Software Citation
Introduction

Daniel S. Katz (d.katz@ieee.org, @danielskatz)
Chief Scientist, NCSA
Associate Research Professor, CS, ECE, iSchool
The Software Citation journey (via FORCE11)

2015
Software Citation WG started

2016
Software Citation Principles published

2017/18
Software Citation Implementation WG started

2019/20
Task forces publish resources

2021/22
Adoption of software citation increases?

Goals:
• Credit for software developers and maintainers
• Better & more sustainable software
• Support for reproducibility

https://www.force11.org/group/software-citation-implementation-working-group
FORCE11 Software Citation Working Group (2015-16)

- Documented differences between software and data; defined software citation challenges

- Created software citation principles

https://www.force11.org/group/software-citation-working-group
Co-Chairs: Arfon M. Smith, Daniel S. Katz, Kyle E. Niemeyer
1. Importance
2. Credit and Attribution
3. Unique Identification
4. Persistence
5. Accessibility
6. Specificity
FORCE11 Software Citation Implementation Working Group (2017-present)

• Initial goals:
  – Write out the “small amount” of detail needed to implement the principles
  – Coordinate research & other work going on in many areas
  – Work with communities to actually implement the principles
• Quickly realized “small amount” of detail wasn’t small, scattered progress wasn't sufficient, underlying challenges not being addressed
  – Technical challenges include complexity of software types and identifiers, where to store metadata, …
  – Social challenges need groups that work on implementation in context (disciplinary communities, publishers, repositories & registries, indexers, funders, institutions) to come together and run pilots to establish norms

https://www.force11.org/group/software-citation-implementation-working-group
Co-Chairs: Neil Chue Hong, Martin Fenner, Daniel S. Katz
**Responses to challenges (1)**

- **Guidance task force**
  - For paper authors who want to cite software
  - For software developers who want to make their software citable

- **CodeMeta task force**
  - Following CodeMeta project
    - In parallel with Software Citation Principles & Implementation Working Groups
    - Some common membership
    - Aiming to understand metadata for software, not just for use in citation
    - Built a crosswalk of existing metadata standards for software
    - Then developed a CodeMeta standard to describe software based on these crosswalks
    - Updating the CodeMeta standard
    - Describing everything in CodeMeta using schema.org properties
    - Moving CodeMeta into a community group, with governance

- **CFF standard (citation metadata), now integrated into GitHub, Zenodo, Zotero**
Responses to challenges (2)

• Software Registries Task Force
  – Developed best practices document
    – Community continuing in SciCodes: Consortium of scientific software registries and repositories, https://scicodes.net/

• Journals Task Force
  – Working with publishers to provide generic guidelines for journals and conferences to provide to authors
    • They then provide specific guidelines, with community-accepted language and examples
    • Tracked by CHORUS in Software Citation Policy Index
  – Also working on publication processing
    • How citation information moves from author provides to internal publisher/contractor systems and then to indices
    • S. Stall, et al., “Journal Production Guidance for Data and Software Citations”, drafted, will be submitted shortly
Responses to challenges (3)

• Considered an Institutions task force, but didn’t get sufficient interest
  – Institutions: places where people work
    • Universities, laboratories, industry, government, etc.
  – Want to affect policies and practices
    • How do they encourage software citation
    • How do they use software citation information in hiring & promotion
  – Collect and share examples
  – Help form communities

• Overall planning
  – Open question: Given progress to date, what else makes sense to do, and who can do it?
  – IMLS-funded software citation workshop this summer addressed this, report coming soon
The Software Citation journey (via FORCE11)

2015
Software Citation WG started
~55 members (researchers, developers, publishers, repositories, librarians)
Reviewed existing community practices & developed use cases

2016
Software Citation Principles published
Started with data citation principles, updated based on software use cases and related work, working group discussions, community feedback
Software citation principles published after community review: 10.7717/peerj-cs.86

2017/18
Software Citation Implementation WG started
Group set up to:
1. endorse the principles
2. develop sets of guidelines for implementing the principles
3. help implement the principles
4. test specific implementations of the principles

2019/20
Task forces publish resources
CodeMeta Task Force provides recommendations for schema changes
Repositories Task Force runs workshop to identify best practices
Journals Task Force starts adoption process with journals and publishers to promote Recognizing the value of software: a software citation guide: 10.12688/f1000research.26932.2

2021/22
Adoption of software citation increases?
How does software citation fit with other related work:
FAIR for Research Software
Open Research / Open Science
Reproducibility
Software catalogs
Work on backend processing of papers & transfer of citation information to indexers

https://www.force11.org/group/software-citation-implementation-working-group
We are here

• We now have
  – Software citation principles
  – Policies from publishers with examples
  – Tools to support processing/indexing
  – Some uptake among software developers, paper authors, editors & reviewers

• What we need
  – More uptake (including from HEP community)
  – Policies from stakeholders (including HEP experiments)
  – Discussion of any blockers/friction that we can work on