

EDITORIAL

Open Access



A look back at a pilot of the citation typing ontology

Rajarshi Guha^{1*}, Barbara Zdrzil², Nina Jeliaskova³ and Karina Martinez-Mayorga^{4,5}

Citations are a foundational component in scholarly writing, and while every citation has a reason for inclusion in a manuscript, the intent behind a particular citation needs to be inferred from the context. The Citation Typing Ontology (CiTO) was first described in [1, 2], motivated by the desire to allow authors to formally specify the intent of a given citation.

While it has seen relatively low usage in the scholarly publishing community since its introduction, there have been efforts to encourage its use, largely in a retrospective manner [3–6]. Given the utility of including machine readable annotations that would enhance the semantics of citations in journal articles, the Editors of the *Journal of Cheminformatics* initiated a pilot program to accept manuscripts containing CiTO annotations, as described in [7].

The pilot program ran from 2020 to 2022 and during this period published 15 articles, which annotated, on average, 75% of the bibliography with CiTO terms. A detailed analysis of the CiTO annotations included in the articles submitted to the *Journal of Cheminformatics* is presented in Wilighagen [8]. This editorial summarizes

our experience with the pilot, in particular, some of the challenges we faced in working with CiTO annotations, and discusses our decision to discontinue the pilot.

While the ontology has existed for more than 10 years, it is not necessarily familiar to authors who would consider submitting to this journal. Thus we provided a landing page (<https://www.biomedcentral.com/collections/cito>) and guidelines (<https://jcheminform.github.io/jcheminform-author-guidelines/cito>) to authors wishing to include annotations in their submitted manuscripts. As noted there, we suggested a list of annotations for authors to consider, basing this list on the assumption that use of these annotations would be most informative. However, we did not restrict authors from using other annotations.

While it was encouraging that authors were responsive to the pilot, 15 articles corresponds to approximately 8% of all articles published in the period 2020–2022 and thus received relatively little attention over a two-year period. We recognize that submitting articles to this pilot involved a number of challenges.

First and foremost is that it required authors to familiarize themselves with the CiTO and be able to identify the most relevant terms for their citations. While we believe that it was not an overly onerous burden, it is still an additional task that must be considered when writing and submitting an article, requiring extra effort on the part of authors.

This leads to the second challenge. We required that submissions specify CiTO annotations as extra text appended to each bibliographic entry that authors intended to annotate. With the exception of CiteULike, we are not aware of any bibliography management tool that supports CiTO annotations. While a number of

*Correspondence:

Rajarshi Guha
rajarshi.guha@gmail.com

¹ Vertex Pharmaceuticals, 50 Northern Ave, Boston, MA 02210, USA

² European Molecular Biology Laboratory, European Bioinformatics Institute (EMBL-EBI), Wellcome Genome Campus, Hinxton, Cambridgeshire CB10 1SD, UK

³ Ideacon Ltd., Sofia, Bulgaria

⁴ Institute of Chemistry, Campus Merida, National Autonomous University of Mexico, Merida-Tetiz Highway, Km. 4.5, Ucu, Yucatán, Mexico

⁵ Institute for Applied Mathematics and Systems, Merida Research Unit, National Autonomous University of Mexico, Sierra Papacal, Merida, Yucatan, Mexico



© The Author(s) 2023. **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>. The Creative Commons Public Domain Dedication waiver (<http://creativecommons.org/publicdomain/zero/1.0/>) applies to the data made available in this article, unless otherwise stated in a credit line to the data.

tools allow the user to the notes section of a bibliographic entry to specify the annotation, this is not native support. This meant authors needed to manually modify their final bibliography and insert CiTO annotations by hand. This is both tedious and an inefficient way to annotate citations. Nonetheless, it was interesting to note that on average authors annotated 75% of their bibliography (with 5 articles containing a 100% annotated bibliography).

The third challenge arose after authors submitted their articles. The journal publishing pipeline is not designed to support CiTO annotations and as a result required manual intervention at multiple points; editors needed to flag articles as part of a CiTO pilot, and processing of annotations also required manual intervention. This led to an inefficient article management and publishing pipeline, and led to errors in the final published form for some articles that subsequently needed to be corrected. As a consequence, the final publication of the articles including CiTO annotations was often delayed.

In summary, the inclusion of CiTO annotations was burdensome on authors and editors. More fundamentally, while the concept of formally annotating intent is commendable, it is unclear as to which is the real-world usage of such annotations. And thus the return on investment—for authors, editors and the journal—is unclear at this point in time.

Based on the relatively low response rate and the challenges described above, we have decided to end the pilot and will no longer officially support CiTO annotations. However, for authors who include CiTO annotations in their submissions, we will make an effort to ensure that they are considered in the publication process.

As recognized in [7], the use of CiTO is a chicken and egg problem, and we anticipate that in 2023, it still remains so. Nonetheless, we see a few actions that could help popularize the use of CiTO annotations and bring it into mainstream publishing.

First, there is a need for more education by journals and publishing organizations (academic and commercial) for authors. Second is CiTO support within bibliography management tools, as well as a recognition of CiTO annotations in standard bibliography formats. Finally, there is a need for a long term view on the part of publishers. While it is true that CiTO annotations do not have an immediate use case, the underlying premise is sound, and this is an effort where it seems that a “critical mass” of annotations must be achieved before any significant community-wide benefits can be seen. Essentially, this means that support for CiTO annotations requires an upfront investment in education and upgrading of publishing pipelines, without expectation of short term returns.

Hopefully publishers (and associated software developers) will recognize the scientific and social benefit that can accrue from a semantic representation of citations, and will be able to drive the uptake of CiTO in the future.

Data availability

The CiTO annotations for the articles in this pilot can be accessed from the individual articles.

Published online: 03 February 2023

References

1. Shotton D (2010) CiTO, the citation typing ontology. *J Biomed Semantics* 1(Suppl 1):S6. <https://doi.org/10.1186/2041-1480-1-S1-S6>
2. Peroni S, Shotton D (2012) FaBio and CiTO: ontologies for describing bibliographic resources and citations. *Web Semant* 17:33–43. <https://doi.org/10.1016/j.websem.2012.08.001>
3. Zhu X, Turney P, Lemire D, Vellino A (2015) Measuring academic influence: not all citations are equal. *J Assoc Inf Sci Technol* 66:408–427. <https://doi.org/10.1002/asi.23179>
4. Matejka J, Grossman T, Fitzmaurice G. (2012) Citeology. In: CHI '12 Extended Abstracts on Human Factors in Computing Systems. ACM, New York
5. Mayernik MS, Hart DL, Maull KE, Weber NM (2017) Assessing and tracing the outcomes and impact of research infrastructures. *J Assoc Inf Sci Technol* 68:1341–1359. <https://doi.org/10.1002/asi.23721>
6. Yu B (2013) Automated citation sentiment analysis: What can we learn from biomedical researchers. *Proc Am Soc Inf Sci Technol* 50:1–9. <https://doi.org/10.1002/meet.14505001084>
7. Willighagen E (2020) Adoption of the citation typing ontology by the journal of cheminformatics. *J Cheminform* 12:47. <https://doi.org/10.1186/s13321-020-00448-1>
8. Willighagen E (2023) Two years of explicit CiTO annotations. *J Cheminform*. <https://doi.org/10.1186/s13321-023-00683-2>

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Ready to submit your research? Choose BMC and benefit from:

- fast, convenient online submission
- thorough peer review by experienced researchers in your field
- rapid publication on acceptance
- support for research data, including large and complex data types
- gold Open Access which fosters wider collaboration and increased citations
- maximum visibility for your research: over 100M website views per year

At BMC, research is always in progress.

Learn more biomedcentral.com/submissions

