chem-bla-ics

CiTO updates: Wakefield and WikiPathways

Published August 7, 2024

Citation

Willighagen, E. (2024, August 7). CiTO updates: Wakefield and WikiPathways. *Chem-bla-ics*. https://doi.org/10.59350/8c1e7-8yp77

Keywords

Cito, Wikipathways, Wikidata

This summer I am trying to finish up some smaller projects that I did not have time for to finish, with mixed successes. I am combing this with a nice Dutch staycation, and I already cycled in Overijssel and in south-west Friesland and learning about their histories. But this post is about an update on my Citation Typing Ontology use cases. And I have to say, a mention by Silvio Peroni is pretty awesome, thanks!

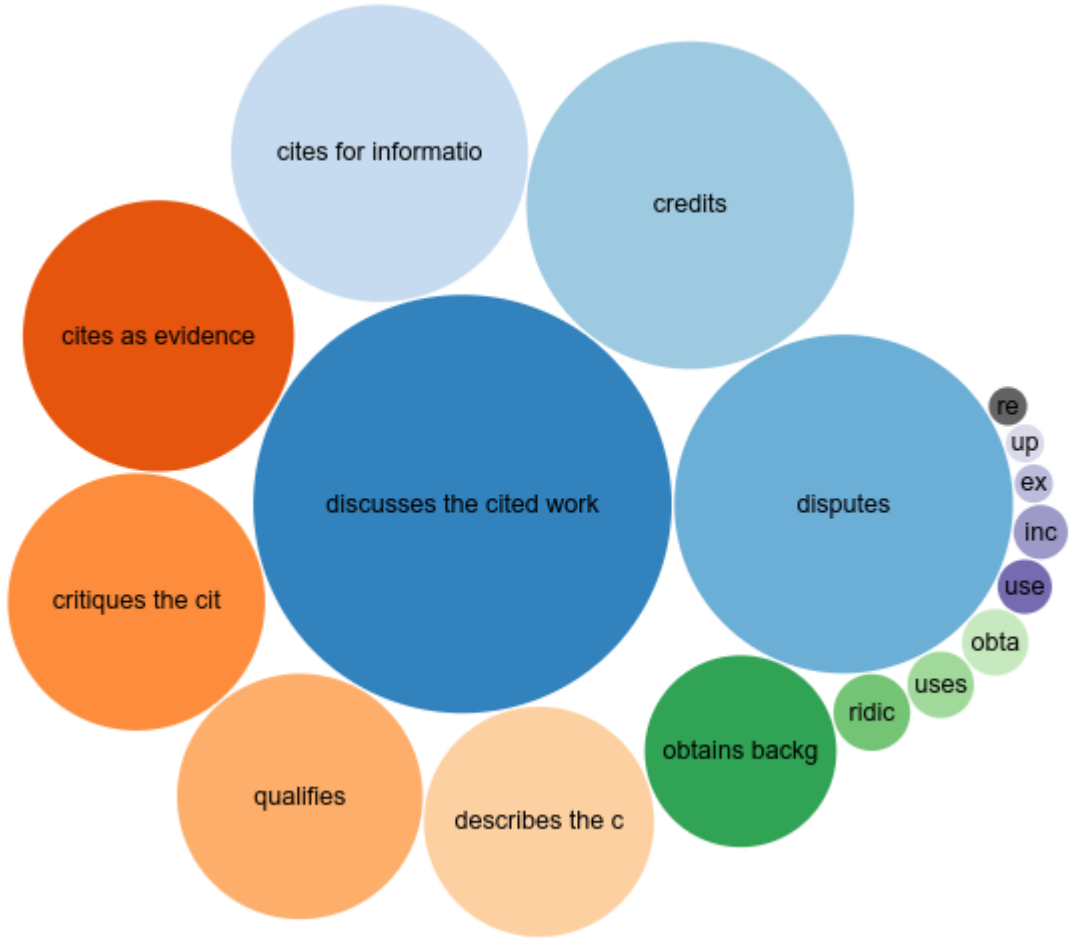
First, the bad news. I still did not get around to the following to tasks I have. First, I need to write up a step-by-step guide how to create CiTO nanopublications and matching draft article. Second, I still need to work out how to update the JATS workflow for CiTO annotation in BioHackrXiv.

Wakefield

Let's first start with a dataset. Peroni mentioned a study they did (10.1007/S11192-021-04097-5) into why the famous Wakefield paper (doi:10.1016/S0140-6736(97)11096-0) is cited. They published their data set on Zenodo (doi:10.5281/zenodo.13166142) with CCZero, so I imported it into Wikidata. Well, at least the citations of articles already in Wikidata. I used a Bacting (doi: 10.21105/joss.02558) script and it actually was quite short. In the end, this added some 500 new citation intentions to Wikidata, now at almost 2000. This is also the third dataset with explicit CiTO intention annotations (see also this post).

This is what the CiTO section of the Wakefield paper in Scholia (doi: 10.1007/978-3-319-70407-4_36) now looks like:

chem-bla-ics



WikiPathways

A second thing I want to show is a potentional CiTO intention annotation dataset. Almost two years ago Alex Pico started a new WikiPathways feature as part of the new website (doi:10.1093/NAR/GKAD960)): a list of citations to specific pathways (in WikiPathways). Alex' setup is fully automated and using PubMed Central and find mentions in figure captions:

Beyond citations to previous WikiPathways journal articles, we have identified 1228 mentions of a total of 582 unique WikiPathways pathway model identifiers, e.g. WP4846, in PubMedCentral articles over the past 13 years.

The file format is a pretty basic YAML file:

chem-bla-ics

main wikipathways-database / downstream / citedin_lookup.yml

egonw Added another article ✓

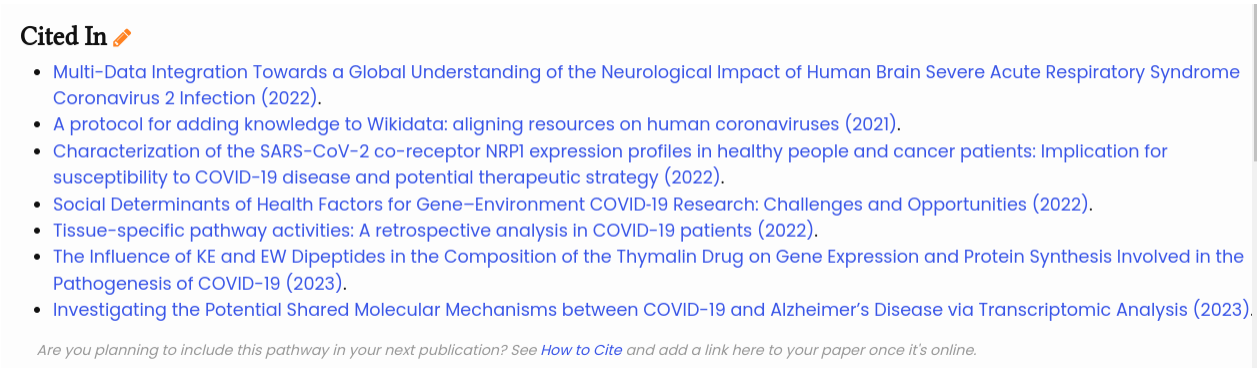
Code Blame 5606 lines (5605 loc) · 256 KB

```yaml
---
last_run: 2024/08/04
WP1:
- link: PMC5075206
  title: Hepatic transcriptome implications for palm fruit juice deterrence of type
    2 diabetes mellitus in young male Nile rats (2016)
- link: PMC4723140
  title: Advanced Running Performance by Genetic Predisposition in Male Dummerstorf
    Marathon Mice (DUhTP) Reveals Higher Sterol Regulatory Element-Binding Protein
    (SREBP) Related mRNA Expression in the Liver and Higher Serum Levels of Progesterone
    (2016)
- link: PMC4546821
  title: Automatically visualise and analyse data on pathways using PathVisioRPC from
    any programming environment (2015)
WP10:
- link: PMC8917653
  title: Transcriptional, epigenetic and metabolic signatures in cardiometabolic syndrome
    defined by extreme phenotypes (2022)
WP100:
- link: PMC8635790
  title: Selenotranscriptome Network in Non-alcoholic Fatty Liver Disease (2021)
- link: PMC8418865
  title: 'Copy Number Variants Captured by the Array Comparative Genomic Hybridization
    in a Cohort of Patients Affected with Hereditary Colorectal Cancer in Sri Lanka:
    The First CNV Analysis Study of the Hereditary Colorectal Cancer in the Sri Lankan
    Population (2021)'
- link: PMC8155553
  title: |-
    Heterogeneity
    of Lipid and Protein Cartilage Profiles
    Associated with Human Osteoarthritis with or without Type 2 Diabetes
    Mellitus (2021)
- link: PMC4761937
  title: Complementary Post Transcriptional Regulatory Information is Detected by
    PUNCH-P and Ribosome Profiling (2016)
WP103:
- link: PMC9329822
```

Additional mentions are found in the main text and tables in the article. These are not always picked up. These can be added manually. Over the past months and the past two weeks particularly, I have been adding additional mentions, not listed yet. We now passed 1500 mentions but I cannot easily give the other statistics.

BTW, anyone can add these citations with the ‘edit’ pencil and some Microsoft GitHub editing (but as far as I am concerned, please feel free to also just mention the paper on the WikiPathways Community Forum):

chem-bla-ics



Cited In

- Multi-Data Integration Towards a Global Understanding of the Neurological Impact of Human Brain Severe Acute Respiratory Syndrome Coronavirus 2 Infection (2022).
- A protocol for adding knowledge to Wikidata: aligning resources on human coronaviruses (2021).
- Characterization of the SARS-CoV-2 co-receptor NRP1 expression profiles in healthy people and cancer patients: Implication for susceptibility to COVID-19 disease and potential therapeutic strategy (2022).
- Social Determinants of Health Factors for Gene–Environment COVID-19 Research: Challenges and Opportunities (2022).
- Tissue-specific pathway activities: A retrospective analysis in COVID-19 patients (2022).
- The Influence of KE and EW Dipeptides in the Composition of the Thymalin Drug on Gene Expression and Protein Synthesis Involved in the Pathogenesis of COVID-19 (2023).
- Investigating the Potential Shared Molecular Mechanisms between COVID-19 and Alzheimer's Disease via Transcriptomic Analysis (2023).

Are you planning to include this pathway in your next publication? See How to Cite and add a link here to your paper once it's online.

So, in the next few days I plan to do two things: 1. generate RDF for the YAML file and make that part of the monthly WikiPathways RDF release; 2. extract citations and offer this back to the OpenCitations project; and, 3. add the citations into Wikidata. Of course, all with `cito:usesDataFrom` :)

There is a fourth things that I am still thinking about. I can also use the above data the annotation citations to the WikiPathways papers if they also mention a WikiPathways identifier as `cito:usesDataFrom`, but I cannot fully oversee the implications of that. What do you think?