



Querying the Wikidata Knowledge Graph

Peter Haase

LDBC Meeting, Barcelona

19.3.2015

metaphacts.

metaphacts GmbH

- Founded 10/2014
- Currently team of four
- Based in Walldorf
- Partnership with Systap

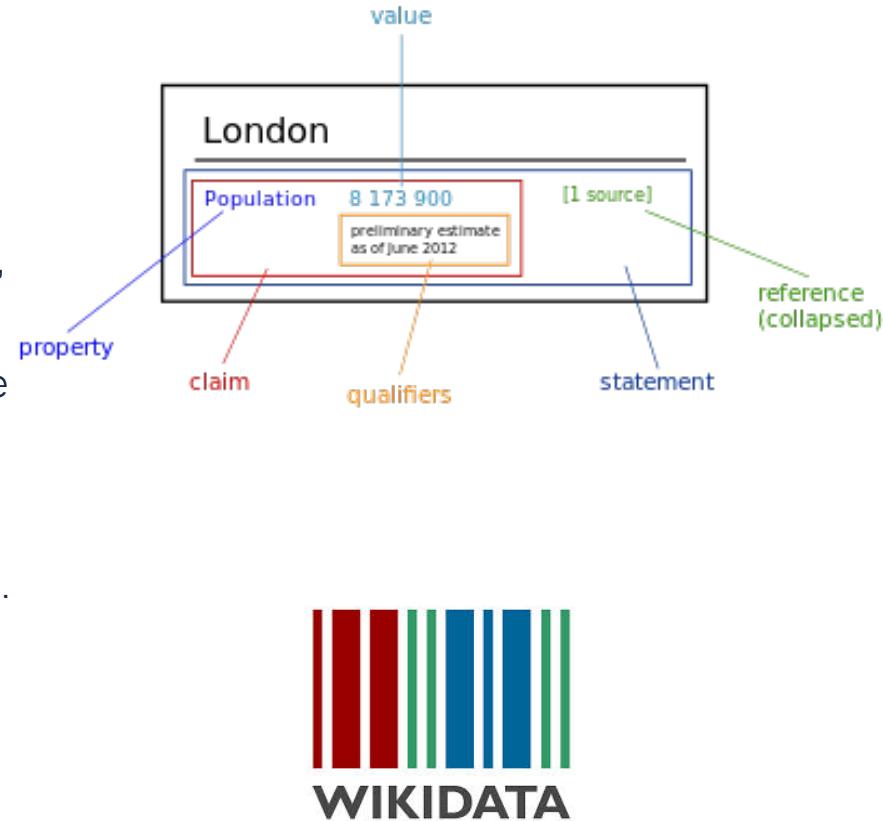


Portfolio: Software, solutions, services for knowledge graphs

- Storing and querying of knowledge graphs
 - Scalable databases for big graphs, building on bigdata/Blazegraph
 - High-performance graph analytics, based on MapGraph / GAS
 - Light-weight reasoning with large-scale knowledge graphs
- Creation and curation of knowledge graphs
 - Semi-automatic creation of knowledge graphs from existing sources
 - Data integration and ontology-based data access
 - Collaborative management of knowledge graphs
- Application development utilizing knowledge graphs
 - Rapid development of end-user oriented applications
 - Visualization of knowledge graphs, semantic search
 - Mobile applications & augmented and virtual reality

Wikidata

- **Collecting structured data.** Unlike the Wikipedias, which produce encyclopedic articles, Wikidata collects data, in a structured form.
- **Collaborative.** The data in Wikidata is entered and maintained by Wikidata editors, who decide on the rules of content creation and management in Wikidata
- **Integration with Wikipedia:** serving language labels and data
- **Multilingual.** Editing, consuming, browsing, and reusing the data is fully multilingual. Data entered in any language is immediately available in all other languages.
- **A secondary database.** Wikidata can record not just statements, but also their sources, thus reflecting the diversity of knowledge available and supporting the notion of verifiability.
- **Free.** The data in Wikidata is published under the [Creative Commons](#)



Wikidata in RDF

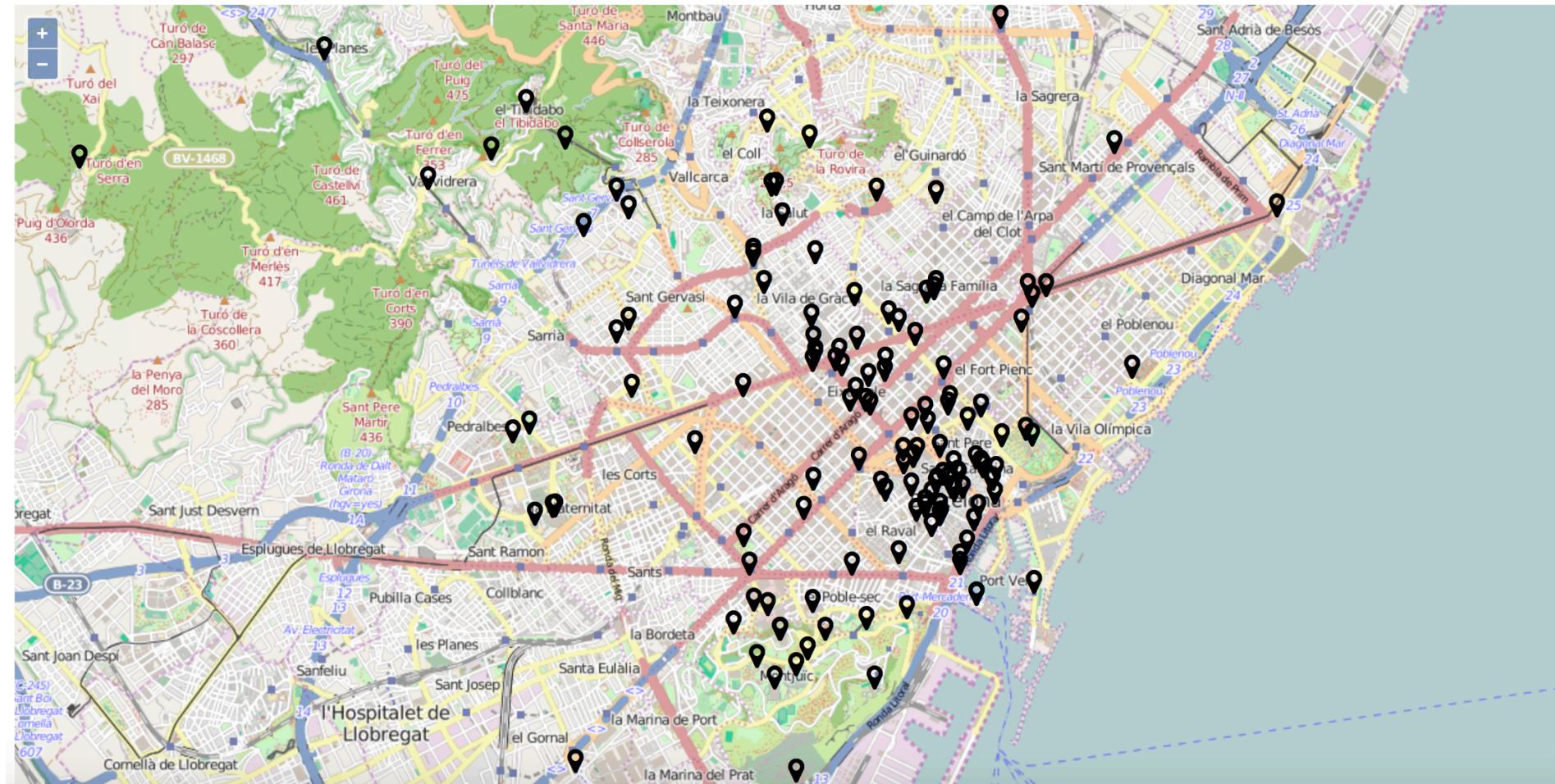
- “Native” Wikidata model data model (not RDF)
 - 16 million entities
 - 34 million statements
 - 80 million labels
 - 350 languages
- RDF exports available
 - Non-trivial mapping to RDF
 - >400 million triples
- Examples:
 - <http://wikidata.metaphacts.com/>
 - <http://wikidata.metaphacts.com/sparql>

Geo/Spatial

metaphacts sparql

Logout

Entities in Barcelona



Geo/Spatial

metaphacts sparql

Logout

= Entities in Barcelona =

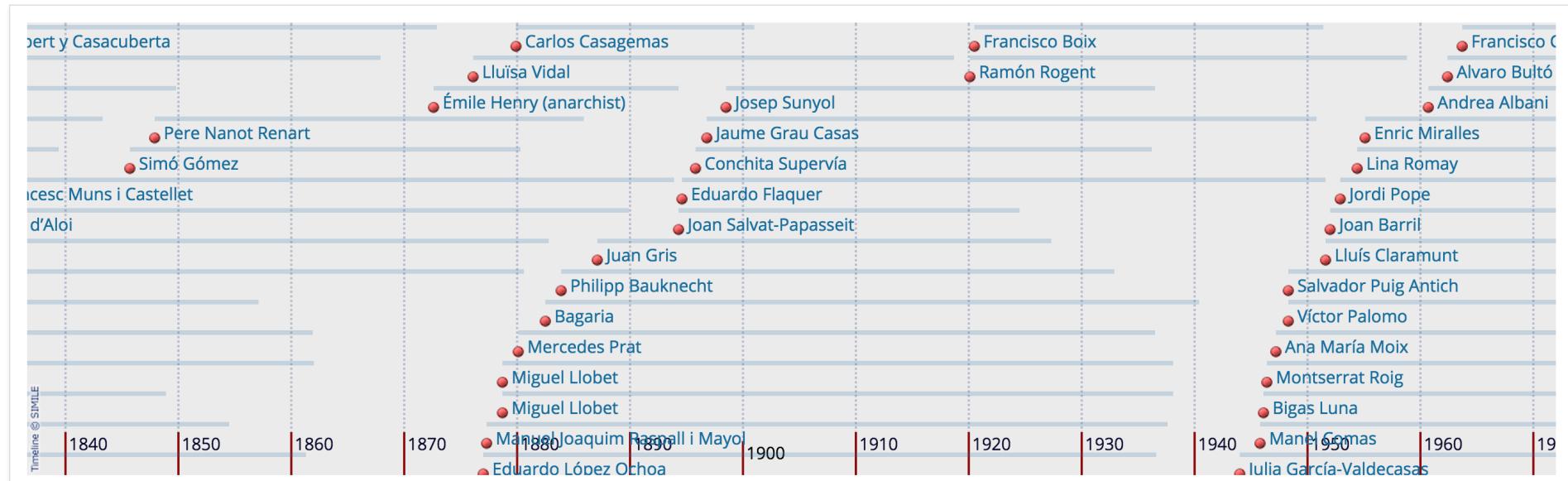
```
 {{#widget: Map | markers = $

SELECT ?lat ?lng ?link ?description
WHERE {
  ?entity wd:P131c wd:Q1492;
    wd:P625c ?coordinates;
    wd:P373c ?description .
  ?coordinates <http://www.wikidata.org/ontology#longitude> ?lng ;
    <http://www.wikidata.org/ontology#latitude> ?lat .
  BIND(?entity as ?link) .
}
$}}
```

Temporal



People born in Barcelona on a Timeline



Temporal

metaphacts sparql

Logout

= People born in Barcelona on a Timeline =

```
{#widget:Timeline
|query =
SELECT DISTINCT ?birth ?death ?label ?uri WHERE {
?uri rdfs:label ?label .
?uri wd:P569c ?birth .
?uri wd:P570c ?death .
?uri wd:P19c wd:Q1492 .
FILTER(lang(?label)='en')
}
start='birth'
end='death'
label='label'
link='uri'
interval='DECADE'
}}
```

Semantic Search

metaphacts sparql

Logout

Semantic Search

place of birth Barcelona [city] + -

occupation actor [profession] + -

Search

Maria Mercader	Loles León	Jinx Falkenburg	Maria Alba	Lola Dueñas	Astrid Berges-Frisbey	Carmen Amaya	Ariadna Gil
							
Sílvia Marsó	Juanjo Puigcorbé	Lluís Homar	Vanessa Incontrada	Santi Millán	Montserrat Carulla	Leticia Dolera	Laia Marull
							

Semantic Search

Semantic Search

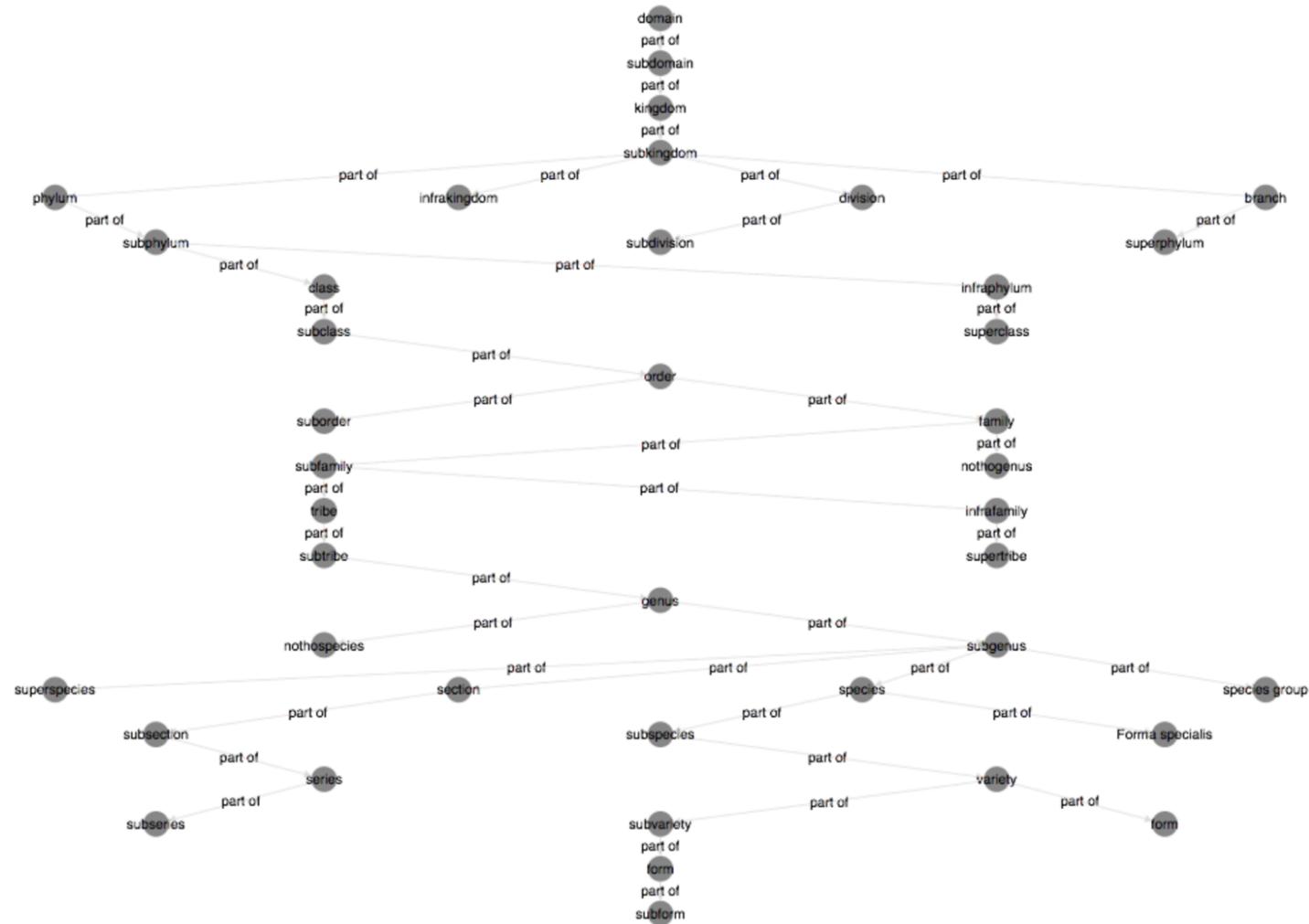
A screenshot of a semantic search interface. On the left, there is a button labeled "select relationship" with a dropdown arrow. Next to it is a search input field containing the word "barcelona". To the right of the input field is a blue "+" button. A dropdown menu is open, listing several options related to "Barcelona": "A Barcelona [poem]", "Barcelona [municipality of Spain]", "Barcelona [film]", "Barcelona [city]" (which is highlighted with a gray background), "Barcelona [village]", "Barcelona [song]", "Barcelona [administrative territorial entity of Ukraine]", "Barcelona [band]", "Barcelona [city with millions of inhabitants]", and "Barcelona [single]". The "Barcelona [city]" option is currently selected.

- A Barcelona [poem]
- Barcelona [municipality of Spain]
- Barcelona [film]
- Barcelona [city]
- Barcelona [village]
- Barcelona [song]
- Barcelona [administrative territorial entity of Ukraine]
- Barcelona [band]
- Barcelona [city with millions of inhabitants]
- Barcelona [single]

Graph algorithms

metaphacts sparql

Taxonomic Ranks (Biological classification)



Graph algorithms

metaphacts sparql

Taxonomic Ranks (Biological classification)



= Taxonomic Ranks (Biological classification) =

```
<row>
<col width="12" height="1000px">
{{#widget: com.metaphacts.widgets.GraphWidget
| query = 'PREFIX gas: <http://www.bigdata.com/rdf/gas#>
CONSTRUCT { ?pred wd:P361c ?out } WHERE {
SERVICE gas:service {
gas:program gas:gasClass "com.bigdata.rdf.graph.analytics.BFS" .
gas:program gas:in wd:Q146481 . # one or more times, specifies the initial frontier.
gas:program gas:out ?out . # will be bound to the visited vertices.
gas:program gas:out1 ?depth . # will be bound to the depth of the visited vertices.
gas:program gas:out2 ?pred . # | will be bound to the predecessor.
gas:program gas:maxVisited 2000 . # optional limit on the #of visited vertices.
gas:program gas:traversalDirection 'Reverse' .
gas:program gas:linkType wd:P361c
}
|
layout = {{
    name = 'breadthfirst' | roots = {{ 'http://www.wikidata.org/entity/Q146481' }}
}}
}}
</col>
</row>
```



Statistics

metaphacts sparql

Logout

Statistics

Most popular classes

Filter Results

class	count
Item	16842535
human	2693145
Wikimedia category page	2422875
taxon	1896799
GlobeCoordinatesValue	1847718
Wikimedia disambiguation page	805146
village-level division in China	588520
owl:Class	191385
Wikimedia list article	186733
Wikimedia template	185499

« 1 2 3 4 5 6 7 8 9 10 »

Most popular properties

Filter Results

property	count
rdf:type	31348748
instance of	12446770
label	8534489
country	2658435
sex or gender	2492158
description	2305311
located in the administrative territorial entity	2036112
taxon rank	1904288
coordinate location	1891469
taxon name	1881588

« 1 2 3 4 5 6 7 8 9 10 »

Qualified Statements and References

Douglas Adams (Q42)

English writer and humorist

[edit]

Also known as:

Douglas Noël Adams

Douglas Noel Adams

DNA

Bop Ad

[edit]

⋮

date of birth



11 March 1952

[edit]

▶ 1 reference

⋮

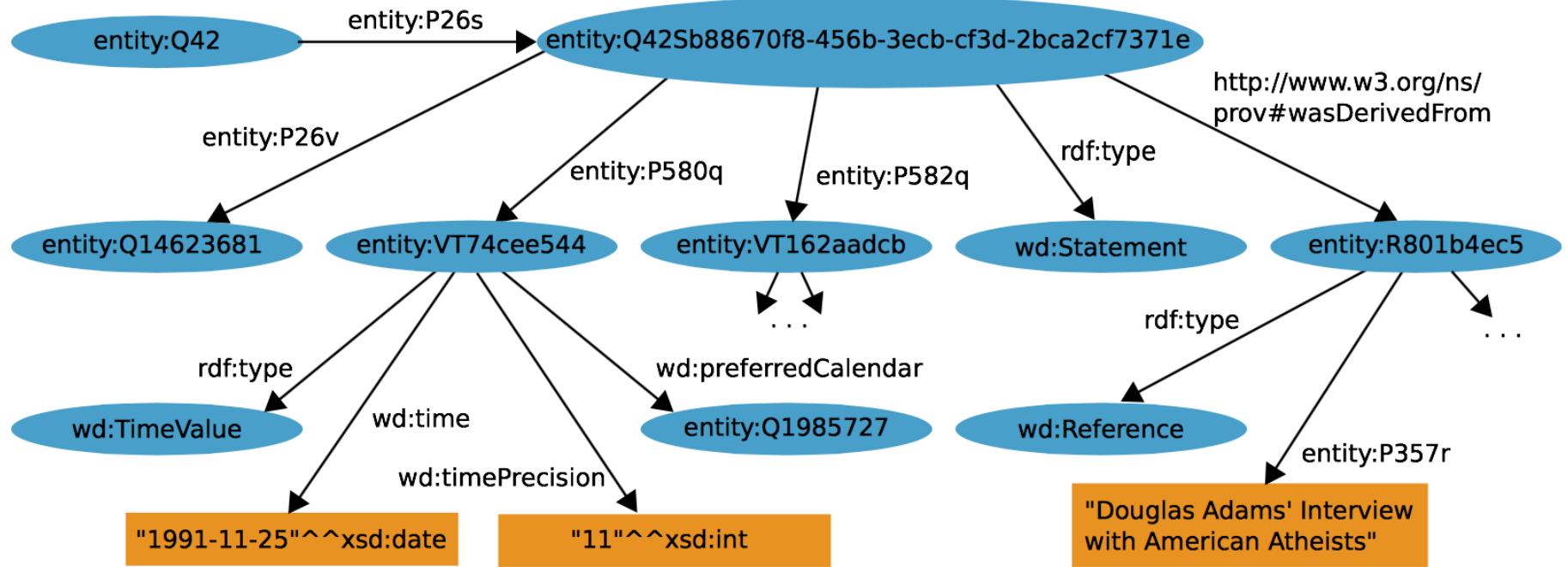
Wikipedia pages linked to this item (64 entries)

Language	Code	Linked page	
العربية	arwiki	دouglas آدمز	[edit]
مصرى	arzwiki	دوجلاس ادامز	[edit]
Boarisch	barwiki	Douglas Adams	[edit]
беларуская	be_x_oldwiki	Дуглас Адамз	[edit]

Qualified Statements and References

spouse	 Jane Belson
	start date 25 November 1991
	end date 11 May 2001
▼ 1 reference	
reference URL	http://www.nndb.com/people/731/000023662/ 
original language	English
title	Douglas Adams

Representing and querying qualified statements and references



Source: <http://korrekt.org/papers/Wikidata-RDF-export-2014.pdf>

```
SELECT ?property ?object ?reference WHERE
{
    entity:Q42 ?s_property ?sid .
    ?sid ?v_property ?object ;
        <http://www.w3.org/ns/prov#wasDerivedFrom> ?reference .
    FILTER ( REPLACE(str(?s_property), ".$", "") =
              REPLACE(str(?v_property), ".$", ""))
}
```

RDR – Reification Done Right

- Simple extension to RDF and SPARQL
- Allows to use a statement as the subject of another statement
- Formalized by Hartig and Thompson in
[Foundations of an Alternative Approach to Reification in RDF](#)
- @prefix : <<http://www.wikidata.org/entity/>> .
@prefix prov: <<http://www.w3.org/ns/prov#>wasDerivedFrom> .
<<:Q42 :p26 :Q14623681>> prov:wasDerivedFrom :R801b4ec5 .
- SELECT ?property ?object ?reference WHERE
{
 << :Q42 ?property ?object >> prov:wasDerivedFrom ?reference .
}

Conclusions

- Wikidata as useful knowledge graph with real life use cases
- Wikidata query services currently being developed based on RDF / SPARQL
- Expressive Wikidata data model that poses challenges for representation in RDF
- Reification Done Right as possible approach for dealing with qualified statements and references
- Standardization desired
- Interesting from benchmarking perspective

Contact us!

metaphacts GmbH

Kautzelweg 13

69190 Walldorf

Germany

p +49 6227 8308660

m +49 157 50152441

e info@metaphacts.com

 [@metaphacts](https://twitter.com/metaphacts)