Software Citation: State of the Practice and Challenges

Stephan Druskat

German Aerospace Center (DLR)

Intelligent and Distributed Systems

Berlin, Germany

Software Engineering and Reuse in Computational
Science and Engineering (ISC 2019 BOF)
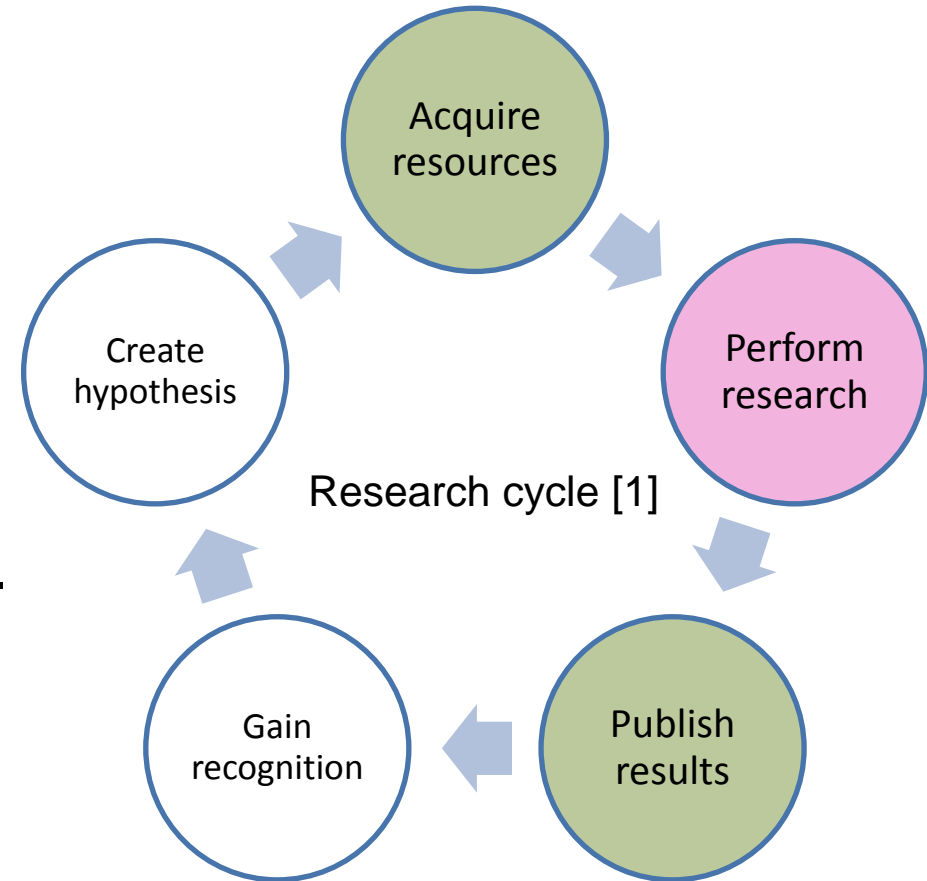
Knowledge for Tomorrow

DLR

Software Citation: General Challenges

Software is essential to research!

Software **contributes** to the research cycle …
1) … as *infrastructure*,
2) … as *research product* in its own right.

In research, **contributions** are measured through citation.

Research cycle [1]

Acquire resources

Perform research

Publish results

Gain recognition

Create hypothesis

Cultural challenge: Software not recognized as research product
Systemic challenge: Current citation system is for papers and books
Overall challenge: Identify software used in research

Software Citation Principles [2]

Principles	Meaning: „Software citation …"
Importance	→ cites software like papers. [*culture*, *system*]
Credit and attribution	→ facilitates credit and legal attribution. [*culture*, *system*]
Unique identification	→ uses machine-actionable UIDs. [*system*]
Persistence	→ lets UID and metadata persist. [*system*]
Accessibility	→ facilitates access to software, metadata, etc. [*system*]
Specificity	→ facilitates identification of/access to versions. [*system*]

Software Citation Principles: Are We There Yet?

Principles	**Example Solutions**
Importance	→ Policies, initiatives, guidance, spreading the word
Credit and attribution	→ Software journals, reference lists, policies, creditability
Unique identification	→ DOIs for software
Persistence	→ Archival repositories, Software Heritage
Accessibility	→ Archival repositories
Specificity	→ Versioned DOIs

Also [3]: Different software types! Metadata! Tooling! Stakeholder challenges! And more …

Are we there yet? Not quite! We're making progress, but there are still (tough) challenges, including culture change (which is hard).

Software Citation - What Works Today? The Citation File Format (CFF) [4]

How to provide complete and correct metadata?

Provision of software citation metadata, in a machine- and human-readable format, following the Software Citation Principles.

```
cff-version: 1.0.3
message: If you use this software, please cite it as below.
authors:
  - family-names: Druskat
    given-names: Stephan
    orcid: https://orcid.org/0000-0003-4925-7248
title: My Research Tool
version: 1.0.4
doi: 10.5281/zenodo.1234
date-released: 2017-12-18
```

citation-file-format.github.io

References

[1] D. S. Katz and N. P. Chue Hong, "Software Citation in Theory and Practice," in *Mathematical Software – ICMS 2018*, 2018, pp. 289–296.

[2] A. M. Smith, D. S. Katz, K. E. Niemeyer, and FORCE11 Software Citation Working Group, "Software citation principles," *PeerJ Comput. Sci.*, vol. 2, no. e86, 2016 [Online]. Available: https://doi.org/10.7717/peerj-cs.86

[3] D. S. Katz *et al.,* "Software Citation Implementation Challenges," *arXiv:1905.08674 [cs]*, May 2019 [Online]. Available: http://arxiv.org/abs/1905.08674.

[4] S. Druskat, N. Chue Hong, R. Haines, and J. Baker, "Citation File Format (CFF)," 2017 [Online]. Available: https://doi.org/10.5281/zenodo.1003149

Figures

- Research cycle diagram adapted from: Daniel S. Katz, "Software Citation in Theory and Practice," [Online]. Available: https://www.slideshare.net/danielskatz/software-citation-in-theory-and-practice.

Thank you!

@stdruskat

stephan.druskat@dlr.de

https://orcid.org/0000-0003-4925-7248