Scholia as of November 2018

Finn Årup Nielsen & Daniel Mietchen

Cognitive Systems, DTU Compute, Technical University of Denmark & Data Science Institute, University of Virginia

27 November 2018
https://tools.wmflabs.org/scholia/
Warning: “Scholia relies on Wikidata, and Wikidata contains only a limited albeit growing . . .”

Example profiles for author, organizations, venues, etc.

Further down: live statistics, e.g., 168523555 Citations, 15439388 Items with a DOI, 333715 Items about authors with an ORCID profile, etc.
Scholia’s search

Scholia’s search searches across all Wikidata items, not just “WikiCite items”.

Uses the standard Wikidata API via JavaScript.

(There is also a hidden search page: https://tools.wmflabs.org/scholia/search)
WikiCite conference

“Event series”: One of several aspects, where other are, e.g., author, organization, topic, publisher, etc.

Event series aspect shows WikiCite 2016, 2017, 2018 in list and timeline.

All created with SPARQL queries to the Wikimedia Foundation’s Wikidata Query Services (WDQS): The list is converted by JavaScript, while the timeline is directly embedded from the WDQS result.
WikiCite conference

We got several other panels on the page for the WikiCite conference:

Maps of the WikiCite conference WDQS and OpenStreetMap.

Topics (what is the WikiCite conferences about?): Wikipedia, WikiCite, Wikidata, ontology, citation analysis, etc. These topics are found by querying for people associated with the event and finding the domain and main subject of their authored works.

“Recent publications” panel.
“Event” aspect: single event with specific time and place.

People associated with the event: organizers, speakers, participants, program committee members, — all pull from Wikidata.

Here with ordering with the number of publications and example publication.
Scholia as of November 2018

WikiCite 2018
WikiCite 2018

Panel with “related events” for WikiCite 2018, where the relatedness is based on time and location.


Second: 2018 IEEE International Conference on Big Data (Seattle, 10 December October)

Third and with first time score: SWIB18 (Bonn, 26 November 2018)
**WikiCite participant**

**Dario Taraborelli** (Q21562060)

Related: Daniel Mietchen · Giuseppe De Simone · David N. Kennedy · Cameron Neylon · Alessandro Vespignani · Lars Kai Hansen · Russell A. Pottorff · Kristoffer Hougaard Madsen · Klaus Kähler Holst · Simona Maria Monti

[ORCID](https://orcid.org/0000-0002-0082-8508)

**List of publications**

<table>
<thead>
<tr>
<th>Date</th>
<th>Work</th>
<th>Type</th>
<th>Pages</th>
<th>Venue</th>
<th>Authors</th>
</tr>
</thead>
<tbody>
<tr>
<td>2018-10-12</td>
<td>A hands-on Introduction to Wikidata and WikiCite</td>
<td>lecture</td>
<td></td>
<td>Figshare</td>
<td>Dario Taraborelli, Daniel Mietchen</td>
</tr>
<tr>
<td>2018-05-14</td>
<td>Conversations Gone Awry: Detecting Early Signs of Conversational Failure</td>
<td>scholarly article</td>
<td>16</td>
<td>Proceedings of the 56th Annual Meeting of the Association for Computational Linguistics</td>
<td>Dario Taraborelli, Cristian Danescu-Niculescu-Mizil, Lucas Dixon, Nithum Thain, Yi-Qing Hua</td>
</tr>
<tr>
<td>2018-04-05</td>
<td>What are the ten most cited sources on Wikipedia? Let's ask the data</td>
<td>blog post</td>
<td></td>
<td>Wikimedia Blog</td>
<td>Dario Taraborelli, Ben Vershinbow, Jake Orlowitz, Miriam Redi</td>
</tr>
<tr>
<td>2018-03-23</td>
<td>Citations with Identifiers in Wikipedia</td>
<td>data set</td>
<td></td>
<td>Figshare</td>
<td>Dario Taraborelli, Aaron Halfaker, Miriam Redi</td>
</tr>
</tbody>
</table>
Work: The altmetrics collection

“Work” aspect for The altmetrics collection.

List of authors, Topic bubble chart, timeline, citations.
Related works

We can find related works based on analysis of the co-citation network around

The central part of the WDQS SPARQL query is:

```
wd:Q21560746
(\(^{\text{\texttt{wdt:P2860}}} \mid \text{\texttt{wdt:P2860}}\))
/ 
(\(^{\text{\texttt{wdt:P2860}}} \mid \text{\texttt{wdt:P2860}}\))? ?work .
```

“Do altmetrics work? Twitter and ten other social web services” most related.
Co-citation graph for The altmetrics collection
WikiCite 2018 is a 3-day conference, summit, and hack day dedicated to the vision of creating an open repository of bibliographic data to support the citation and fact-checking needs of Wikimedia projects, and possibly, to serve as an open infrastructure for research, education, and information quality across the web.

WikiCite 2018 expands efforts started with WikiCite 2016 and WikiCite 2017 to explore these possibilities by leveraging Wikidata—Wikimedia’s structured knowledge base—and to develop open source tools to improve citation management and verifiability strategies for free knowledge. Since then, the amount of bibliographic data in Wikidata has grown exponentially, allowing us to glimpse the possibilities of a truly open, universal library and citation graph, while also exposing significant social and technical challenges.

This year presents a pivotal moment for WikiCite, wherein its emergent community — consisting of Wikimedians, librarians, LODLAM practitioners, software engineers, data scientists, and open knowledge advocates — must grapple with possible growth scenarios, address critical gaps, and set a course for the project’s future.

If you are passionate about tending Wikipedia’s root system (reference!), or believe in the broader possibilities of contributing to the bibliographic commons, WikiCite 2018 presents an unprecedented opportunity to meet fellow dreamers and hackers, and to help shape this vital effort.

This year’s event will be hosted at the David Brower Center (map) in Berkeley, California, USA, November 27-29, 2018. Applications to attend the event (including travel support requests) are open until September 17, 2018.
Text-to-topics

Text

Copy and paste an abstract to the text area and press submit to extract Wikidata labels and query.

WikiCite 2018 is a 3-day conference, summit, and hack day dedicated to the vision of creating an open repository of bibliographic data to support the citation and fact-checking needs of Wikimedia projects, and possibly, to serve as an open infrastructure for research, education, and information quality across the web.

WikiCite 2018 expands efforts started with WikiCite 2016 and WikiCite 2017 to explore these possibilities by leveraging Wikidata—Wikimedia's structured knowledge base—and to develop open source tools to improve citation management and verifiability strategies for free knowledge. Since then, the amount of bibliographic data in Wikidata has grown exponentially, allowing us to glimpse the possibilities of a truly open, universal library and citation graph, while also exposing significant social and technical challenges.

This year presents a pivotal moment for WikiCite, wherein its emergent community — consisting of Wikimeanders, librarians, LODLAM practitioners, software engineers, data scientists, and open knowledge advocates — must grapple with possible growth scenarios, address critical gaps, and set a course for the project’s future.

If you are passionate about tending Wikipedia's root system (references!), or believe in the broader possibilities of contributing to the bibliographic commons, WikiCite 2018 presents an unprecedented opportunity to meet fellow dreamers and hackers, and to help shape this vital effort.

Keyword extraction in Scholia

At https://tools.wmflabs.org/scholia/text-to-topics
# Text-to-topics

## Topics

<table>
<thead>
<tr>
<th>Topic</th>
<th>Example work</th>
</tr>
</thead>
<tbody>
<tr>
<td>travel</td>
<td>Zika Virus Infection Among U.S. Pregnant Travelers — August 2015–February 2016</td>
</tr>
<tr>
<td>academic conference</td>
<td>Women's visibility in academic seminars: women ask fewer questions than men</td>
</tr>
<tr>
<td>community</td>
<td>Women on women: lesbian identity, lesbian community, and lesbian comics.</td>
</tr>
<tr>
<td>Wikidata</td>
<td>Wikidata: A Free Collaborative Knowledgebase</td>
</tr>
<tr>
<td>research</td>
<td>Why the impact factor of journals should not be used for evaluating research</td>
</tr>
<tr>
<td>information quality</td>
<td>Third Molars on the Internet: A Guide for Assessing Information Quality and Readability</td>
</tr>
<tr>
<td>education</td>
<td>The Tragedy of the Unexamined Cat: Why K-12 and University Education Are Still in the Dark Ages and How Citizen Science Allows for a Renaissance</td>
</tr>
<tr>
<td>volition</td>
<td>The psychology of volition</td>
</tr>
<tr>
<td>Open Knowledge</td>
<td>The Open Knowledge Foundation: open data means better science</td>
</tr>
<tr>
<td>International</td>
<td></td>
</tr>
<tr>
<td>future</td>
<td>The Future of Data Analysis</td>
</tr>
</tbody>
</table>
## Text-to-topics

### List of works

<table>
<thead>
<tr>
<th>Count</th>
<th>Publication date</th>
<th>Work</th>
<th>Topics</th>
</tr>
</thead>
<tbody>
<tr>
<td>4</td>
<td>2017-08-08</td>
<td>What do Wikidata and Wikipedia have in common? An analysis of their use of external references</td>
<td>Wikipedia // citation // Wikidata // WikiCite</td>
</tr>
<tr>
<td>3</td>
<td>2018-04-05</td>
<td>What are the ten most cited sources on Wikipedia? Let's ask the data</td>
<td>Wikipedia // citation // WikiCite</td>
</tr>
<tr>
<td>3</td>
<td>2018-01-01</td>
<td>Meet the Data</td>
<td>Wikipedia // Wikidata // data</td>
</tr>
<tr>
<td>3</td>
<td>2017-09-23</td>
<td>Analysis of References Across Wikipedia Languages</td>
<td>Wikipedia // citation // WikiCite</td>
</tr>
<tr>
<td>3</td>
<td>2017-08-08</td>
<td>Provenance information in a collaborative knowledge graph: an evaluation of Wikidata external references</td>
<td>citation // Wikidata // WikiCite</td>
</tr>
<tr>
<td>3</td>
<td></td>
<td>WikiCon</td>
<td>Wikipedia // Wikidata // Wikimedia Movement</td>
</tr>
<tr>
<td>2</td>
<td>2018-11-15</td>
<td>200 plus 1: Forschungsmethoden für offenes #Radfahrenwissen</td>
<td>knowledge // research</td>
</tr>
<tr>
<td>2</td>
<td>2018-10-12</td>
<td>A hands-on introduction to Wikidata and WikiCite</td>
<td>Wikidata // WikiCite</td>
</tr>
</tbody>
</table>
Scholia as of November 2018

Scholia usage statistics

Monthly pageview for Scholia has increased and has been over 300,000.

The latest increase is likely due to inclusion of link to Scholia from Wikimedia Commons templates. Whether page view coming this way are bots or users are not known.
Thanks!
Licenses

Maps are copyrighted by OpenStreetMap creators and licensed under their license Open Data Commons Open Database License by OpenStreetMap Foundation.

Text taken from the the WikiCite 2018 homepage at https://meta.wikimedia.org/wiki/WikiCite_2018 by Jake Orlowitz, Dario Taraborelli and other. Under CC BY-SA