

Byte-sized RSE: Season 2, Session 5

Citation File Format (CFF) / cffinit

Making your code citable

Stephan Druskat, German Aerospace Center (DLR)
Abel Siquera, Reggie Cushing, Carlos Martinez Ortiz,
Netherlands eScience Center

In collaboration with the UNIVERSE-HPC project



Session Overview

- [Session intro]
- [Practical activity]
- [Q&A]

Session Introduction: Code of Conduct

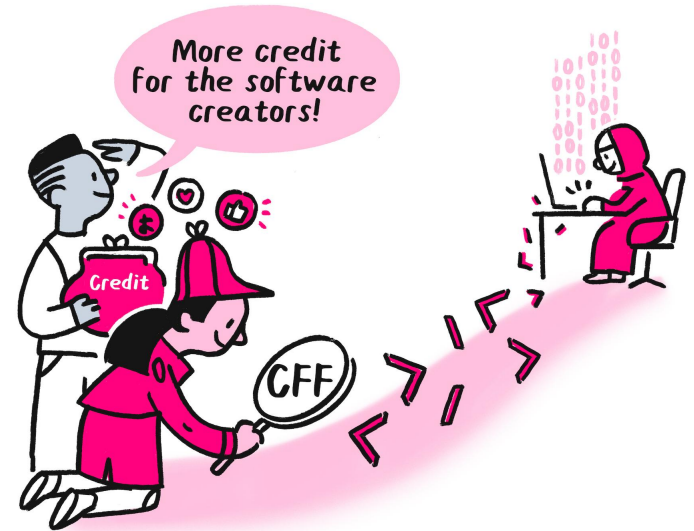
Our byte-sized RSE sessions are run under the UNIVERSE-HPC Events Code of Conduct:

<https://www.universe-hpc.ac.uk/events/code-of-conduct/>

- Please familiarise yourself with the Code of Conduct which applies to all activities in relation to byte-sized RSE sessions.

Why make software citable, and cite software?

- Principles of research apply:
disclosure of research process
- Accountability: taking credit
- **Formal citation** as established method:
 - Importance of software in research
 - Findability (FAIR)
 - Reproducibility
 - Credit



Scriberia

Software citation principles

1. Importance

Software is cited like papers are cited.

2. Credit and attribution

3. Unique identification

4. Persistence

5. Accessibility

Citation allows access to software and metadata.

6. Specificity

Citation identifies the software version used in research.

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, doi: [10.7717/peerj-cs.86](https://doi.org/10.7717/peerj-cs.86).



Software citation principles

Arfon M. Smith^{1*}, Daniel S. Katz^{2,3*}, Kyle E. Niemeyer^{4*} and
FORCE11 Software Citation Working Group

¹GitHub, Inc., San Francisco, California, United States

²National Center for Supercomputing Applications & Electrical and Computer Engineering
Department & School of Information Sciences, University of Illinois at Urbana-Champaign,
Urbana, Illinois, United States

³School of Mechanical, Industrial, and Manufacturing Engineering, Oregon State University,
Covallis, Oregon, United States

* These authors contributed equally to this work.

ABSTRACT

Software is a critical part of modern research and yet there is little support across the scholarly ecosystem for its acknowledgement and citation. Inspired by the activities of the FORCE11 working group focused on data citation, this document summarizes the recommendations of the FORCE11 Software Citation Working Group and its activities between June 2015 and April 2016. Based on a review of existing community practices, the goal of the working group was to produce a consolidated set of citation principles that may encourage broad adoption of a consistent policy for software citation across disciplines and venues. Our work is presented here as a set of software citation principles, a discussion of the motivations for developing the principles, reviews of existing community practice, and a discussion of the requirements these principles would place upon different stakeholders. Working examples and possible technical solutions for how these principles can be implemented will be discussed in a separate paper.

Subjects Digital Libraries, Software Engineering

Keywords Software citation, Software credit, Attribution

SOFTWARE CITATION PRINCIPLES

The main contribution of this document are the software citation principles, written fairly concisely in this section and discussed further later in the document (see Discussion). In addition, we also motivate the creation of these principles (see Motivation), describe the process by which they were created (see Process of Creating Principles), summarize use cases related to software citation (see Use Cases), and review related work (see Related Work). We also lay out the work needed to lead to these software citation principles being applied (see Future Work).

1. **Importance:** Software should be considered a legitimate and citable product of research. Software citations should be accorded the same importance in the scholarly record as citations of other research products, such as publications and data; they should be included in the metadata of the citing work, for example in the reference list of a journal article, and should not be omitted or separated. Software should be cited on the same basis as any other research product such as a paper or a book, that is, authors should cite the appropriate set of software products just as they cite the appropriate set of papers.

Submitted 24 June 2016
Accepted 23 August 2016
Published 19 September 2016
Corresponding author
Daniel S. Katz, d.katz@ieee.org
Academic editor
Silvia Peroni
DOI [10.7717/peerj-cs.86](https://doi.org/10.7717/peerj-cs.86)
© Copyright
2016 Smith et al.
Distributed under
Creative Commons CC BY 4.0
OPEN ACCESS

[How to cite this article: Smith et al. \(2016\), Software citation principles, PeerJ Comput. Sci., 2\(86\) DOI 10.7717/peerj-cs.86](https://doi.org/10.7717/peerj-cs.86)

Software citation needs metadata!

- What is the software name?
- Who are the software authors?
- What version have I used?
- Who do I give as the publisher?
- What is the publication date?

• ... 

```
stephan@computer:~/src/cff-corporus$ ./script.py | sed -e "s/.\{42\}/&\n/g"
##### Statistics #####

CFF files (total): ..... 56,848

GitHub users/orgs (total): ... 11,520
Repositories (total): ..... 16,175

Commits (total): ..... 56,848
Unique commits (total): ..... 46,450
N-plicate commits (total): ... 10,398

Unique commit dates (total): . 1,589

Nested CFF file paths (total): 1,676

Earliest commit date: ..... 2017-11-02
Earliest committed file: .... /home/stephan/src/cff-corporus/data/cffs/force11/force11-sciwg/2017-11-02/10-07-24/5b94cacca5ec1c61cc962e04e14213bc449bc7a/examples/swh-in-dexer_CITATION.cff
Latest commit date: ..... 2024-03-17

#####

stephan@computer:~/src/cff-corporus$
```

What is the software citation metadata?

Only you/your project knows the correct metadata:

- Software name, versions, where can the artifacts be found?
- Who are the authors?
 - Is the person who fixed that typo / designed the UI / wrote unit tests a "software author"? www.researchsoft.org/tf-authorship-contribution/

<ReSA>

Research Software Alliance

About ▾ Task Forces ▾ Forums ▾ Resources ▾ News Events Blog

Research Software Authorship and Contribution Task Force

Recording software citation metadata

```
cff-version: 1.2.0
message: "If you use this software, please cite it as below."
title: "My Research Software"
abstract: "MRS does many research things."
version: 2.0.4
authors:
  - family-names: Druskat
    given-names: Stephan
    orcid: https://orcid.org/1234-5678-9101-1121
identifiers:
  - type: doi
    value: 10.5281/zenodo.1234
    description: "This is the archived snapshot of MRS version 2.0.4"
license: Apache-2.0
repository-code: https://gitlab.institution.edu/druskat/my-res-soft
date-released: 2021-08-11
```

citation-file-format.github.io/

The Citation File Format

- `CITATION.cff` in the source code repository root
- YAML
- `references`
- `preferred-citation`
- `type: dataset`



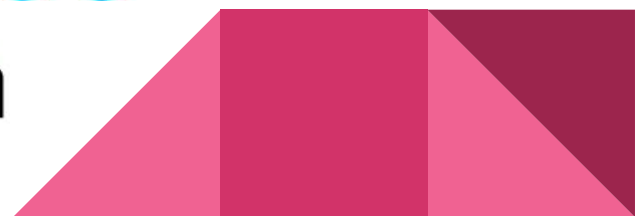
Tool ecosystem & usage

- Creation
- Editing & updating
- Validation
- Conversion



~15,000 CITATION.cff files on GitHub

Integrations: Zenodo, GitHub, Zotero, JabRef, IDEs



Demo: Create a CITATION.cff file with cffinit

Over to Abel



Demo: Harnessing the GitHub & Zenodo integrations



1: Repository w/out a CITATION.cff

The screenshot shows a GitHub repository interface. At the top, there are navigation elements: a dropdown menu for 'main', '1 Branch', and '0 Tags'; a search bar 'Go to file'; an 'Add file' button; and a 'Code' button. Below this is a commit summary for 'citable-s0ftware' with the message 'Add note aobut non-existence of CITATION.cff file', commit hash 'cb5991c', and '3 Commits'. A file list shows 'README.md' with the same commit message and 'now' as the time. The main content area displays the README for 'cff-github-zenodo', which states: 'Demonstrating the Citation File Format integrations with GitHub and Zenodo. This repository does not contain a CITATION.cff file.' On the right sidebar, there are sections for 'About' (demonstration of integrations), 'Releases' (no releases published), and 'Packages' (no packages published).

Create a GitHub repository WITHOUT a CITATION.cff.

2: Activate Zenodo's GitHub integration (Sandbox!)

The screenshot shows the Zenodo Settings page for GitHub integration. On the left is a sidebar with 'Settings' and a list of options: Profile, Change password, Security, Notifications, Linked accounts, Applications, and GitHub (highlighted). The main content area is titled 'GitHub Repositories' and includes a 'Sync now' button. It features a 'Get started' section with three steps: 1. Flip the switch (with an 'ON' toggle and '(example)') and instructions to select a repository and toggle automatic preservation; 2. Create a release (with instructions to go to GitHub and create a release, and that Zenodo will download a .zip-ball and register a DOI); 3. Get the badge (with instructions that a DOI badge will appear next to the repository, and an example DOI: 10.5281/zenodo.8475). Below this is an 'Enabled Repositories' section with one entry: 'citable-s0ftware/cff-github-zenodo' with an 'ON' toggle. At the bottom, a 'Repositories' section contains a note: 'If your organization's repositories do not show up in the list, please ensure you have enabled [third-party access](#) to the Zenodo application. Private repositories are not supported.'

Activate the GitHub integration for the repository in the Settings on Zenodo Sandbox.

3: Make a release without having a CITATION.cff

Releases Tags

without-CITATION.cff Target: main Previous tag: auto Generate release notes

Excellent! This tag will be created from the target when you publish this release.

Release without a CITATION.cff

Write Preview

This is a release of this repository which is automatically published on Zenodo Sandbox.

Note that for this release, this repository does NOT contain a CITATION.cff file.

↓ Attach binaries by dropping them here or selecting them.

Set as a pre-release
This release will be labeled as non-production ready

Publish release Save draft

Make a release in the repository.

Create a tag
without-CITATION.cff
on publication.



4: View the Zenodo record

citable-s0ftware/cff-github-zenodo: Release without a CITATION.cff

citable-s0ftware

This is a release of this repository which is automatically published on Zenodo Sandbox.

Note that for this release, this repository does NOT contain a CITATION.cff file.

Files

citable-s0ftware/cff-github-zenodo-without-CITATION.cff.zip


citable-s0ftware/cff-github-zenodo-without-CITATION.cff.zip

citable-s0ftware-cff-github-zenodo-cb5991c

README.md

Most metadata for the record is retrieved from the GitHub API.

Rights

 Creative Commons Attribution 4.0 International

Citation

citable-s0ftware (2024). citable-s0ftware/cff-github-zenodo: Release without a CITATION.cff (without-CITATION.cff). Zenodo. <https://doi.org/10.5072/zenodo.37554>

Style

APA

Share

0

VIEWS

0

DOWNLOADS

Show more details

Versions

Version without-CITATION.cff

Mar 18, 2024

10.5072/zenodo.37554

Cite all versions? You can cite all versions by using the DOI 10.5072/zenodo.37553. This DOI represents all versions, and will always resolve to the latest one. Read more.

External resources

Available in



citable-s0ftware/cff-github-zenodo
Release: without-CITATION.cff

Indexed in

5: Add a CITATION.cff file

The screenshot shows a GitHub repository for 'citable-s0ftware'. The commit history shows a recent commit 'Add software citation metadata' for the file 'CITATION.cff'. The README file contains the text 'This repository does contain a CITATION.cff file.' A popup window titled 'Cite this repository' is displayed, showing citation information for 'Druskat, S. (2024), pseudoSoftware (Version 0)' in both APA and BibTeX formats. The BibTeX string is 'Druskat, S. (2024), pseudoSoftware (Version 0)'. The popup also includes a 'View citation file' button.

main 1 Branch 1 Tags

Go to file Add file Code

About

Demonstration of the Citation File Format integrations with GitHub and Zenodo

Readme

Cite this repository

Cite this repository

If you use this software in your work, please cite it using the following metadata. [Learn more about CITATION files.](#)

APA BibTeX

Druskat, S. (2024), pseudoSoftware (Version 0)

CITATION.cff Latest

View citation file

No packages published [Publish your first package](#)

Contributors 2

citable-s0ftware

GitHub reads the metadata from CITATION.cff and renders a formatted citation and a BibTeX string.

6: Make a new release with a CITATION.cff

Releases Tags

with-CITATION.cff Target: main Previous tag: auto Generate release notes

Excellent! This tag will be created from the target when you publish this release.

Release WITH a CITATION.cff

Write Preview

This is a release of this repository which is automatically published on Zenodo Sandbox.

Note that for this release, this repository DOES contain a CITATION.cff file.

↓ Attach binaries by dropping them here or selecting them.

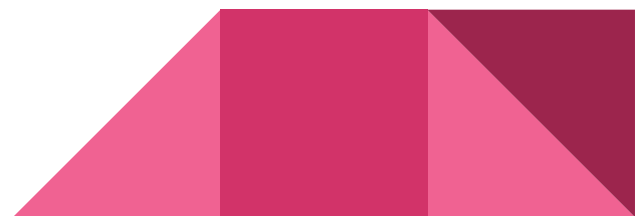
Set as a pre-release
This release will be labeled as non-production ready

Set as the latest release
This release will be labeled as the latest for this repository.

Publish release Save draft

Make a release in the repository.

Create a tag with-CITATION.cff on publication.



7: View the Zenodo record

pseudoSoftware

Druskat, Stephan¹

Druskat, Stephan's ORCID profile

Hide affiliations

1. German Aerospace Center (DLR)

This "software" is a demonstration-only repository with two releases. For both releases, the Zenodo GitHub integration (<https://help.zenodo.org/faq/>) automatically publishes artifacts and metadata with a DOI (on the Zenodo Sandbox). 1. The first release contains no CITATION.cff file. Hence, the Zenodo Sandbox record is populated with metadata retrieved from the GitHub API. 2. The second release contains a CITATION.cff file. Hence, the Zenodo Sandbox record is populated with metadata retrieved from the CITATION.cff file.

Notes

If you use this software, please cite it using the metadata from this file.

Files

citable-s0ftware/cff-github-zenodo-with-CITATION.cff.zip

citable-s0ftware/cff-github-zenodo-with-CITATION.cff.zip

▾ citable-s0ftware-cff-github-zenodo-a497f1d

📄 CITATION.cff

📄 README.md


Most metadata for the record is retrieved from the CITATION.cff file.

Keywords and subjects

Citation File Format Demo Zenodo integration

GitHub integration

Rights

 Creative Commons Zero v1.0 Universal

Citation

Druskat, S. (2024). pseudoSoftware (with-CITATION.cff). Zenodo. <https://doi.org/10.5072/zenodo.37555>

Style

Share

0
EYES
VIEWS

0
DOWNLOADS

Show more details

Versions

Version with-CITATION.cff Mar 18, 2024
10.5072/zenodo.37555

Version without-CITATION.cff Mar 18, 2024
10.5072/zenodo.37554

View all 2 versions

Cite all versions? You can cite all versions by using the DOI 10.5072/zenodo.37553. This DOI represents all versions, and will always resolve to the latest one. Read more.

External resources

Available in

 citable-s0ftware/cff-github-zenodo
Release: with-CITATION.cff

8: Thanks!

The **Citation File Format** has been funded by the *German Aerospace Center (DLR)*, the *Netherlands eScience Center*, the *Software Sustainability Institute*, *Code for Science & Society*, and *CampusSOURCE e.V.* 🙏

 <https://citation-file-format.github.io>

Contact

stephan.druskat@dlr.de | Fediverse: [@sdruskat@fosstodon.org](https://fosstodon.org/@sdruskat)

ORCID: <https://orcid.org/0000-0003-4925-7248>

