

WikiCite

Abstract

WikiCite is an ongoing effort to build an extensive bibliographic database in Wikidata to serve all Wikimedia projects. The technology needed to sustain this effort is maturing, the modeling of a schema to represent scholarly article metadata is relatively stable, whereas similar schemas for other forms of publications are being worked on as well. To the extent possible, these modeling efforts at Wikidata build on existing controlled vocabularies and map to external identifiers. WikiCite participants and other Wikidata users have catalogued over 8 million scientific articles on Wikidata, along with over 36 million citations between them. These articles have begun to be used as references for Wikidata statements and could be used to improve the handling of references elsewhere, both on other Wikimedia projects and beyond.

Various tools and platforms have been created that allow a rapid integration of large numbers of references into the WikiCite corpus. Currently, there are various reference corpora that have been assembled on Wikidata using these frameworks. Examples are all references from Wikipathways, which is an open database of biological pathways maintained by and for the scientific community all references from the CIViC database, which captures clinical interpretations of variants in cancer.

Furthermore, efforts are ongoing to cover all citations from Reactome - another open repository of biological pathways - and to keep track of the literature about the Zika virus.

While the main purpose of including these references in WikiCite is to be able to provide evidence for statements in Wikidata, the references can also be queried via Wikidata's SPARQL endpoint, which provides new ways of browsing the literature or related information like co-authorship networks.

Aims

In this tutorial, participants will learn about:

- the basics of editing Wikidata
- bibliographic information in Wikidata
- tools that allow to automate certain kinds of edits
- schemas employed by these automated tools
- mechanisms for integrating Wikidata with external biomedical resources
- the basics of querying Wikidata via its SPARQL endpoint

Prerequisites

Participants are expected to create a Wikidata account at least one week prior to the tutorial, and to attend it with their own mobile devices.

Logistics

Further materials will be made available via

https://www.wikidata.org/wiki/Wikidata:WikiCite/SWAT4LS_2017_tutorial .