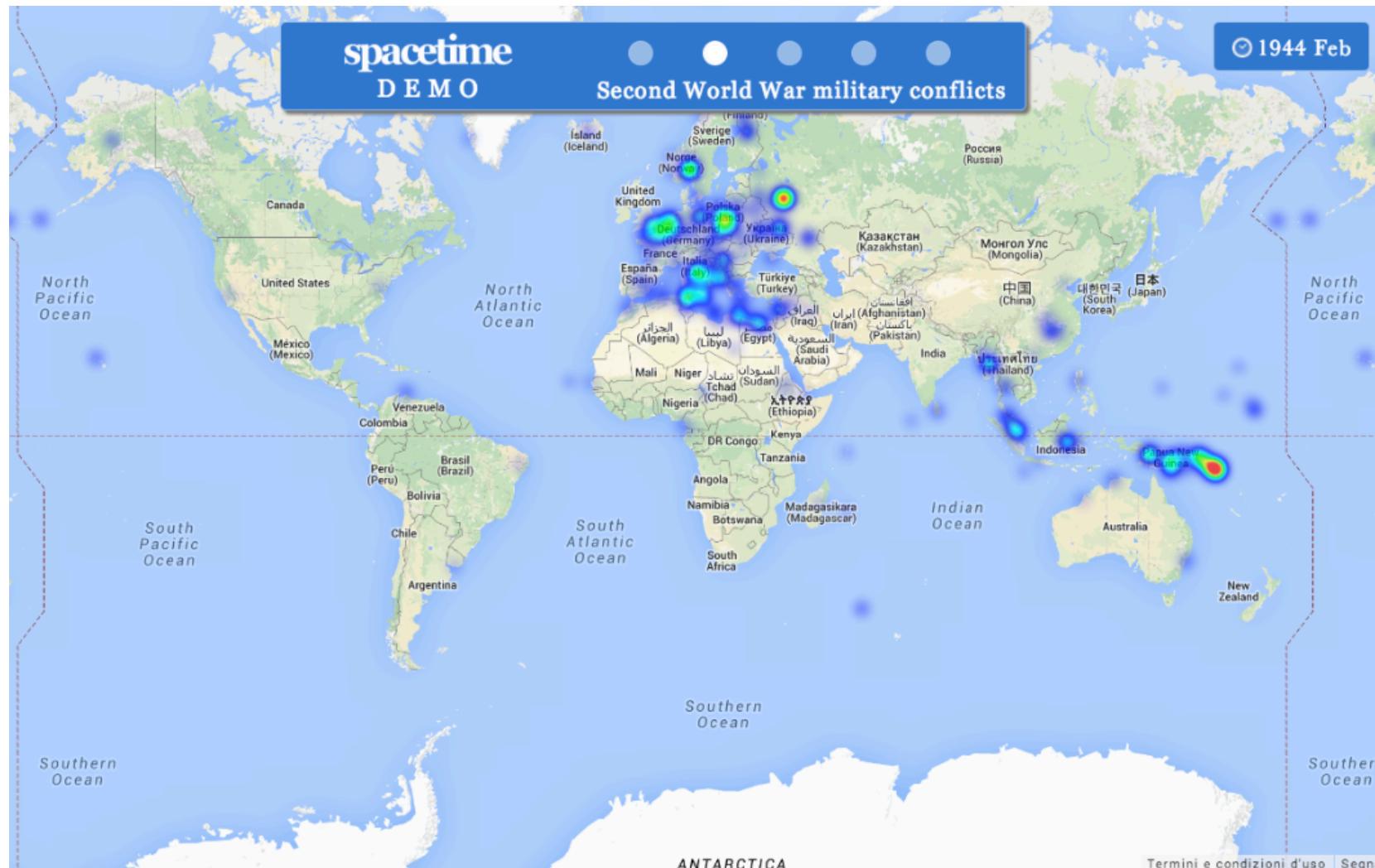
# Query results



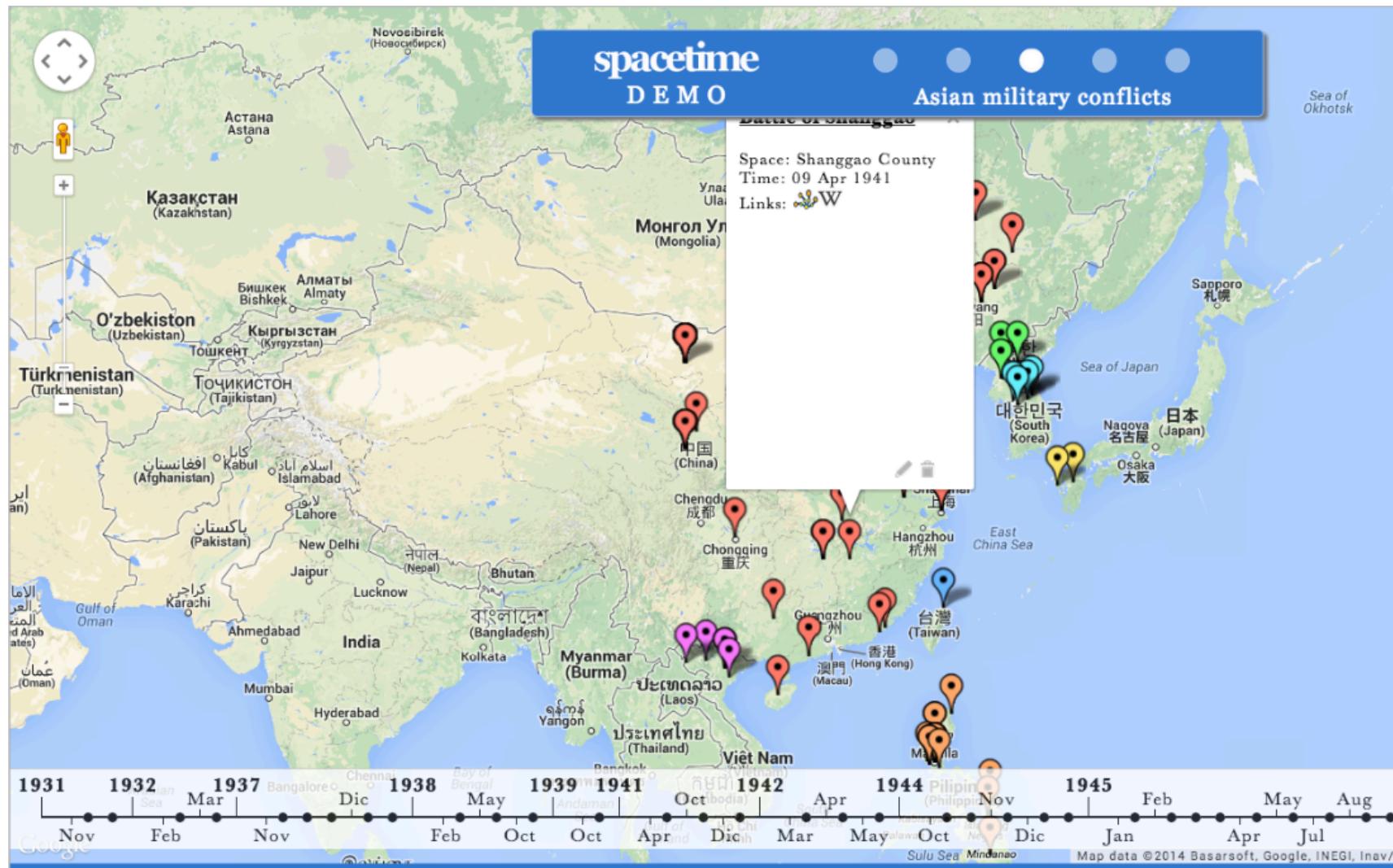
# Sequence of results



# Density maps



# Grouping different queries



# Creating a story



# Technologies

- SPARQL: Queries are composed by the Javascript engine, and are executed through the SPARQL endpoint;
- JSON: the results of the SPARQL queries are returned as JSON strings;
- Google Maps JavaScript API v.3: the Google Maps API used for populating a map with the data extracted in the JSON file
- JavaScript and JQuery library: the scripting language and its library define a set of functions that are the core of the application JQuery allows the creation of animations inside Spacetime;
- AJAX: this technology is used to have a responsive user interface compliant with the Rich Internet Application paradigm;
- CSS: for designing the graphical aspect of Spacetime;
- HTML5: used for developing certain parts of the application, such as the map saving operation, and some graphical feature.

# Conclusion

- The system is up and running
  - (google for Timespace DBpedia)
- Included in the DBpedia front end demos  
<http://wiki.dbpedia.org/Applications>
- No user assessment done (yet)

