

Geospatial data and Scholia

Finn Årup Nielsen, Daniel Mietchen, Egon Willighagen

Cognitive Systems, DTU Compute, Technical University of Denmark;
Data Science Institute, University of Virginia; Dept of Bioinformatics -
BiGCaT, NUTRIM, Maastricht University

3 June 2018



WIKIDATA

How much geospatial data do we have?



Daniel Mietchen
@EvoMRI

Following



#Wikidata now knows of over 50 000 cases where creative works have been tagged with a topic for which it knows a #geolocation

#SPARQL query:

query.wikidata.org/#%23defaultView

...

#LOD

3:44 PM - 29 May 2018

5'855'337 Wikidata geocoordinate links according to the query:

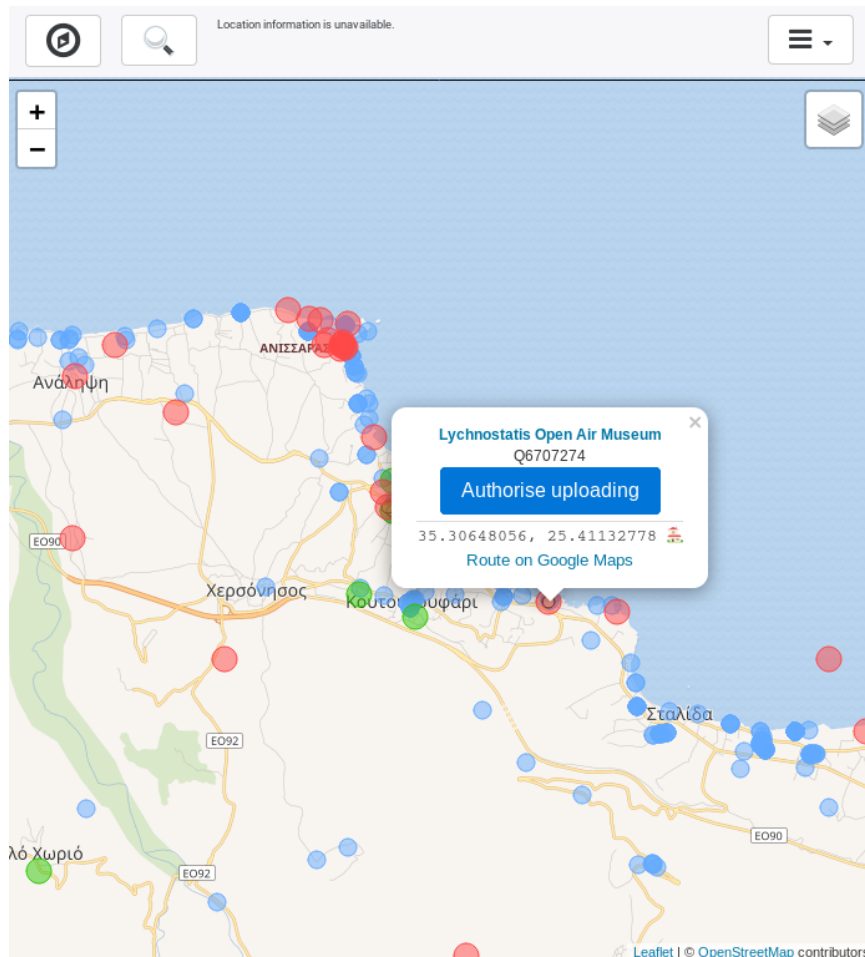
```
SELECT (COUNT(*) AS ?cnt)
WHERE { [] wdt:P625 [] . }
```

From Chinese village, Dutch roads, Danish restaurants, ...

The numbers: around 48 million Wikidata items, over 5 milliard triples, over 13 million DOI links, around 39 thousand geolocatable topics of works with DOI:

```
SELECT (COUNT(*) AS ?count) WHERE { [] wdt:P356 [] ;
wdt:P921 / wdt:P625 [] }
```

Application of Wikidata geospatial data



There are several applications using geospatial data from Wikidata.

Magnus Manske has produced *Reasonator* that displays a map for a specific Wikidata item and *Wikishootme* that shows a map with geolocatable Wikidata items missing an image (see screenshot).

You can also discover Wikidata items near your with the special URL <https://www.wikidata.org/wiki/Special:Nearby> in the MediaWiki software.

Scholia

Scholia Author Work Organization Location Event Award Topic Tools Help

Search

Search for a scientist, paper, organization, venue, event, topic, etc.

Examples

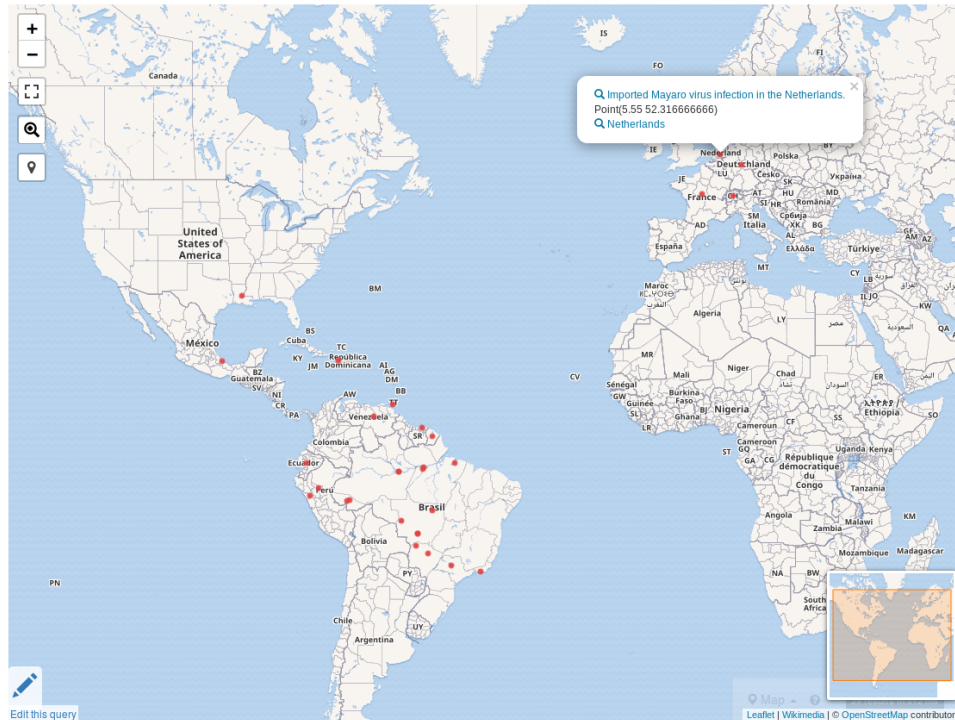
Profiles	Comparisons	Redirects
<p>Denny Vrandečić View the researcher profile for the Semantic Web researcher Denny Vrandečić. It shows his papers, co-authors, etc.</p> <p>Technical University of Denmark View the profile for an organization: People associated with the organization, their publications, the co-author patterns, etc.</p> <p>NeuroImage View information about a venue, e.g., a scientific journal or scientific conference. Here, the <i>NeuroImage</i> journal, its recently published paper, authors, topics, citation pattern, etc.</p> <p>Public Library of Science View information about a publisher, here <i>Public Library of Science</i>, with, e.g., the journals it publishes.</p> <p>Zika virus View information about a topic, e.g. <i>Zika virus</i>, e.g., the authors or journals publishing on the topic.</p>	<p>Scholia can show multiple items together.</p> <p>Technical University of Denmark and UCL Compare two or more organizations. Here a comparison between two universities with collaborating researchers, number of publications and citations.</p> <p>Tim Berners-Lee, James Hendler and Ruben Verborgh Compare three Semantic Web researchers.</p>	<p>If you know the external identifier of a concept, then Scholia can make a lookup based on it:</p> <p>twitter/utafrih Look up by Twitter username @utafrih. This will identify the London-based researcher Uta Frith and redirect to her Scholia page.</p> <p>twitter/mitpress Redirect also works for organizations, here MIT Press</p> <p>orcid/0000-0001-7542-0286 Lookup 0000-0001-7542-0286 that is identifying Egon Willighagen.</p> <p>github/vedina Redirect via GitHub username, here @vedina to Nina Jeliazkova.</p> <p>doi/10.1186/S13321-016-0161-3 Redirect via a DOI.</p> <p>viaf/59976288 Redirect via VIAF identifier, here to Ben Feringa</p>

Scholia is a webservice running from <https://tools.wmflabs.org/scholia/>

Display information from Wikidata about researchers, works and their citations, organizations, venues, events, topics, etc.

Panels for each Wikidata item constructed with calls to the *Wikidata Query Service* (WDQS), showing tables and plots such as bubble charts and line plot as well as OpenStreetMap-based maps.

Geospatial data with Scholia: topic



Maps established with simple queries to WDQS.

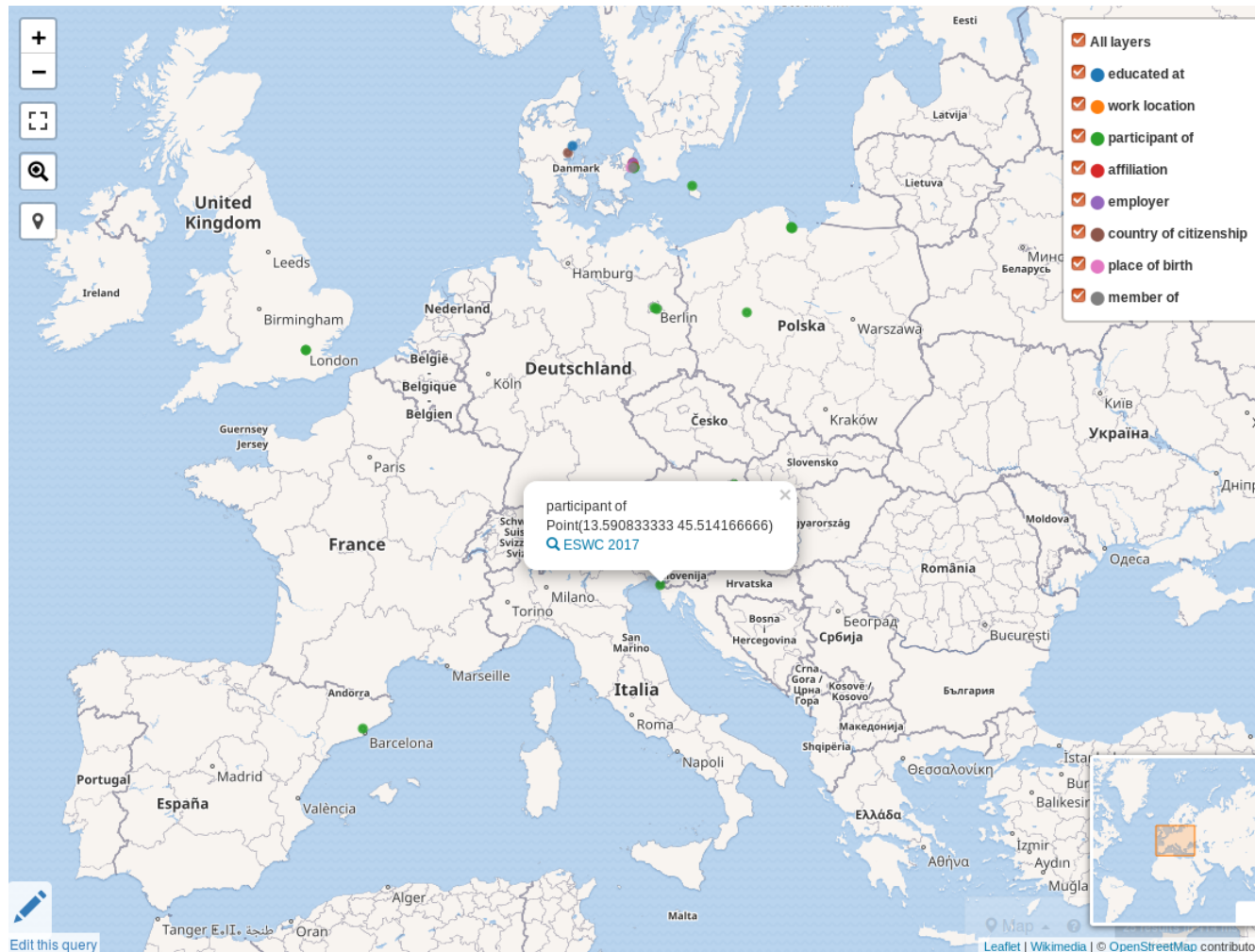
Find works about a topic (here **Mayaro virus**) with a SPARQL path query:

```
?work wdt:P921 /  
( wdt:P31*/wdt:P279*  
| wdt:P361+ | wdt:P1269+ )  
wd:Q18863953 .
```

Identify co-occurring topic that is geo-locatable.

```
?work wdt:P921 ?location .  
?location wdt:P625 ?geo .
```

Geospatial data with Scholia: author



Author aspect: </scholia/author/Q20980928>

With node coloring controlled by the type of property.

New Scholia aspects: location and country

Geospatial data with Scholia: location

[location](#)

Aldemar Knossos Royal ([Q47259960](#))

Nearby organizations

Show entries

Search:

Distance	Organization
25.57356420872047	Centre for Technological Research of Crete
27.49884916774176	University of Crete
27.616884625684804	Library and Information Center of the University of Crete
46.175972083464266	Skinakas Observatory
121.23846544518143	Technical University of Crete

[Edit on query.Wikidata.org](#)

Showing 1 to 5 of 5 entries

Previous

1

Next

Scholia location aspect for a Cretian hotel:

[/location/Q47259960](#)

SPARQL query with the distance function called `geof:distance` showing nearby academic institutions.

Other panel: Nearby locations as topics in works. Identifies, e.g., *Tomb Robbing and the Transformation of Social Memory in Roman Knossos* ([Grigoropoulos, 2004](#)) as an article with a nearby topic.

Geospatial data with Scholia: location

SPARQL query for identifying nearby academic institutions

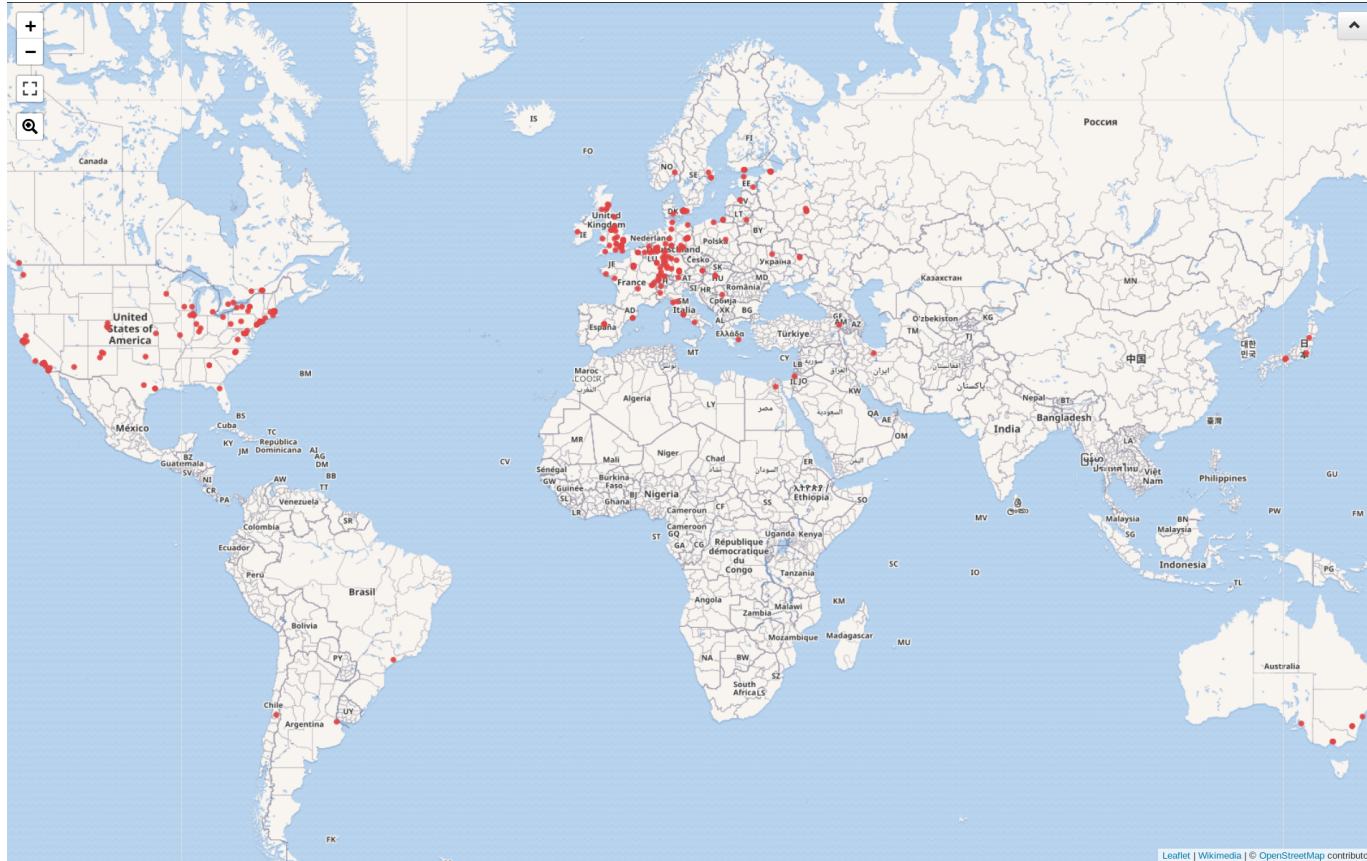
```
# Academic institution
VALUES ?university { wd:Q3918 wd:Q1371037 wd:Q7315155
                      wd:Q31855 }

...
# Find individual universities and departments
# and the geocoordinate
?organization wdt:P361* / wdt:P31 / wdt:P279* ?university .
?organization wdt:P625 ?other_geo .

...
# Compute distance between academic institution
# and the query location
wd:Q47259960 wdt:P625 ?geo .
BIND(geof:distance(?other_geo, ?geo) AS ?distance)
FILTER(?distance < 250)

...
ORDER BY ?distance
```

Countries in Scholia



Map in Scholia with international collaborators of authors based in the Netherlands: </scholia/country/Q55>

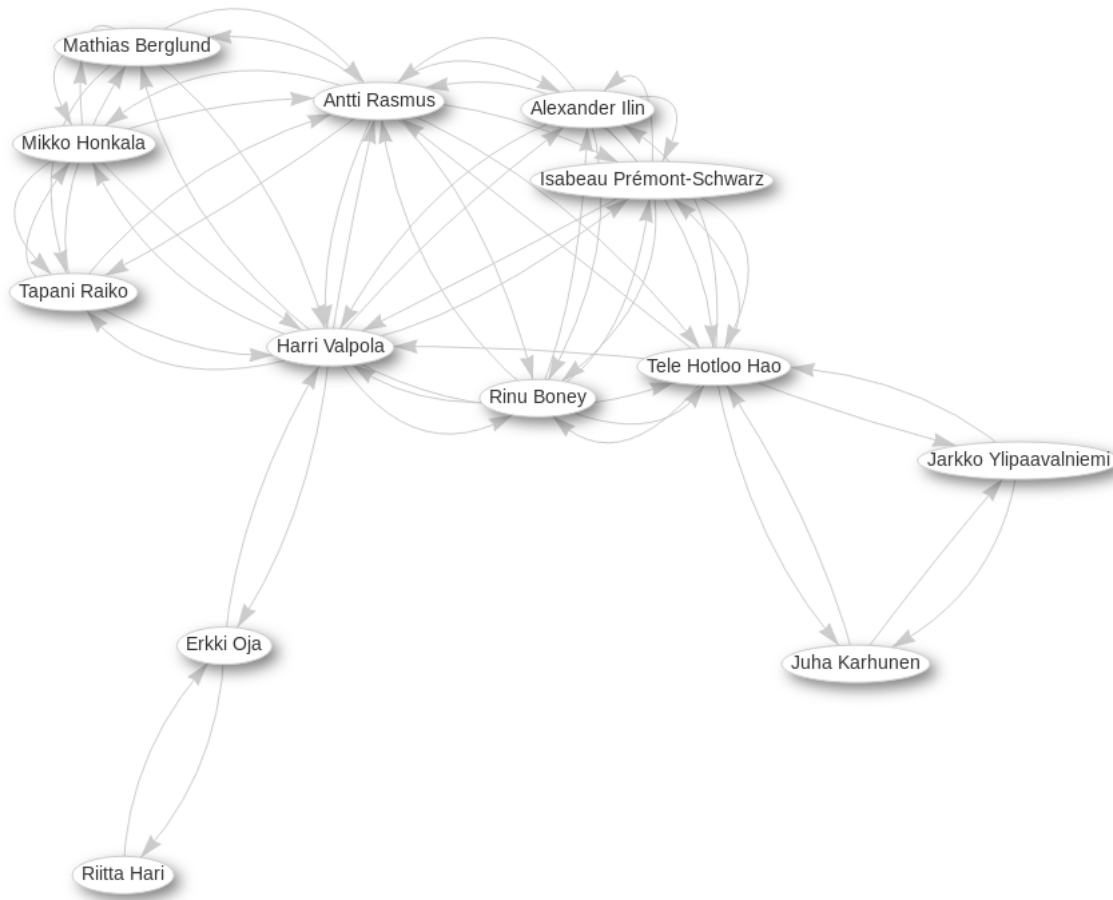
Other map panels for the country aspect display narrative locations within works and location as topics in works.

User stories

User story: Finnish machine learning

You are to review research applications from Finland about machine learning and related research fields. You are based outside Finland and would like to get an overview of Finnish researchers and research organizations in that research area, their works as well as their collaboration and citation patterns.

User story: Finnish machine learning



One of the panels on the Finnish machine learning country–topic aspect page with display of the co-author graph.

Using item links between researchers, affiliation and country to identify Finnish researchers with no need for query on geocoordinate data.

Combination of country and topic: </scholia/country/Q33/topic/Q2539>

User story: Wikipedia researchers in Tübingen

You are a researcher interested in Wikipedia research and planning a visit to Tübingen where you would like to meet other Wikipedia researchers.

User story: Wikipedia researchers in Tübingen

Nearby researchers

Show entries Search:

Score	Author	Example work
18.13370167064972	Ulrike Cress	A productive clash of perspectives? The interplay between articles' and authors' perspectives and their impact on Wikipedia edits in a controversial domain
4.909317231401991	lassen Halatchliyski	A productive clash of perspectives? The interplay between articles' and authors' perspectives and their impact on Wikipedia edits in a controversial domain
1.604942154393766	Jason Weston	Reading Wikipedia to Answer Open-Domain Questions
0.08335000742150632	Denny Vrandečić	Wikipedia and the Semantic Web - The Missing Links
0.04167500371075316	Rudi Studer	Semantic Wikipedia

Combination of location and topic: </scholia/location/Q3806/topic/Q52>.

```
(SUM(?topic_score) * MAX(?inverse_distance) AS ?score)
```

...

```
?work wdt:P921 / wdt:P279* wd:Q52 . BIND(3 AS ?topic_score) }
UNION { ?author wdt:P101 wd:Q52 . BIND(20 AS ?topic_score) }
```


User story: conference hunter

You are going to The Web Conference in April 2018 in Lyon. You want to know if there is any other relevant scientific meeting in the local area at that time, preferably just before or just after the conference.

User story: conference hunter

Related events

Related events in terms of time and location.

Show entries

Search:

Score	Time score	Location score	Event
2000.0	10	200.0	Semantics, Analytics, Visualisation: Enhancing Scholarly Dissemination Workshop co-located with The Web Conference 2018
2000.0	10	200.0	Wiki Workshop 2018
0.48289635681516646	2	0.24144817840758323	Coding da Vinci Ost 2018
0.3528879254047528	0.7547169811320754717	0.37969864810615994	Wikimedia Hackathon 2018
0.2419491754542301	0.85106382978723404255	0.28429028115872035	4th Dutch Bioinformatics & Systems Biology Conference
0.22856288438578842	6.66666666666666666667	0.03428443265786826	CHI 2018
0.15614832986186872	0.56338028169014084507	0.27716328550481695	11th International Conference on Chemical Structures
0.1395750399245413	0.22857142857142857143	0.2058556595973786	1st Workshop on Quality of Open Data
0.10699910223060025	1.33333333333333333333	0.08024932667295019	21st International Conference on Artificial Intelligence and Statistics
0.09521655169911777	0.21621621621621621622	0.44037655160841965	PIDapalooza 2018

[Edit on query.Wikidata.org](#)

Showing 1 to 10 of 93 entries

Previous 2 3 4 5 ... 10 Next

Related events panel for *The Web Conference 2018* in Scholia.

User story: conference hunter

Event aspect for *The Web Conference 2018*: </scholia/event/Q48910401> where the SPARQL query combines inverse distance and inverse time separation:

```
BIND(20 / (1 + ABS(?day - ?day0)) AS ?time_score_)
...
# inverse distance
BIND((200 / (1 + geof:distance(?geo, ?geo0))) AS
      ?inverse_distance)
...
BIND((?time_score_ * ?location_score_) AS ?score_)
...
ORDER BY DESC(?score)
```

Summary

Large amount of geospatial data in Wikidata, including geospatial data tied to scientific items (articles, researchers, organizations).

... and continuous expansion of the data.

WDQS makes it easy to create maps of the data in Wikidata.

Scholia uses the capabilities of WDQS to render maps and compute distances for a range of different scholar-associated data.

References

Grigoropulous, D. (2004). [Tomb Robbing and the Transformation of Social Memory in Roman Knossos](#). pages 62–77.

Copyright and license

OpenStreetMap maps are [Map © OpenStreetMap contributors. CC BY-SA 2.0.](#)

Wikidata logo by Arun Ganesh (Planemad). It is a trademark of the Wikimedia Foundation.

Wikishootme is the work of Magnus Manske.