



Software Citation Update

Open Source Software Health Index Workshop, 18 June 2019

Daniel S. Katz, Martin Fenner, Neil Chue Hong
FORCE11 Software Citation Implementation Working Group Co-Chairs

(working on [GitHub](#), linked from
[our FORCE11 WG page](#))



Happy Days

July 2015: starting FORCE11 Software Citation working group, with Data Citation group as an model to follow

September 2015: built community of researchers, developers, publishers, repositories, librarians,
w/ co-leads: Arfon Smith, Kyle Niemeyer

Goal: enable citation of software itself for credit, support reproducibility

... lot of great discussions, arrived at a consensus ...



More Happy Days

September 2016: published Smith AM, Katz DS, Niemeyer KE, FORCE11 Software Citation Working Group. (2016) Software Citation Principles. PeerJ Computer Science 2:e86. DOI: [10.7717/peerj-cs.86](https://doi.org/10.7717/peerj-cs.86) and <https://www.force11.org/software-citation-principles>

Early 2017: Declared success



Software Citation Principles

Started with data citation principles, updated based on software use cases and related work, updated based working group discussions, community feedback and review of draft, workshop at FORCE2016

1. Importance

2. Credit and Attribution

3. Unique Identification

4. Persistence

5. Accessibility

6. Specificity

Paper also included lots of discussion to help use principles



Less Happy Days

Early 2017: Realized that principles were not enough

Started Software Citation Implementation Working Group to

- Write out the “small amount” of detail needed to implement the principles
- Work with communities to actually implement them
 - Publishers, conferences, repositories, indexers, funders, etc.



Moderately Happy Days

- Formed a good group, with diverse interests and expertise
- Lots of good work done, and good coordination of ongoing activities
 - Metadata standards and translation (DataCite Schema 4.1, CodeMeta, citation.cff)
 - Open source archiving and identification (Software Heritage)
 - Good work and initial acceptance in communities (astronomy, Earth science, math, HEP, ...)



Unsettled Days

Mid 2018 - today: Realized “small amount” of detail wasn’t small, scattered progress wasn't sufficient, underlying challenges not being addressed

Wrote document to identify challenges:

D. S. Katz et al. (2019) Software Citation Implementation Challenges. [arXiv:1905.08674](https://arxiv.org/abs/1905.08674)



Technical Challenges

- Complexity of software types: open source, closed source; published, unpublished; versioned, unversioned; developed by citer, not developed by citer; services, containers, executables
- How to uniquely identify software of each type (ideally as uniformly as possible)
 - Including via new [Joint FORCE11 RDA Software Source Code Identification WG](#)
- How to define and store citation metadata for each type
- How to access metadata and convert it as needed
- How to count citations across versions
- Realization: metadata is fundamental



Social Challenges

Need groups that work on implementation in context

- Disciplinary communities
- Publishers
- Repositories
- Indexers
- Funders
- Institutions

Groups need to come together, run pilots to establish norms



Mixed progress in 2019

Positives in 2019:

- [Software Citation Checklist for Authors](#) document drafted and under review, led by Neil Chue Hong
- [Software Citation Checklist for Reviewers](#) document, started by Neil but on hold until the author document is completed
- Repositories task force led by Alice Allen started, good community participation
- [CodeMeta](#) gaining more community acceptance as a metadata standard, “aligned” with schema.org

Negatives in 2019:

- Work on these documents brings up technical issues/challenges that are not resolved; not clear if they will be useful or will fail to be completed, or perhaps will just need later iterations to improve
- Publisher task force not yet started



Next steps

Continue to coordinate ongoing activities

Try to address the challenges & underlying problems, realizing that this is a long term activity

Find funding for meeting(s) and work to plan and implement