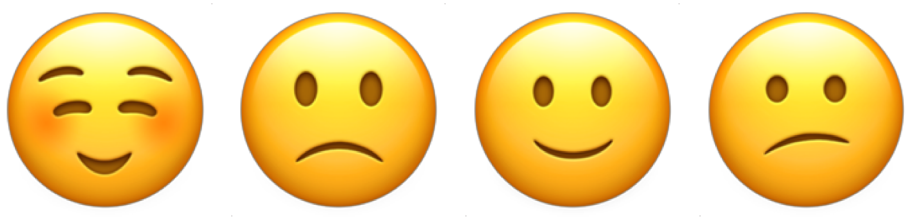


FORCE11 Software Citation Implementation WG Update

ESIP Winter Meeting, 15 Jan 2019

Daniel S. Katz, Martin Fenner, Neil Chue Hong
Working Group Co-Chairs

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

FORCE11
The Future of Research Communications and e-Scholarship

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

… 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 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, …)

Earth Science Details

- ESIP Software and Services Citation Cluster, led by Jessica Hausman
  - Specific recommendations and examples for software citation, aligned with FORCE11 work
  - Effective balance of principles and practices
- Enabling FAIR Data project, led by Shelley Stall
  - Mostly focused on data, but considering also other research objects such as software
- Learned at AGU Fall meeting (reinforcing other external inputs)
  - Researchers (& other stakeholders) looking for guidance & evidence that doing this is a good idea for them and for their field

Unsettled Days

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

Started document to identify challenges

Discussed at FORCE2018 Software Citation workshop - led to better understanding

Have plan to move to completion by Jan 2019

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) [also see 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

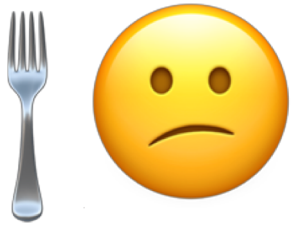
Social Challenges

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

At a Fork in the Road



Today: How should the FORCE11 group go forward?

1. Continue to coordinate activities, declare success in a year or so
2. Try to address the challenges & underlying problems, realizing that this is a long term activity

In both cases, need some funding for work and meeting(s)