

Wikicite discussion and lightning talks

**WIKIMANIA
SINGAPORE**

Friday, August 18 2023



Session program

Introduction to Wikicite – Phoebe Ayers

Lightning talks:

- Cita: Citation Graphs in Zotero with Wikidata – Diego de la Hera
- Integrating Scholia and ORKG for real-time research assessment – Houcemeddine Turki
- WD:CRAP - WikiProject-based source reliability & WP 2.0 – SJ Klein
- Visualizing Wikidata by using ALEC – Rod Page

Discussion – all

Slides & videos for this session: <https://is.gd/OEjg8X>

Etherpad: <https://etherpad.wikimedia.org/p/8TFHHG>

WIKIMANIA
SINGAPORE

Introduction to Wikicite

Phoebe Ayers

Slides & videos for this session: <https://is.gd/OEjg8X>
Etherpad: <https://etherpad.wikimedia.org/p/8TFHHG>

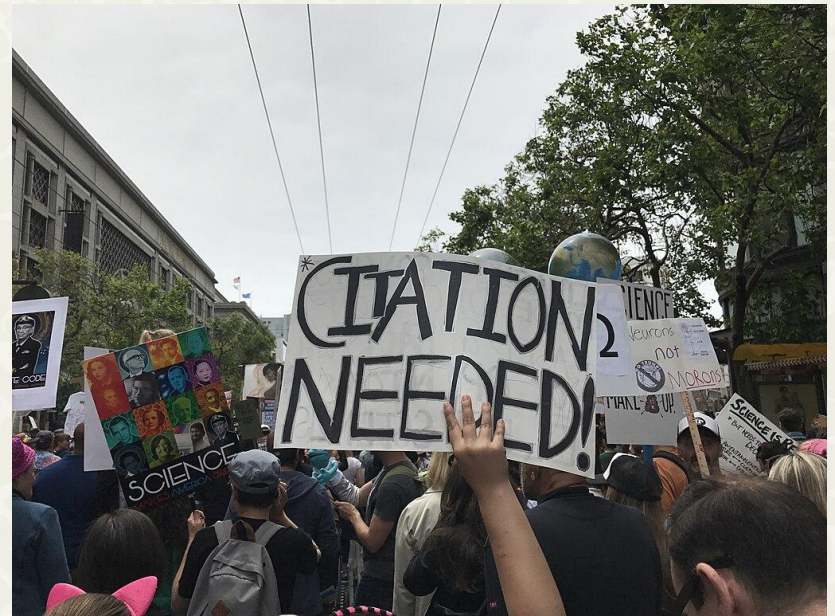
WIKIMANIA
SINGAPORE



Wikicite is...

- The community who is interested in bibliographic metadata and citations in the Wikimedia projects
- The community supporting citation data projects in Wikidata
- A series of conferences and grants (from 2016-2020)

**WIKIMANIA
SINGAPORE**



(a few of the) current challenges and questions

- Where should bibliographic metadata live (in Wikidata, another Wikibase) ?
- How can we share citation data across projects?
- What tools can we build to understand citation use on the projects and improve quality?
- What bibliographic metadata should be in the Wikipedia projects?



How to get in touch and join Wikicite discussions

Mailing list, Telegram, social media:

<https://meta.wikimedia.org/wiki/WikiCite>

Monthly calls - announced on mailing list

**WIKIMANIA
SINGAPORE**



Some other Wikicite related talks on this year's program

- Lane Raspberry - “WikiCite and Scholia for medicine and beyond”
- Houcemeddine Turki - “Empowering Wikidata and Wikipedia with Generative AI: Unlocking the Potential of Scholarly Publications”
- Mark Graham - “Internet Archive: The Annual Wikimania Update”
- Diego de la Hera - “Collaboratively fixing automatic citations in Wikipedia with Web2Cit”

**WIKIMANIA
SINGAPORE**



Cita: Citation Graphs in Zotero with Wikidata

Diego de la Hera

Slides & videos for this session: <https://is.gd/OEjg8X>
Etherpad: <https://etherpad.wikimedia.org/p/8TFHHG>

WIKIMANIA
SINGAPORE



Cita

Citation graphs in Zotero
with Wikidata

WikiCite discussion
18/Aug/23

Diego de la Hera
Wikimedia Argentina /
Wikimedistas Calamuchita

[C]ita

a **WIKIDATA**
addon
for **Zotero**

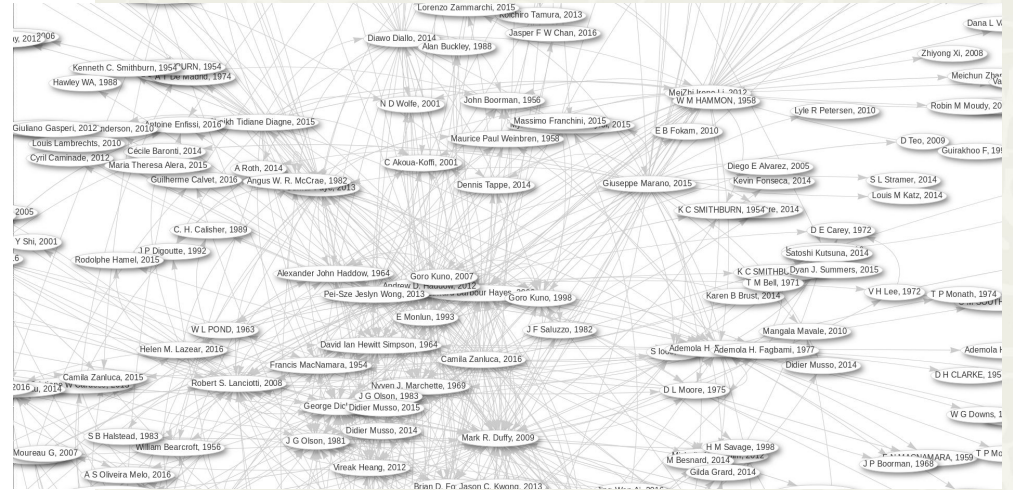
WIKIMANIA
SINGAPORE

{ }
wikicite

On the shoulders of giants

Citations speak of knowledge evolution

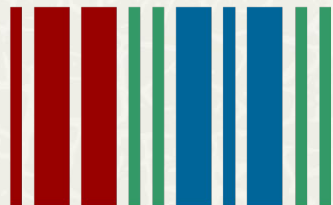
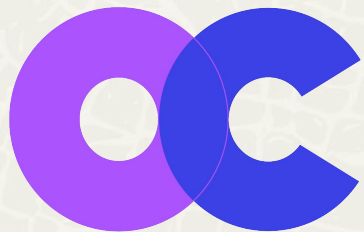
**WIKIMANIA
SINGAPORE**



Citation graph for Zika Virus Outbreak -
CC0 Daniel Mietchen (Wikimedia Commons)

CC0 - Diego de la Hera

I40C



WIKIDATA

**WIKIMANIA
SINGAPORE**

Initiative for Open Citations

“to promote the unrestricted availability of scholarly citation data”

Lack of reference manager support

References as bibliographic metadata

The screenshot shows a reference manager interface with a list of references on the left and a detailed view of a selected reference on the right.

| Title | Creator |
|--|------------------------|
| ▼ A Glimpse into Babel: An Analysis ... | Kaffee et al. |
| Full Text | |
| Cross-Lingual Natural Language Q... | Aggarwal et al. |
| ▶ DBpedia – A large-scale, multiling... | Lehmann et al. |
| ▶ Freebase, Freebase: a collaborati... | Bollacker et al. |
| ▶ From Freebase to Wikidata: The G... | Tanon et al. |
| ▶ Peer-production system or collab... | Müller-Birn et al. |
| ▶ Reifying RDF: What Works Well W... | Hernández et al. |
| ▶ Scholia and scientometrics with W... | Nielsen et al. |
| ▶ Scholia, Scientometrics and Wikid... | Nielsen et al. |
| ▶ The Semantic Web | Berners-Lee et al. |
| ▶ Wikidata as a linking hub for kno... | Neubert |
| ▶ Wikidata as a semantic framewor... | Burgstaller-Muehlb... |
| ▶ Wikidata: a free collaborative kno... | Vrandečić and Kröt... |
| ▶ Wikidata: A New Paradigm of Hu... | Piscopo |
| ▶ Wikidata: A New Platform for Coll... | Vrandečić |
| ▶ Wikidata: A platform for data int... | Mitraka et al. |
| ▶ Wikidata: a platform for your libr... | Allison-Cassin and ... |
| ▶ Wikidatians are born: paths to full... | Piscopo et al. |
| ▶ Yago | Suchanek et al. |
| ▶ YAGO2: A spatially and temporall... | Hoffart et al. |

| Info | Notes | Tags | Related |
|--------------|---|--------------------------|--------------------------|
| Item Type | Journal Article | | |
| Title | A Glimpse into Babel: An Analysis of Multilinguality in Wikidata | | |
| ▼ Author | Kaffee, Lucie-Aimée | <input type="checkbox"/> | <input type="checkbox"/> |
| ▼ Author | Piscopo, Alessandro | <input type="checkbox"/> | <input type="checkbox"/> |
| ▼ Author | Vougiouklis, Pavlos | <input type="checkbox"/> | <input type="checkbox"/> |
| ▼ Author | Simperl, Elena | <input type="checkbox"/> | <input type="checkbox"/> |
| ▼ Author | Carr, Leslie | <input type="checkbox"/> | <input type="checkbox"/> |
| ▼ Author | Pintscher, Lydia | <input type="checkbox"/> | <input type="checkbox"/> |
| Abstract | | | |
| Publication | Proceedings of the 13th International Symposium on Open Collaboration | | |
| Volume | | | |
| Issue | | | |
| Pages | | | |
| Date | 2017-08 | | ym |
| Series | | | |
| Series Title | | | |
| Series Text | | | |
| Journal Abbr | | | |
| Language | En | | |
| DOI | 10. | | |
| ISSN | | | |
| Short Title | A C | | |
| URL | | | |

**WIKIMANIA
SINGAPORE**



CC0 - Diego de la Hera

[c]ita

a **WIKIDATA**
addon
for **Zotero**

{ }
wikicite

WIKIMANIA
SINGAPORE

Cita

A Wikidata addon for Zotero

wikidata.org/wiki/Wikidata:Zotero/Cita

Reference list





































































































































Fetch from Wikidata or add manually

WIKIMANIA
SINGAPORE

[c]ita

Info Notes Tags Related Citations

22 citation(s):

- Ashish and Knoblock (1997) Wrapper generation for semi-struct...      
- Breiman et al. (1984) Classification and regression trees      
- Califf and Mooney (1997) Relational learning of pattern-match ...      
- Cohen (1996) Learning trees and rules with set-valued features      
- Cohn et al. (1994) Improving generalization with active learning      
- Dagan and Engelson (1996) Sample selection in natural languag...      
- Domingos (1994) The RISE system: Conquering without separat...      
- Fisher et al. (1995) Description of the UMass system as used fo...      
- Freitag (1998) Multistrategy learning for information extraction      
- Huffman (1996) Learning information extraction patterns from...      
- Kim and Moldovan (1993) Acquisition of semantic patterns for ...      
- Krupka (1995) Description of the SRA system as used for MUC-6      
- Kushmerick et al. (1997) Wrapper induction for information ext...      
- Lewis and Gale (1994) A sequential algorithm for training text c...      
- Michalski (1983) A theory and methodology of inductive learning      
- Quinlan (1990) Learning logical definitions from relations      
- Quinlan (1993) C4. 5: Programs for machine learning      
- Riloff (1993) Automatically constructing a dictionary for infor...      
- Soderland (1997) Learning text analysis rules for domain-specif...      
- Soderland (1997) Learning to extract text-based information fr...      
- Soderland et al. (1995) CRYSTAL: Inducing a conceptual dictionary      
- Valiant (1984) A theory of the learnable      

Citation graph

Understand connections
and discover relevant works

Local Citation Network for Cita

Input articles (18) Incoming suggestions (▲ 10)

Incoming suggestions are the locally most cited references from the input articles. They are generally older.
Maximum: 10 ▾

Title & Abstract ▾ Filter articles... 10 / 10

| Title | First Author | Year | InDeg | OutDeg |
|---|--------------|------|-------|--------|
| Semantic Wikipedia ↔ | Kröttsch | 2007 | 3 | 0 |
| DBpedia: A Nucleus for a Web of Open Data ↔ | Auer | 2007 | 2 | 0 |
| A novel family of mammalian taste receptors ↔ | Adler | 2000 | 2 | 0 |
| A principal component analysis of 39 scientific impact measures ↔ | Bollen | 2009 | 2 | 0 |
| Time evolution of Wikipedia network ranking ↔ | Frahm | 2013 | 2 | 0 |
| Molecular Structure of Nucleic Acids: A Structure for Deoxyribose Nucleic Acid ↔ | Watson | 1953 | 2 | 0 |
| NeuroVault.org: A web-based repository for collecting and sharing unthresholded statistical maps of the human brain ↔ | Gorgolewski | 2015 | 2 | 0 |
| Flask Web Development: Developing Web Applications with Python ↔ | Grinberg | 2014 | 2 | 0 |
| Defrosting the digital library: bibliographic tools for the next generation web. ↔ | Hull | 2008 | 2 | 0 |
| Functional Requirements for Bibliographic Records: Final Report ↔ | Records | 2009 | 2 | 0 |

Citation network (fullscreen network) Co-authorship network ↑ newer articles

(how to read this network) ↓ older articles

Adapted from Tim Wölfe's Local Citation Network ↔

WIKIMANIA
SINGAPORE

[c]ita

CC0 - Diego de la Hera

Contribute to Wikidata

Lower the barrier to Wikidata contribution

WIKIMANIA
SINGAPORE

[c]ita

The screenshot shows a citation management interface with a list of citations and a dropdown menu. The citations list includes:

- Adler et al. (2000) C.S.: A novel family of...
- Alpher et al. (1948) The origin of chemi...
- Bollen et al. (2009) R.: A principal comp...
- Eom et al. (2013) Time evolution of Wik...
- Erxleben et al. (2014) Introducing Wiki...
- Gorgolewski et al. (2015) NeuroVault.c...
- Grinberg and Flask (2014)
- Heller et al. (2013) I.: InChI - the worldw...
- Hernandez et al. (2015) Reifying RDF: w...
- Hernandez et al. (2016) Querying Wikid...
- Hull et al. (2008) Defrosting the digital library: bibliographic...
- (2009) IFLA Study Group on the Functional Requirements fo...
- Khabsa and Giles (2014) The number of scholarly documents...
- Kikkawa et al. - DOI links on Wikipedia
- Springer (2016)
- Kousha and Thelwall (2016) Are Wikipedia citations importa...
- Krotzsch et al. (2006) Semantic MediaWiki
- Lin and Fenner (2014) An analysis of Wikipedia references a...
- Maqqio et al. (2017) Wikipedia as a gateway to biomedical r...

The dropdown menu is open, showing the following options:

- Sync citations with Wikidata
- Fetch QIDs for cited items
- Get citations from Crossref
- Get citations from OpenCitations Corpus
- Extract citations from attachments
- Add citation(s) by identifier
- Import citations
- Export citations to file
- Export citations to CROCI
- Sort by
- Auto link citations with Zotero items

At the bottom of the interface, the following identifiers are displayed:

- DOI 10.1007/978-3-319-70407-4_36
- QID Q41799194
- OCC

There are also three green circular icons at the bottom right of the interface.

Present and future



**WIKIMANIA
SINGAPORE**

- ~2800 users
- Bibliographic metadata on Wikidata
 - Support for other repositories
 - Cita's wide scope
- Long-term maintenance



Dominic Dall'Osto
University of Zurich,
Institute for Neuroinformatics

Thank you!

**WIKIMANIA
SINGAPORE**

{ }
wikicite

[c]ita

a **WIKIDATA**
addon
for **Zotero**

Diego de la Hera
(User:Diegodlh)

delahera AT gmail DOT com
@diegodlh

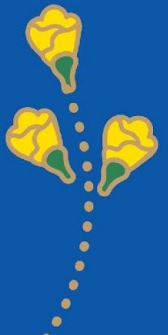
CC0 - Diego de la Hera

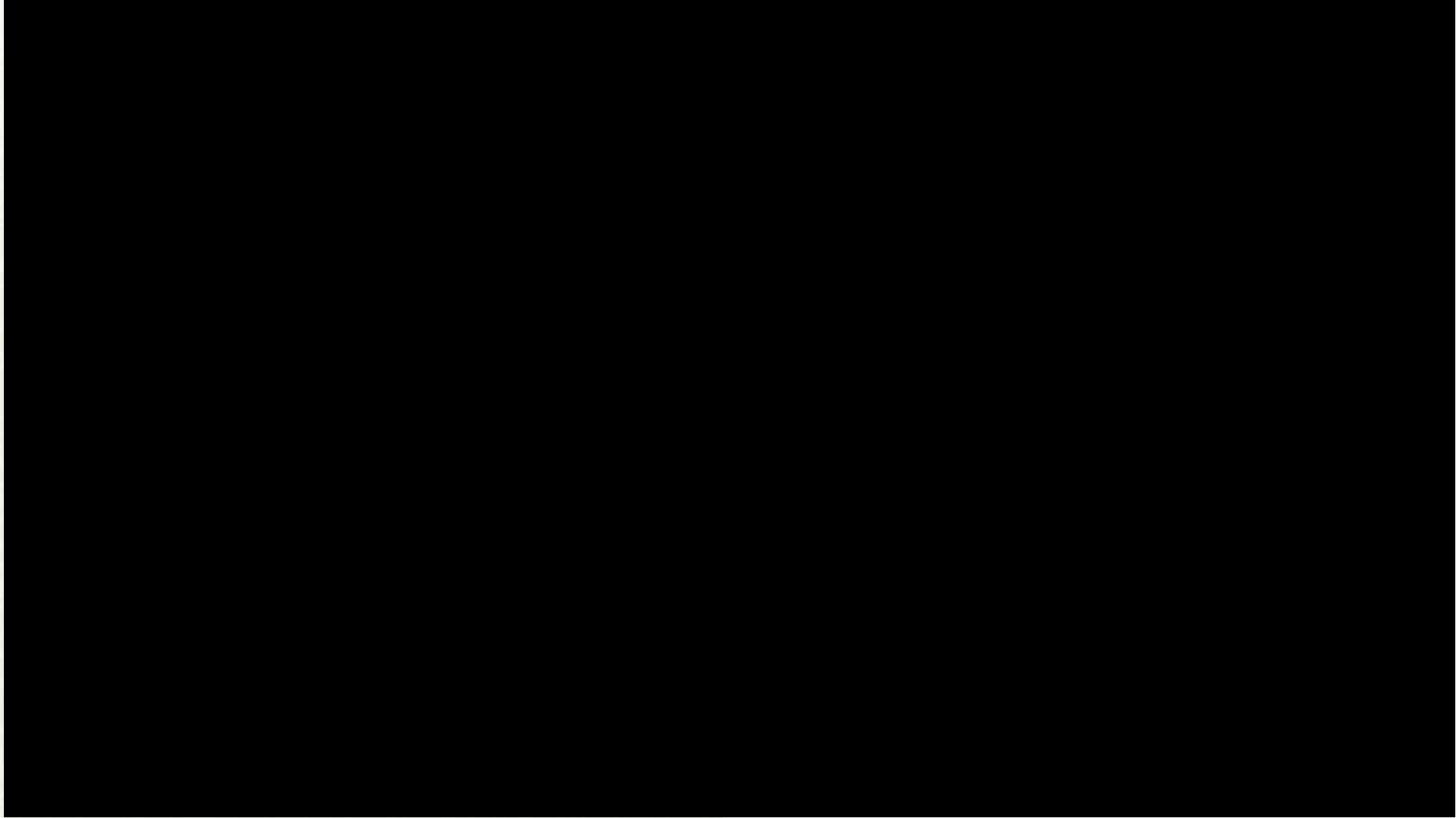
Integrating Scholia and ORKG for real-time research assessment

Houcemeddine Turki

Slides & videos for this session: <https://is.gd/OEjg8X>
Etherpad: <https://etherpad.wikimedia.org/p/8TFHHG>

WIKIMANIA
SINGAPORE





WD:CRAP – WikiProject-based source reliability ... and WP:2.0

SJ Klein

Slides & videos for this session: <https://is.gd/OEjg8X>
Etherpad: <https://etherpad.wikimedia.org/p/8TFHHG>

WIKIMANIA
SINGAPORE



Wikipedia ~~0.9~~ 1.0

Wikipedia 1.0 Server



Use the autocomplete search box below to search for a project in order to display its summary table.

Spanish Africa

Select Project

| Quality | Importance | | | | | | Total |
|-------------------|------------|------|-----|-----|-----|-----|-------|
| | Top | High | Mid | Low | NA | ??? | |
| ⊕ GA | | | | 1 | | | 1 |
| B | 3 | 1 | 1 | 7 | | 7 | 19 |
| C | 2 | | | 8 | | 19 | 29 |
| Start | 5 | 4 | 5 | 16 | | 142 | 172 |
| Stub | | 2 | 1 | 18 | | 95 | 116 |
| List | 2 | | | | | 5 | 7 |
| Category | | | | | 157 | | 157 |
| Disambig | | | | | 3 | | 3 |
| File | | | | | 1 | | 1 |
| Project | | | | | 2 | | 2 |
| Redirect | | | | | 6 | | 6 |
| Template | | | | | 7 | | 7 |
| Assessed | 12 | 7 | 7 | 50 | 176 | 268 | 520 |
| Unassessed | | | | | | 2 | 2 |
| Total | 12 | 7 | 7 | 50 | 176 | 270 | 522 |

Article quality + importance

- Rating of articles, by WikiProjects
- Article metadata in talk-page categories
- Tools: [WP 1.0 bot](#), OpenZim [server](#)
- Simple report-card summaries showing WikiProject-level progress

Wikipedia 2.0*

Citation quality + coverage

- Rating of citations,, by WikiProjects
- Article-level metadata on talk pages
- Citation-level metadata on reliability in Wikidata / Wikibase / tabular data
- Claim-level metadata on sourcedness*
- Tools: [Citation bot](#), [Credibility bot](#)



**WIKIMANIA
SINGAPORE**

tools

- [What links here](#)
- [Related changes](#)
- [Upload file](#)
- [Special pages](#)
- [Permanent link](#)
- [Page information](#)
- [Expand citations](#)

Wikipedia 2.0

Wikidata:WikiProject Source Reliability

[Shortcut](#): `WD:CRAP`

(Redirected from [Wikidata:CRAP](#))

WikiProject source reliability, aka the **Credibility Ratings + Assessments Project**, is an effort to identify and aggregate online sources of assessments of the reliability and credibility of sources. These assessments include estimates of bias, verifiability of claims, level of editorial oversight or peer review, and expertise in specific topics. Assessments may be of individual documents, or aggregate assessments of authors or domains.

Source reliability data

- [Credibility bot](#) (for aggregate stats)
- [Citations Wikibase](#) (for source reliability and bibliometrics data)
- [Underlay](#) (for versioned tabular data, like perennial sources tables per topic)
- Ex: Vaccine Safety WikiProject
WD:CRAP - Source Reliability

805

articles

5372

domains

3.3%

known reliable sources


0.0%

flagged sources

96.7%

unrated domains

Reliability data based on [/Perennial sources](#) • View more data in [/Reports](#).

Last updated on 19:01, 15 August 2023 (UTC) by  [Credibility bot](#)

Wikipedia 2.0

📍 Maniphest > T337013

! [Epic] Splitting the graph in WDQS

Open, High

Public

Description

In order to stabilize the Wikidata Query Service we are looking into splitting the graph inside Blazegraph into 2 (or potentially more) subgraphs. This ticket is for tracking the investigation into what a sensible split would be, what the consequences are and then making it happen.

Next steps

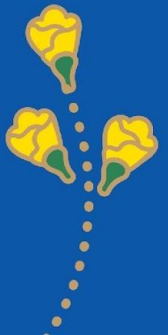
- **Enterprise:** Work w/ Enterprise streams?
- **Wikidata split:** Citation data may be split from Wikidata; help define its home
- **Campaigns:** highlighting how WP's citation density and precision is one of the best models on the Web.
- **Partnersions:** Get source-reliability annotations from large search engines and AI labs for the "cited on WP" corpus
- Large Citation Models?
Automated estimates of source relevance + related sources?
... your ideas here~!
- join WD:CRAP today !

Visualizing Wikidata by using ALEC

Rod Page

Slides & videos for this session: <https://is.gd/OEjg8X>
Etherpad: <https://etherpad.wikimedia.org/p/8TFHHG>

WIKIMANIA
SINGAPORE



W
SI

Final thoughts / discussion

Slides & videos for this session: <https://is.gd/OEjg8X>
Etherpad: <https://etherpad.wikimedia.org/p/8TFHHG>

WIKIMANIA
SINGAPORE

