



*Citations for the sum of all human knowledge.  
Final Report, 2020–2021.*

---

L. Wyatt – *Program manager*

P. Ayers, M. Proffitt, D. Mietchen, A. Stinson, D. Taraborelli – *Steering committee*

A. Bittaker, J. Curiel, J. Tud, C. Virtue – *Wikimedia Foundation staff*

*WikiCite 2020-2021: Citations for the sum of all human knowledge*

<https://doi.org/10.5281/zenodo.5363757> · [CC BY 4.0](#)

# Table of Contents

About	3
Introduction	4
Grant programs	6
WikiCite Virtual conference 2020	15
Responding to the pandemic	17
News from the movement	18
Reflecting Back and Looking Forward: The Impact of WikiCite	22
Financial Report	23



Photo by [Eka343](#), CC-BY-SA 4.0

A collection of [lontar](#) [palm leaf manuscripts]. The digitisation and transcription of Balinese lontar is the subject of one of this year's Wikimedia community grants funded through the WikiCite program.

# About

[WikiCite](#) is an initiative aiming to build **a comprehensive knowledge base of sources**, to serve the sum of all human knowledge. The [Alfred P. Sloan Foundation](#) generously supported the 2018 WikiCite conference and WikiCite activities from 2019 through 2021. This report examines the impact, key milestones, and reach the WikiCite community has achieved over the course of July 2020 to June 2021 – the third and final year of the current grant.

Additional financial and logistical support was provided by the [Wikimedia Foundation](#). Previous annual reports can be found at the [WikiCite homepage on Meta Wiki](#).



Alfred P. Sloan  
FOUNDATION



WIKIMEDIA  
FOUNDATION

*Citations are a simple, critical interconnection mechanism for all modern knowledge in the digital, Internet-connected world. Arguably the most important ingredient of open knowledge, sources and references have ironically received little technical attention in the Wikimedia movement up until now.*

[2016 WikiCite annual report](#)

# Introduction

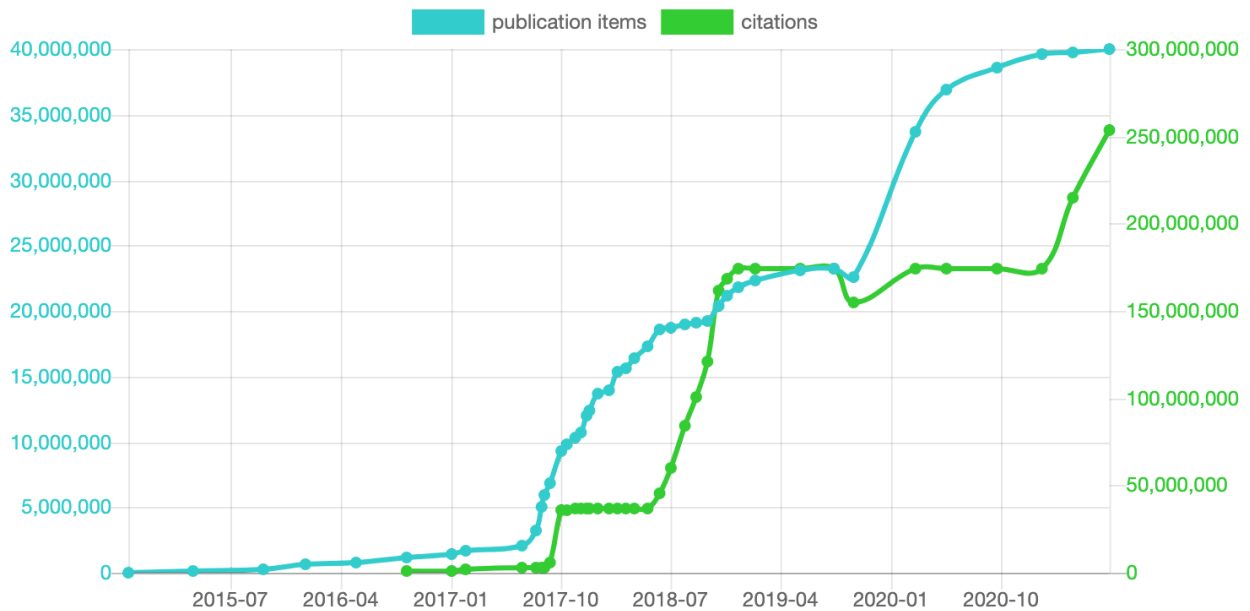
Citations form the basis of how we know what we know. Citing prior work allows new scholarship to acknowledge what has gone before, identify gaps, and extend existing knowledge. On Wikipedia, citations to reliable sources provide the fundamental layer of verification and fact-checking that is required for an accurate and reliable encyclopedia that is read by millions every day. But despite their importance, citations had largely been ignored in discussions of types of open knowledge as well as in Wikimedia’s technical infrastructure.

Over the five years of the WikiCite initiative, that landscape has changed. With the creation of a rich, human-curated, and machine-readable knowledgebase of sources in Wikidata, the WikiCite initiative is crowdsourcing the process of vetting information and its provenance. Beyond Wikidata, initiatives to open citations more broadly have resulted in over a billion citations from one article to another now available in open databases, a major change from just a few years ago.<sup>1</sup>

This year, with the help of a grants program funded by the WikiCite initiative, members of the WikiCite community worked on improving the rich ecosystem of tools that have developed to work with citations on the Wikimedia projects. Community members improved documentation and worked on representing the diversity of the world’s bibliographic formats, going beyond books and articles. WikiCite is a global, multilingual project, and this year’s projects and virtual conference highlight and celebrate that. As Wikipedia celebrates its twentieth anniversary in 2021 with over a billion edits made in total, reliable sources continue to be the core of verifiability, anti-disinformation and knowledge integrity in our movement.

---

<sup>1</sup> B. Ian Hutchins; A tipping point for open citation data. *Quantitative Science Studies* 2021; 2 (2): 433–437. doi: [https://doi.org/10.1162/qss\\_c\\_00138](https://doi.org/10.1162/qss_c_00138) . Wikidata/ Scholia: [Q107912806](https://www.wikidata.org/wiki/Q107912806).



[Wikicite statistics](#) derived from Wikidata information by James Hare et al. [CC0](#)

The number of items in Wikidata representing sources has grown dramatically since the first WikiCite conference in 2016. There are [now ~40 million Wikidata items about individual publications and over 250 million known citations](#) – where one of those known publications references another – making an ever more interconnected knowledge graph. This means that **~43% of Wikidata is now open bibliographic citation information**. With this growth has come a large number of innovative tools to work with this data, to help curate, share and explore it. These tools and experiments have been developed by a community of volunteer Wikimedia developers, librarians, and linked data experts and enthusiasts, and together show the power and potential for this new source of open bibliographic data. The library and GLAM community has shown increasing interest in using Wikidata and Wikibase (the underlying software of Wikidata) in library-related linked data systems, with several experimental projects continuing to be developed, notably by several European national libraries. The WikiCite community uses an [active mailing list](#) (with 288 members) and social media platforms (with over 5,000 followers of the [@WikiCite twitter account](#), and 170 members of the recently created [Telegram discussion group](#)), as well as discussions on Wikidata itself. WikiCite-related projects have been presented and discussed all over the world.

# Grant programs

In August 2020, the WikiCite steering committee [announced](#) a simultaneous call for proposals for two community grant programs – eScholarships and Project grants. These grants, funded by WikiCite, were available to individuals, groups, and organizations with projects that supported WikiCite’s goals. The project grants represented larger-scale activities (between \$2,000 and \$10,000 per project), while the e-scholarship program pioneered a new kind of grant for Wikimedia, created in response to the [2030 Movement strategy recommendations](#) to “provide for safety and inclusion” and “ensure equity in decision making” in an era of COVID-19 quarantines.

An eScholarship provides a per-diem equivalent allowance for 1-5 people to stay at their respective home(s), and work for 2-4 days on a project supporting WikiCite’s mission. eScholarship recipients’ projects could be the kinds of things they might have previously undertaken at an in-person hackathon, unconference, or research trip. While the financial value of these individual eScholarships was small, the introduction of this program was a significant step in **providing equity in grantmaking**. It was the first Wikimedia Foundation program to pay a living allowance, not require expense reports, and be valued according to location and time rather than purchases made (calculated according to the Wikimedia Foundation’s own living allowance calculations for travel to their city). It facilitated remote group-work, and acknowledged that a requirement of pure volunteerism is often a barrier to entry for some – particularly those disproportionately in marginalized knowledge communities. The precedent of this project has now led to the concept of eScholarships being repeated at the most recent Wikimania, and forthcoming Wikidata conferences. The eScholarship recipients contributed valuable citations and citation tools that might not have otherwise been possible.

The grant programs had a strong response from the WikiCite and GLAM communities, [22 groups and individuals across 15 countries received and eventually completed WikiCite grants](#) – the majority of which were outside the OECD. Project categories included content creation & upload, outreach & training, software development, and documentation/localization.

A **full list** of project and event grants funded can be [found here](#). Some **highlights** from these grants are as follows:

### **Balinese WikiLontar** – Indonesia

A program to collect metadata, digitize and catalog Balinese **palm-leaf manuscript** on the island of Bali – and subsequently to add references for Wikipedia, items to Wikidata, images to Commons, and transcripts to Wikisource.

*“Until now **only 108 items** about palm-leaf manuscripts or lontar exist in Wikidata. This project will help to add more palm-leaf manuscripts, create first projects in palm-leaf manuscript catalogues, contribute quality reference into Balinese Wikipedia and upcoming Balinese Wikisource along with Indonesian Wikipedia . . . Our target is lontar which [are] already preserved by many in our society but still not catalogued and digitized due to a lack of people who can read and write Balinese script anymore.”*



Photo by Eka343, [CC-BY-SA 4.0](#)

[Members of the project team documenting the lontar, February 2021](#)

Although disrupted by the pandemic, the initial goals were surpassed with the collection of 600 titles and the creation of 776 Wikidata items. [A 650-page catalogue](#) is being prepared for professional publication ([transcriptions](#), [place of original publication](#)). This catalog will enable them to be specifically cited as references in their own right for the facts that they contain in Wikipedia articles.

### **[Modelling legislation in Wikidata](#)** – Brazil

A process to model and import the entire framework of Brazilian laws (and “law-decrees”) in Wikidata, documenting in detail the process and methodology for future replicability.

*“For the first time, the Brazilian legislation (about 28 thousand laws and decree-laws) is available in an open, structured and semantic platform”*

The project required developing scraper tools to extract data from various sources, “[schema crosswalk](#)” to ingest it to Wikidata, and a schema in Wikidata to model the metadata for each type of legislation. Finally, [a web-app was created](#) that allows anyone to create new Wikidata items based on an entry from the legislative website, and to automatically reconcile the law's subjects to Wikidata.

### **[WikiCite add-on for Zotero with citation graph support](#)** – Argentina

A project to develop a Wikidata plugin for [Zotero](#) is the most popular open source reference management software too, enabling users to understand how publications connect with one another and discover new works. [The resulting software “Cita”](#) adds citations metadata support to Zotero, using information available from Wikidata, and enabling users to easily contribute missing data. Open source, the software is now also fully translated by the community into eight languages.

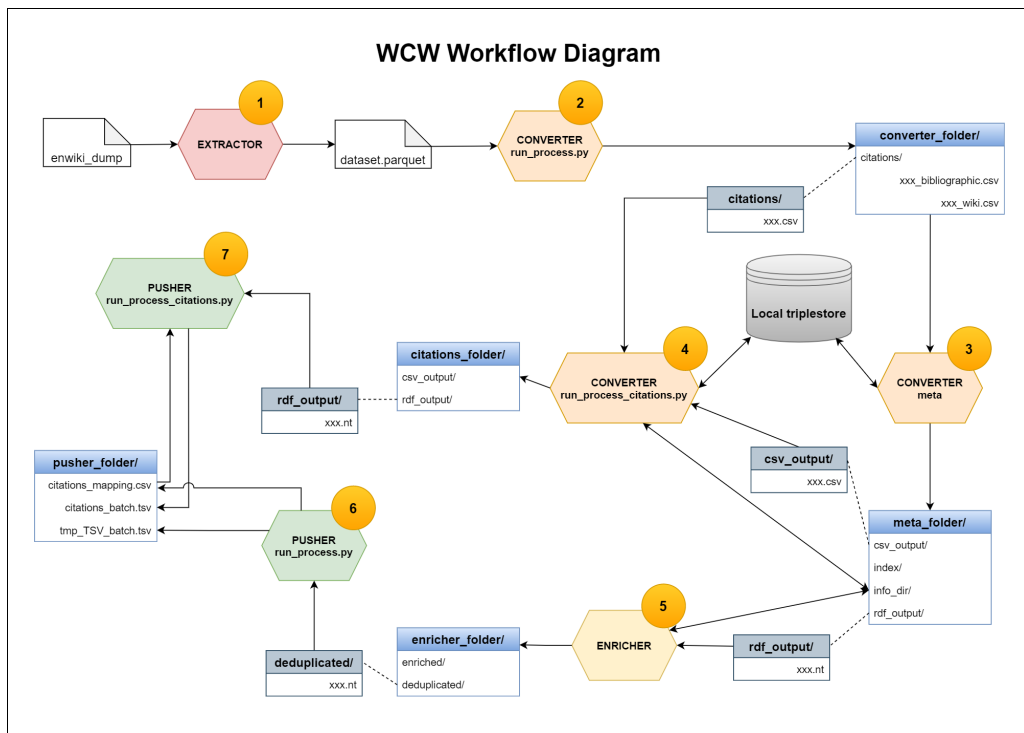
*“Researchers and writers often use reference management software to organize their literature review. These collections tend to grow fast and it is often unclear how the different items relate to one another... The plugin fetch[es] citation information from Wikidata, that lets the user easily fill in the gaps (and upload this information back to Wikidata), and that uses this information to easily show how the items in the user’s collection connect to one another, would expand how both projects talk to each other.”*



On the back of the success of this project, the author [has now obtained separate grant funding](#) to build on the functionality of Zotero’s web translators to increase the user-friendliness of the most used automated reference creation tool in Wikipedia, “Citoid.”

**Wikipedia Citations in Wikidata** – Italy

A project to develop a codebase to enrich Wikidata with citations to scholarly publications that are currently referenced in the English Wikipedia. This **Citations dataset** currently includes around 30 million citations from Wikipedia pages to a variety of sources – of which four million are to scientific publication. This codebase consists of four software modules in Python and integrates new components (a classifier to distinguish citations by cited source and a look-up module to equip citations with identifiers from Crossref or other APIs).



*The Wikipedia Citations in Wikidata workflow: Extractor, Converter, Enricher, Pusher* - by OpenCitations team, [CC-BY-SA 4.0](#)

“These citations in Wikidata would make Wikipedia contents better discoverable and enrich Wikidata with a ready-to-use corpus for further analysis or for developing new services. In addition, Wikimedia projects and GLAM services that already leverage Wikidata knowledge base or alignments to Wikipedia pages, would benefit from having mechanisms that allow discovery of relevant works related to entities described in Wikipedia and distilled in Wikidata.”

### ***Histo Cita-thon*** – Ghana

An event that seeks to educate Wikimedia volunteers and other open communities on how to cite using books, journals and newspapers in creating and improving Wikipedia articles and Wikidata items. Specifically focusing on the Independence era.

*“Previous attempts by the Wikimedia community in Ghana to help increase content on the pre-independence history have been met with difficulty due to poor citations.” “At the end of the contest, 25 people participated, 43 articles were created, 233 articles were improved, 228 references were improved.*



*Photo* by NanaYawBotar, CC-BY-SA 4.0

*Participants of the Histo-Cita-Thon training event, Ghana*

### ***Research Records of Tāmaki Paenga Hira*** – New Zealand

This project engaged a Wikimedian in Residence (WiR) help unlock the potential knowledge held within Tāmaki Paenga Hira’s – the Auckland Museum – academic outputs and research publications.

*“The primary focus [is] loading the 450+ articles of the **Records of the Auckland Museum** into Wikidata as well as other publications that the museum holds the copyright for, such as **AWB Powell’s Native Animals of New Zealand**. This material contains all the interdisciplinary research*

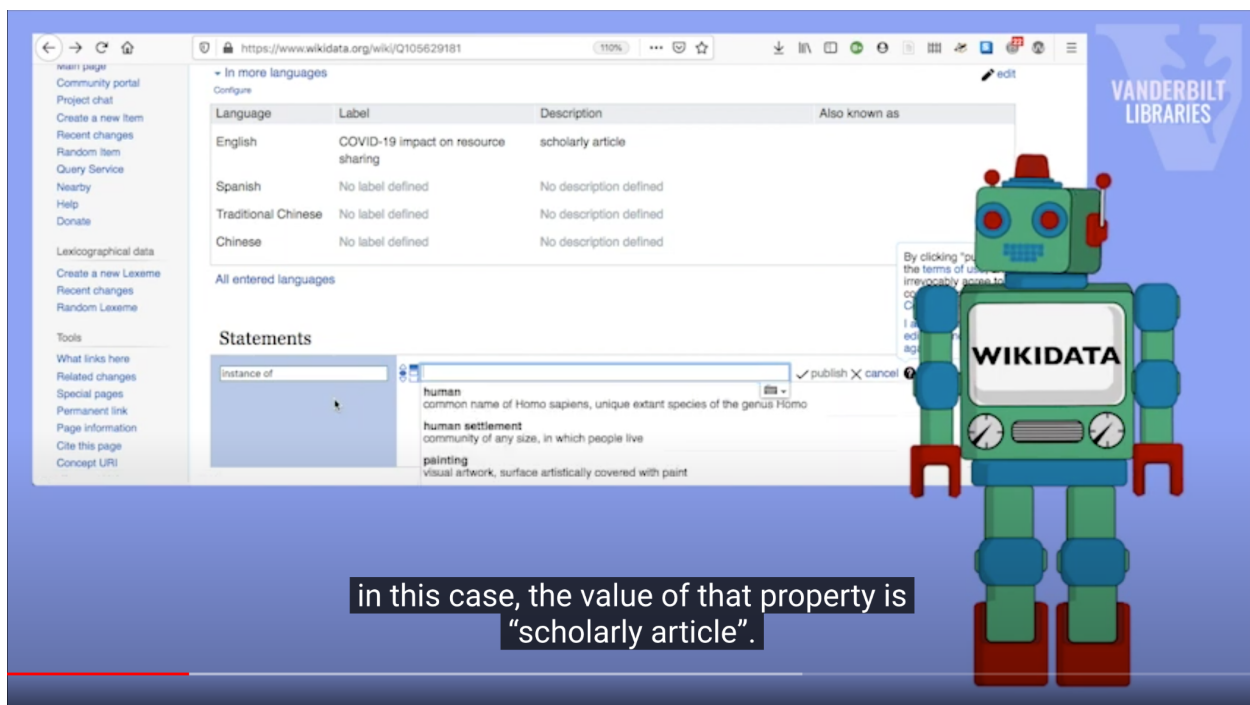
the Museum has undertaken in the last 165 years; including world leading research on New Zealand’s biodiversity, Mātauranga Māori, the wider Moana Pacific.” “Every article in the Records, from 1930 to the most recent issue, now has a **Wikidata entry**. 49 new Wikipedia articles were created, and 882 references were added to existing articles, enhancing the quality of information on Wikipedia and further dispersing the Museum’s research output. The topics covered, such as **Pitcairn Island**, **Niue**, **Adzes**, the **Yellow-bellied sea snake**, and **tapa cloth** cover the range of the Museum’s collecting focus on Aotearoa and the Pacific.”

### **Interactive Learning Pathways for Information Professionals** – USA

A project to create a series of multimedia learning pathways designed to help beginners install and use Wikidata tools and get started editing – using the interactive story software.

“The goal of the project is to develop and implement interactive learning pathways for information professionals to learn the fundamentals of WikiCite . . . few of the Wikidata training resources are interactive or allow for guided exploration of beginning Wikidata topics.”

The team scripted, produced, and transcribed 60 videos, 20 videos each in Chinese, English, and Spanish. The open educational resource is now publicly available at [LearnWikidata.net](https://www.learnwikidata.net).



The image shows a screenshot of a Wikidata edit page. The browser address bar shows the URL <https://www.wikidata.org/wiki/Q105629181>. The page title is "COVID-19 impact on resource sharing". The "In more languages" table is visible, showing the English label and description. The "Statements" section shows a property being edited, with a dropdown menu open showing options like "human", "human settlement", and "painting". A blue robot character with "WIKIDATA" on its chest is overlaid on the right side of the page. A text box at the bottom of the screenshot reads: "in this case, the value of that property is 'scholarly article'".

Language	Label	Description	Also known as
English	COVID-19 impact on resource sharing	scholarly article	
Spanish	No label defined	No description defined	
Traditional Chinese	No label defined	No description defined	
Chinese	No label defined	No description defined	

Statements

instance of

- human: common name of Homo sapiens, unique extant species of the genus Homo
- human settlement: community of any size, in which people live
- painting: visual artwork, surface artistically covered with paint

Screenshot of “[editing a journal article item](#)”, Vanderbilt University, [CC-BY-SA 4.0](#)

## ***Bibliography and citations of Hausa folklore*** – Nigeria

This project endeavored to provide reliability to **Hausa Wikipedia** articles and Wikidata by providing references and making accurate citations to articles that are related to Hausa folklore, and related biographies, culture and religion.

*“Currently, many articles in Hausa Wikipedia are lacking citations and bibliography, almost 70% of the articles lack references, citations and bibliography, while the remaining 30% lack reliable sources.” “At the end of the project almost 50 new articles were created, and 187 articles were improved by four Wikimedians in the space of four to five days.”*



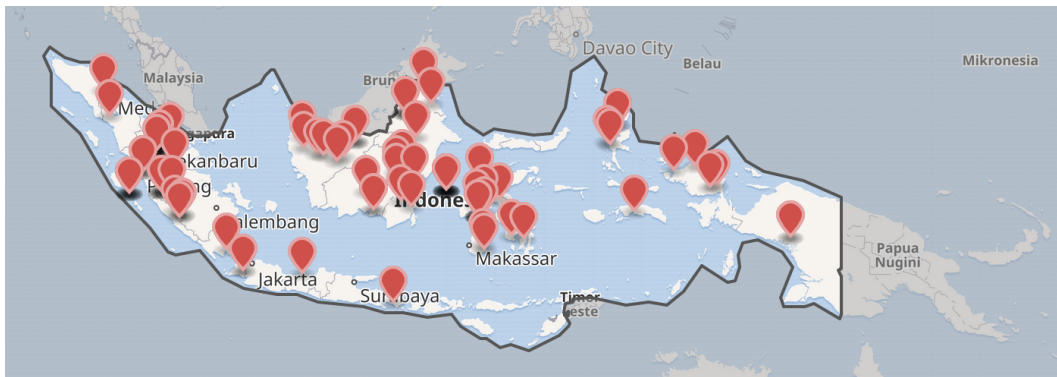
Photo by Anasskoko, [CC-BY-SA 4.0](#)

*[Participants of the editathon, using in the library collection of Hausa Folklore reference materials.](#)*

## eScholarships

**Wiki Masyarakat Adat**, awarded to *Wirjadisastra, Elicefa, Moentatoz, Afalranggajati, and RXerself* – Indonesia.

A project to collect Indonesian local regulations about indigenous groups across local governmental websites into Wikimedia projects. “Indonesian society is comprised of hundreds of indigenous groups . . . However, there has not been any national-wide law regulating specifically the indigenous people, especially about their ancestral land and cultural rights.” The project resulted in the cataloguing of 128 distinct laws across various jurisdictions in both Wikimedia Commons ([as images of documents](#)) and in Wikidata as structured information.



Map [CC-BY-SA 4.0](#). Data available under ODC [Open Database License v1.0](#).

[A Wikidata query for the location of publication of the laws pertaining to indigenous rights.](#)

**Cite Q improvements**, awarded to *Mike Peel, Andy Mabbett, RextxS, Adamant.pwn, and Ederporto* – of Spain, UK, Germany and Brazil.

A project to rewrite the *Cite Q* reference template on English Wikipedia, so that it uses contemporary best practices - using Wikimedia modules in *Lua*. Simultaneously improving the functionality and enabling it to be used in other languages more easily. Furthermore, this eScholarship featured a multinational team, emphasizing the benefits and flexibility of the funding program.

### Examples

```
{{cite q|Q15625490}}
```

Jeffrey T. Williams; Kent E. Carpenter; James L. van Tassell; Paul Hoetjes; Wes Toller; Peter Etnoyer; Michael Smith (21 May 2010). "Biodiversity Assessment of the Fishes of Saba Bank Atoll, Netherlands Antilles". *PLoS ONE*. **5** (5). doi:10.1371/JOURNAL.PONE.0010676. ISSN 1932-6203. PMC 2873961. PMID 20505760. Wikidata Q15625490.

```
{{cite q|Q15625490|page=42}}
```

Jeffrey T. Williams; Kent E. Carpenter; James L. van Tassell; Paul Hoetjes; Wes Toller; Peter Etnoyer; Michael Smith (21 May 2010). "Biodiversity Assessment of the Fishes of Saba Bank Atoll, Netherlands Antilles". *PLoS ONE*. **5** (5): 42. doi:10.1371/JOURNAL.PONE.0010676. ISSN 1932-6203. PMC 2873961. PMID 20505760. Wikidata Q15625490.

```
{{cite q|Q15625490|access-date=18 May 2017}}
```

Jeffrey T. Williams; Kent E. Carpenter; James L. van Tassell; Paul Hoetjes; Wes Toller; Peter Etnoyer; Michael Smith (21 May 2010). "Biodiversity Assessment of the Fishes of Saba Bank Atoll, Netherlands Antilles". *PLoS ONE*. **5** (5). doi:10.1371/JOURNAL.PONE.0010676. ISSN 1932-6203. PMC 2873961. PMID 20505760. Wikidata Q15625490. Retrieved 18 May 2017.

```
{{cite q|Q15625490|quote=lorem ipsum}}
```

Jeffrey T. Williams; Kent E. Carpenter; James L. van Tassell; Paul Hoetjes; Wes Toller; Peter Etnoyer; Michael Smith (21 May 2010). "Biodiversity Assessment of the Fishes of Saba Bank Atoll, Netherlands Antilles". *PLoS ONE*. **5** (5). doi:10.1371/JOURNAL.PONE.0010676. ISSN 1932-6203. PMC 2873961. PMID 20505760. Wikidata Q15625490. "lorem ipsum"

Examples of the “Cite Q” Template allows for highly esoteric reference information to be included in a standardized manner in a Wikipedia footnote, simply by calling upon the Wikidata item number for the desired reference work. ([Examples, CC-BY-SA 3.0](#))

***[Tools to add DOI information to Chinese articles](#)***, awarded to *Stevenliuyi* – USA

In recent years, China has become **the world’s largest producer of scientific articles**. However, bibliographic data of Chinese scientific articles are still very limited on Wikidata, because most commonly used bibliographic databases on Wikidata, such as Crossref, do not include articles published in most Chinese academic journals. There is no available database containing comprehensive DOIs for Chinese articles, and many DOIs can only be found on journals’ official websites. This project developed **[an open source scraper, validation, matching, and generation toolset](#)** to collect those data scattered across different journal websites. During the development, **more than 20,000 DOIs** have now been added to Wikidata.

***[Documentation of workflows for the ingestion of bibliographic data into Wikidata](#)***, awarded to *Walkuraxx* – Netherlands.

The project, based on the workflow employed by the African Studies Centre Leiden, created screencasts and documentation for the process of adding articles via Zotero and QuickStatements, and books via OpenRefine and QuickStatements. The ASCL workflow is applicable and replicable by other institutions, as its constituents (e.g., bibliographic records in MARC 21) are used in many libraries.

# WikiCite Virtual conference 2020

From 2016 to 2018, an annual WikiCite conference was held to support this community: [Berlin 2016](#), [Vienna 2017](#), and [Berkeley 2018](#). In early 2020, the format was adjusted to fund “satellite events,” large and small, around the world. As the COVID-19 pandemic developed, these satellite events had to be re-imagined again, as detailed in the 2020 WikiCite annual report.<sup>2</sup>



*The 2020 Virtual conference logo, in association with [Wikidata's 8th birthday and broader series of events](#).*

In October 2020, a [virtual WikiCite community conference](#) was held. This was a multilingual and global event that spanned **three days, with 82 speakers, hosted in five languages** (English, French, German, Indonesian, and Portuguese) and was streamed across nine platforms at times appropriate for all timezones. The success of this conference – one of the first fully online and parallel-track events in the Wikimedia community – served as inspiration for the first fully online Wikimania conference, which was held in August 2021.

Sessions included hands-on technical workshops on tools and library workflows, as well as discussion and demonstration sessions on modelling and importing specific types of sources, mapping identifiers, conducting scientometrics, and connecting Wikidata to the other Wikimedia projects. Session videos are archived on Wikimedia Commons for future viewing.

---

<sup>2</sup> See [WikiCite annual report 2019-2020](#). p4.

To support the breadth of the conference, two of the eScholarships were also awarded to session hosts of the virtual conference – covering [Estruturação de projetos no Wikidata](#) (in Portuguese) and the [OpenVirus research project](#). Other sessions included such diverse issues as *Citations in Swedish Parliamentary documents*, *Research output items*, and *Advancing Librarianship and Scholarly Communications*.

## Responding to the pandemic

In mid-March 2020, [the Wikimedia Foundation sent notice](#) to all its affiliates and grant recipients that, subsequent to the WHO declaration of a global coronavirus pandemic, all in-person events must be cancelled or postponed until *at least* mid-September of that year. This was later expanded such that no *new* grants for in-person events would be approved until 2021.

Continuing the precedent and health regulations begun during that year, no international in-person conferences were held during 2020-21 – including those in which WikiCite (and related topics) usually hold prominence – WikidataCon and Wikimania.

Depending on the local regulations enforced at the time, the Wikimedia Foundation supported *local* in-person events with the creation of a comprehensive “[risks & mitigations assessment protocol](#)”. This system allowed a cautious and highly contextual approach to the approval of local in-person events. In addition to developing the grant program and virtual conference, in 2020, we worked with communities and groups that had previously been approved to host events with WikiCite funds in 2020-21 to develop alternate plans. Several events that had been previously approved as “Satellite events” under the 2019-20 WikiCite program were paused, and after further risk assessment review, the funds returned.

The International Federation of Library Associations (IFLA) *Wikidata working group* had previously been awarded a grant to produce Wikidata workshops during their annual conference to be held in Dublin, Ireland. With in-person events cancelled, they successfully converted their grant to instead produce a series of long-form training videos, to be released at the end of 2020 and with a smaller budget. As the grantee described:



*While the delivery formats of the event have changed, the theme of the event remains the same, that is to support and activate the professional library communities' efforts to foster knowledge equity through the development of open bibliographic data to increase the visibility of knowledge and intellectual traditions that remain marginalized on an internet dominated by certain cultures as well as decrease barriers to linked open data in libraries around the world.*

*The digital resources to be developed and shared will be for sharing 'best practices,' tools, and techniques which could further help communities share knowledge by overcoming technical and resource constraints and become contributors themselves.*

*Proposed activities keep close to the originally proposed activities, consisting of a mix of workshops and talks. However, we proposed to deliver content in an asynchronous format to better support participation of a distributed, international community. Packaging material this way also removes dependence on a stable, continuous internet connection.*

The resulting [IFLA video discussion series](#) of six sessions was uploaded as both a [YouTube playlist](#) and on [Wikimedia Commons](#) – with videos in English and professionally transcribed/translated subtitles in English, French, Spanish, Arabic, Chinese, and Portuguese. The topics focused on issues of diversity and equity of access to Open Citation information such as “Critical issues in Knowledge Equity”, “Scholarly Profiles, Wikidata and Academic Libraries” and “How Wikidata complements and completes metadata work”.

Like many other communities, all Wikimedia projects were [engaged with the COVID-19 response in some way](#). The efforts of [WikiProject COVID-19](#) on Wikidata built on previous initiatives like [WikiProject Zika corpus](#) and have been described in a [paper](#) accepted for publication in the Semantic Web Journal. The WikiCite community contributed to these efforts in various ways, e.g. by curating [Wikisource entries](#) about pandemic-associated policy documents or by [enriching Scholia profiles](#) of researchers or institutions working on various aspects of the virus, the disease, the pandemic, policy responses or vaccination.

## News from the movement

The WikiCite community grants and conference are only a small part of the story. Over the past year, the WikiCite community, partners, and the professional and volunteer contributors to Wikidata (and its underlying software Wikibase) were part of a number of important events, publications, and projects, a sampling of which are noted below.

## Events and publications

- [Wikimania 2021](#) – the first wholly virtual edition of the annual flagship conference – featured multiple WikiCite related presentations, including “WikiCite: Recent achievements, what happens next?” ([Video](#)) and “Automatically maintained citations using Wikidata” ([Video](#)), as well as many sessions pertaining to libraries/library science.
- [2021 LD4 Conference on Linked Data](#) for Libraries, featured more than a dozen sessions focused on Wikidata and Wikibase, from introductory workshops to advanced matters such as “Standard Citation Forms for Rare Materials Cataloging in Wikidata” at Yale Library, and “The Program for Cooperative Cataloging Wikidata” at NYU.
- [Wikimedia Hackathon 2021](#), the [Wikidata Workshop 2021](#) and the [WikiWorkshop 2021](#) each featured participation from many members of the extended WikiCite community.
- [The Invisible Citation Commons: Unsolved challenges and future directions for open citations](#), discussing WikiCite and the open citations landscape generally, was co-authored by Steering Committee member Phoebe Ayers and SJ Klein, both former members of the Wikimedia Foundation Board of Trustees. The essay was published in [Commonplace](#) as part of a collection called "[The Business of Knowing](#)" about reimagining scholarly communication business models, and frames open citations as a critical but understudied part of the open access landscape.
- [A tipping point for open citation data](#) by Ian Hutchins, and [further elaborated by OpenCitations](#) themselves, recognizing the significant milestone of over 1 billion citations (more specifically, DOI-to-DOI citation links derived from open references within Crossref) now available in public databases. Noting:

“The competitive benefits of closing access to citation data diminish with each new citation released to the public domain, but the benefits of open data remain. Going forward, citation data is almost completely public domain.”
- [Verifiability of Wikimedia](#) by Sarah Gelb-Wiegand at the Spiegel Institut discusses what is well-cited content, the research that goes into creating it, and how it is incorporated across the Wikimedia projects.
- [Wikipedia citations: A comprehensive data set of citations with identifiers extracted from English Wikipedia](#) by Harshdeep Singh, Robert West, Giovanni Colavizza provided a dataset of 29M citations extracted from the English Wikipedia, representing one of the most comprehensive overviews of what is cited in Wikipedia to date.
- [Measuring the quality of scientific references in Wikipedia: an analysis of more than 115M citations to over 800 000 scientific articles](#) by [Joshua M. Nicholson](#), [Ashish Uppala](#), [Matthias Sieber](#), [Peter Grabitz](#), [Milo Mordaunt](#), and [Sean C. Rife](#) used a novel technique to analyze the quality of articles cited in Wikipedia.

## Projects

The global volunteer and professional communities that contribute to and use Wikidata and Wikibase – both directly and through third-party platforms – continue to grow, and so does the density and quality of data relating to citations. A few recent projects with connections to WikiCite worthy of note include:

- [\*Shared Citations\*](#), a proposal for a new Wikimedia Foundation hosted database of structured citation information that could be used in any Wikimedia page, was developed under the auspices of the WikiCite program throughout the year. Following extensive qualitative research and stakeholder interviews, a proposal was developed for an in-house citation management system that would permit all Wikimedia projects to centralize and re-use the millions of footnotes and references across all projects as structured data, with concomitant benefits for usability, data quality and consistency, and knowledge equity. Widespread community public endorsement was obtained for the proposal, which will require considerable software development and re-architecting of the MediaWiki platform. This proposal was based on synthesized input from WikiCite community members and the broader Wikimedia community, technical contributors, and years of discussion about Wikimedia projects' in-house needs for improved citation management. The Shared Citations proposal was not prioritized in this year's annual funding for the Wikimedia Foundation, but was well received by WMF leadership and community members, and we hope it will be a good candidate for future technical focus if the opportunity presents itself. Investigations on whether elements of that work can be undertaken during the current financial year are underway.

## Ileal-lymphoid-nodular hyperplasia, non-specific colitis, and pervasive developmental disorder in children (C44583) [English]

<b>Citation</b>	<a href="#">APA</a>   <a href="#">MLA</a>   <a href="#">MHRA</a>   <a href="#">Chicago</a>   <a href="#">CSE</a>   <a href="#">Bluebook</a>   <a href="#">AMA</a>   <a href="#">BibTeX</a>   <a href="#">wiki</a>
Author	String as published: "Dr AJ Wakefield, FRCS" (Q508568) First: Andrew Last: Wakefield Initial: J
URL	<a href="https://www.thelancet.com/journals/lancet/article/PIIS0140-6736(97)11096-0/fulltext">https://www.thelancet.com/journals/lancet/article/PIIS0140-6736(97)11096-0/fulltext</a>
Archive URL	<a href="https://web.archive.org/web/*/https://www.thelancet.com/journals/lancet/article/PIIS0140-6736(97)11096-0/fulltext">https://web.archive.org/web/*/https://www.thelancet.com/journals/lancet/article/PIIS0140-6736(97)11096-0/fulltext</a> Saved 207 times between March 30, 2009 and November 18, 2020.
Access	Closed repository [ <a href="#">link</a> . Access via <a href="#">The Wikipedia Library</a> ]; OA Preprint [ <a href="#">link</a> ]; Shadow Library [ <a href="#">link</a> ]
Source type	<b>Scientific article</b> – Primary source
Publication date	28 February 1998
Publication status	Retracted. Date: 6 February 2010
<b>Citelinks</b>	0 current (Former links: 150 across 27 wikis <a href="#">expand</a> )

New with Shared Citations

## Vindication: A Life of Mary Wollstonecraft (C60384) [English]

<b>Citation</b>	<a href="#">APA</a>   <a href="#">MLA</a>   <a href="#">MHRA</a>   <a href="#">Chicago</a>   <a href="#">CSE</a>   <a href="#">Bluebook</a>   <a href="#">AMA</a>   <a href="#">BibTeX</a>   <a href="#">wiki</a>
Author	String as published: "Lyndall Gordon" First: Lyndall Last: Gordon Initial: L (Q6708612)
Publisher	String as published: "Harper Perennial" (Q5663419)
Source type	<b>Book</b> – Secondary source
Date/Edition	<b>2005. 3rd, Hardcover.</b> (4 other editions cited across Wikimedia - with 12 citations)
Edition of	Wikidata item of overarching work: <a href="#">Q123456789</a>
<b>Citelinks</b>	Total: 105. Current: 98.
By time	First: 11 October 2007, <a href="#">User:CaptainSaru</a> , on English Wikipedia - <a href="#">Feminist philosophy</a> (104 other links <a href="#">expand</a> )
by section	Page 34: <a href="#">Mary Wollstonecraft</a> on en.wikipedia.org (38 more links to page 34 <a href="#">expand</a> ) Pagerange 56-60: <a href="#">Mary Wollstonecraft</a> on de.wikipedia.org (9 more links to page 56-60 <a href="#">expand</a> ) Chapter 3: <a href="#">Feminist philosophy</a> on en.wikipedia.org (12 more links to Chapter 3 <a href="#">expand</a> )
by project	En.Wikipedia.org (9 links <a href="#">expand</a> ) Wikidata.org (54 links <a href="#">expand</a> )

New with Shared Citations

[Examples](#) by Liam Wyatt, CC-BY-SA

Wireframes with examples from the [Shared Citations](#) proposal

- [The Wikilibrary Manifesto](#), an advocacy campaign of Wikimedia Deutschland, and supported by dozens of National and institutional libraries. Its aim:

“Our vision is to create a reliable, machine-readable, collaboratively maintained Linked Open Data network for the arts, culture and science as a solid base for FAIR knowledge. Yet fully applying the [FAIR data](#) principles to knowledge requires a shared framework. The signers of the Wikimedia Library Manifesto will work together to achieve the vision of a common knowledge graph embedded in a shared framework...”

- The [InternetArchiveBot](#) has now received community permission to operate on all Wikipedia editions. This high-impact and high-volume tool extracts URLs used in Wikipedia references to ensure they do not suffer link-rot by automatically including them in the Internet Archive’s Wayback Machine, and re-introducing the archive URL to the Wikipedia reference. By obtaining this editorial right, the linguistic diversity and speed of the reliable source websites that are archived will be greatly improved.
- [Integration of scholarly databases with Wikidata](#) is ongoing, and collaborative curation of the associated literature via Wikidata is becoming more common. For instance, three-way links between natural products, the species they were found in and the publications about this have been [implemented](#) via the Natural Products subgroup of Wikidata’s WikiProject Chemistry, and a similar integration process [has begun](#) for hypotheses described in the invasion biology literature, overseen by WikiProject Invasive Species.

## Reflecting Back and Looking Forward: The Impact of WikiCite

This document is the final in a five-year series of annual reports. Prior reports can be found here: [2016-17](#); [2017-18](#); [2018-19](#); [2019-20](#).

Over the course of these years, the WikiCite program in the Wikimedia Foundation has received financial support from OCLC; CrossRef; ORCID; Gordon and Betty Moore Foundation; Simons Foundation; and the Alfred P. Sloan Foundation. In the last three years, the Alfred P. Sloan Foundation and the Wikimedia Foundation have been the sole financial sponsors of the program. This document represents the third, and final, annual report for the three-year support grant from the Alfred P. Sloan Foundation to the Wikimedia Foundation – listed as [grant number 8518](#) in the sub-program for *Scholarly Communication*.

Over the course of the three years of this latest WikiCite grant, two community conferences have been held – in Berkeley and virtually – continuing the momentum of the first two in Berlin and Vienna. These events brought together a diverse and passionate community of Wikimedia developers, Wikidata enthusiasts, open science advocates, sociologists and humanists, librarians and archivists, and Wikipedians. Conference participants did hands-on work to model and develop citations on Wikidata and held big-picture discussions that helped shape how citations are treated in the Wikimedia ecosystem. The focus on a single topic at these events allowed them to be extraordinarily productive. Wikimedians were able to discuss and collaborate with “fellow traveler” organizations like Crossref and the Internet Archive, leading to fruitful projects that have lasted beyond the conferences. In 2019, the focus shifted from a central event to supporting distributed and local events, to make the WikiCite experience more equitable (since not all members of our community are able to travel to a central conference). This resulted in creative proposals from all over the world. As the COVID pandemic developed, many of these were shifted to online events and trainings as the 2020-21 project-grant and eScholarship programs were developed.

-----

WikiCite’s influence has extended beyond the Wikimedia projects; WikiCite projects have been presented and discussed at events all over the world, particularly in the library community. Projects at national libraries, library organizations and individual institutions have explored cross-walks to local knowledge bases, storing archival metadata in Wikidata, and more, with the goal of making bibliographic metadata about works and collections more accessible, open and linkable. Alongside the growth of bibliographic metadata on Wikidata, Wikidata is increasingly seen as a shared, trusted and open platform that can provide infrastructure for open citations. The WikiCite community remains vibrant and active, with technical discussions occurring daily and enthusiasm for new projects and bibliographic data on Wikidata continuing to grow. **The future for open citations is bright.**