

# THE CITATION FILE FORMAT: PROVIDING CITATION METADATA FOR RESEARCH SOFTWARE

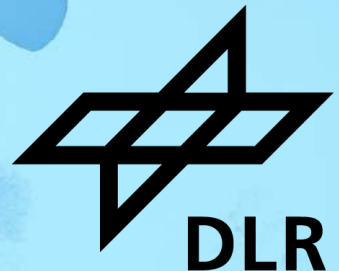
**Stephan Druskat**

German Aerospace Center (DLR), Institute for Software Technology  
Humboldt-Universität zu Berlin, Department of Computer Science

DOI [10.5281/zenodo.7655140](https://doi.org/10.5281/zenodo.7655140)

License [CC-BY-4.0 International](https://creativecommons.org/licenses/by/4.0/)

netherlands  
**eScience center**



# Software citation needs metadata



- Good research practice requires software to be cited [1]
- Software does not usually come with metadata included:  
software projects must define and provide correct citation metadata
- The Citation File Format [2] lets projects provide software citation metadata

[citation-file-format.github.io](https://citation-file-format.github.io) | [github.com/citation-file-format/citation-file-format](https://github.com/citation-file-format/citation-file-format)

**Open licenses:** Apache-2.0, MIT, CC-BY-4.0, CC0-1.0

# Citation File Format

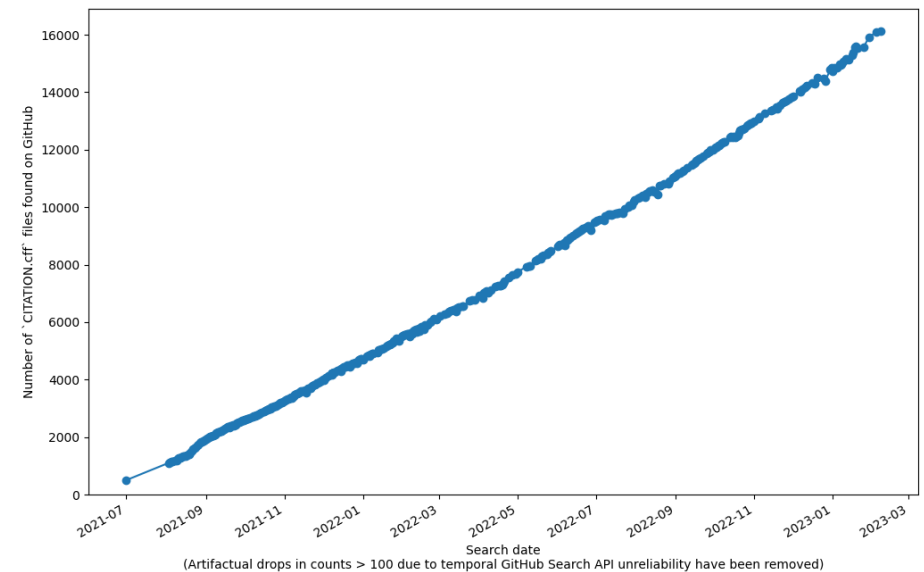


- Human-/machine-readable software citation metadata in YAML
- `CITATION.cff` file in the source code repository

```
cff-version: 1.2.0
message: If you use this software, please cite it using these metadata.
title: My Research Software
abstract: This is my awesome research software. It does many things.
authors:
  - family-names: Druskat
    given-names: Stephan
    orcid: "https://orcid.org/0000-0003-4925-7248"
version: 0.11.2
date-released: "2021-07-18"
identifiers:
  - description: This is the collection of archived snapshots of all versions of My Research Software
    type: doi
    value: "10.5281/zenodo.123456"
  - description: This is the archived snapshot of version 0.11.2 of My Research Software
    type: doi
    value: "10.5281/zenodo.123457"
license: Apache-2.0
repository-code: "https://github.com/citation-file-format/my-research-software"
```

# Citation File Format Status

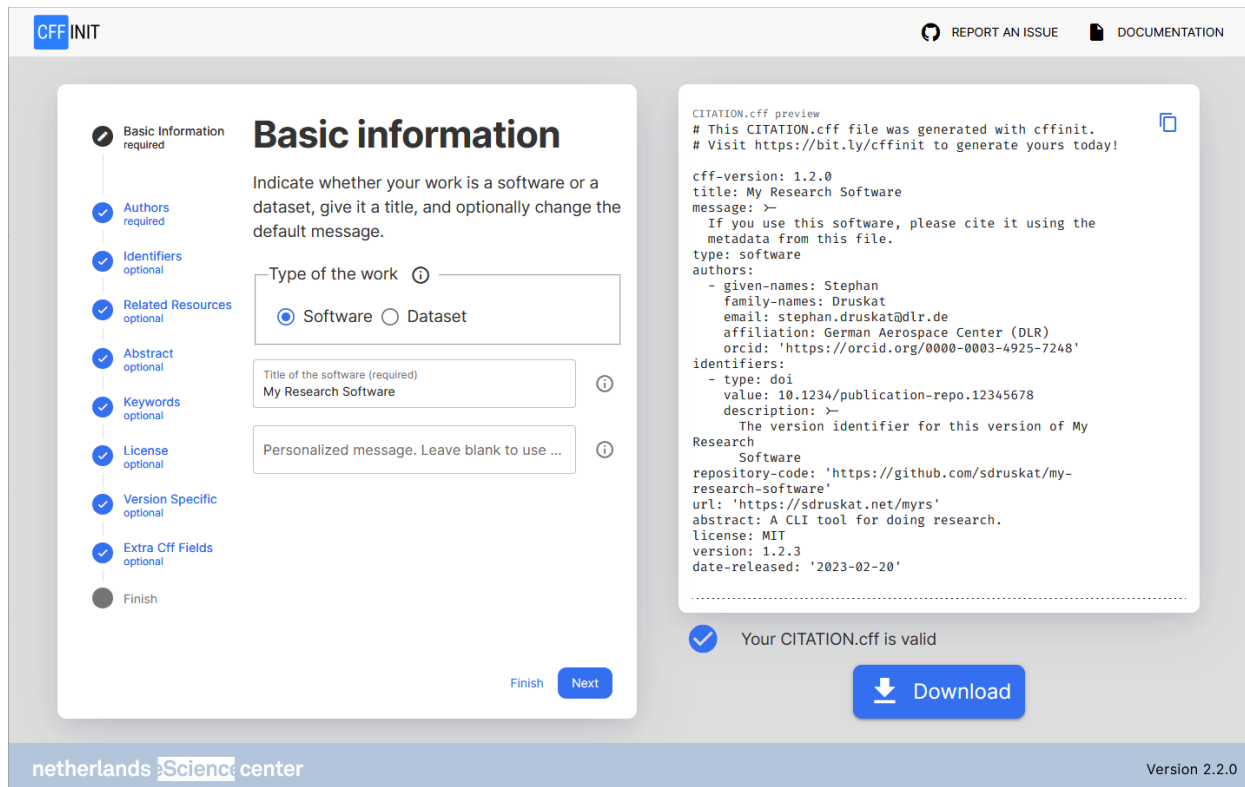
- Covers ~~everyone's~~ a **single use case**: provide citation metadata for software
- A ~~universal~~ **community standard**: 16,000+ CITATION.cff files on GitHub
- A community project (co-lead: Jurriaan H. Spaaks; governance tbd)



# Citation File Format Tooling

- `cffinit`: create, edit, validate online
- [bit.ly/cffinit](https://bit.ly/cffinit)

- More tools for creating, editing, updating, validating, converting CITATION.cff files



The screenshot shows the CFFINIT web interface. On the left, a sidebar lists steps: Basic Information (required), Authors (required), Identifiers (optional), Related Resources (optional), Abstract (optional), Keywords (optional), License (optional), Version Specific (optional), Extra Cff Fields (optional), and Finish. The main area is titled 'Basic information' and contains a form with the following fields: 'Type of the work' (radio buttons for Software and Dataset), 'Title of the software (required)' (text input with 'My Research Software'), and 'Personalized message. Leave blank to use ...' (text input). A 'Next' button is at the bottom right of the form. On the right, a 'CITATION.cff preview' box shows the generated text: 

```
CITATION.cff preview
# This CITATION.cff file was generated with cffinit.
# Visit https://bit.ly/cffinit to generate yours today!

cff-version: 1.2.0
title: My Research Software
message: >
  If you use this software, please cite it using the
  metadata from this file.
type: software
authors:
  - given-names: Stephan
    family-names: Druskat
    email: stephan.druskat@dlr.de
    affiliation: German Aerospace Center (DLR)
    orcid: 'https://orcid.org/0000-0003-4925-7248'
identifiers:
  - type: doi
    value: 10.1234/publication-repo.12345678
    description: >
      The version identifier for this version of My
      Research Software
repository-code: 'https://github.com/sdruskat/my-research-software'
url: 'https://sdruskat.net/myrs'
abstract: A CLI tool for doing research.
license: MIT
version: 1.2.3
date-released: '2023-02-20'
```

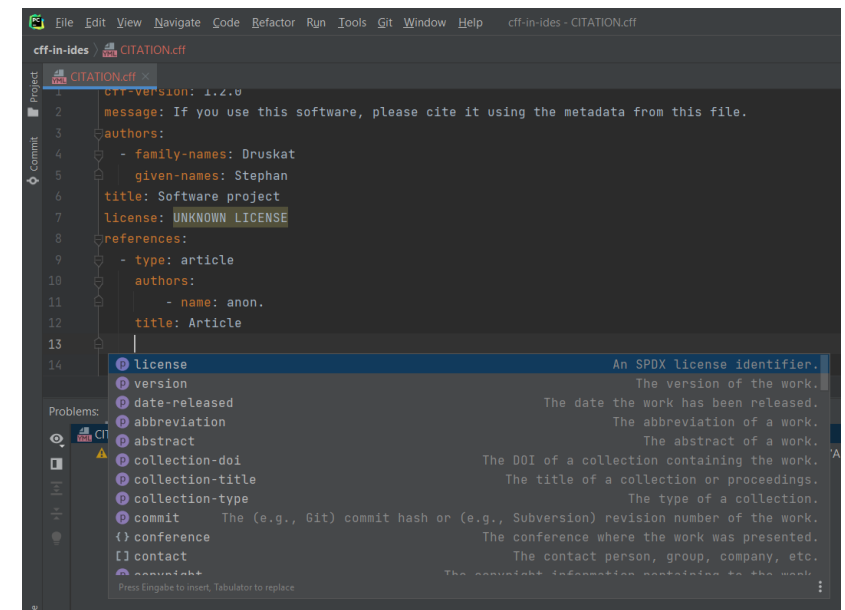
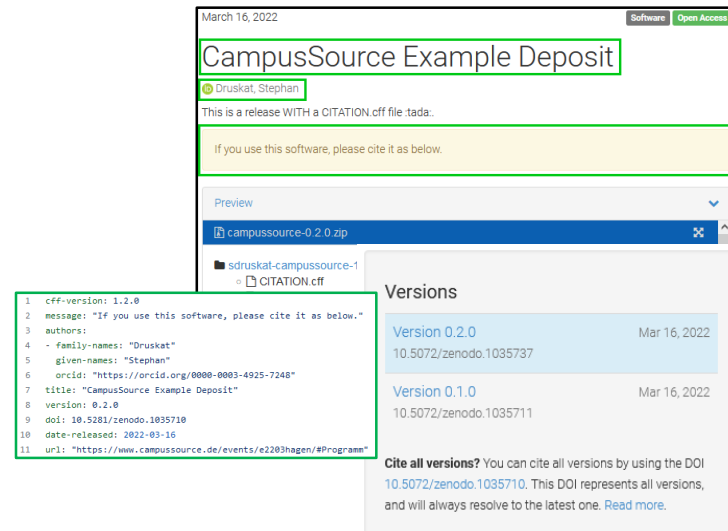
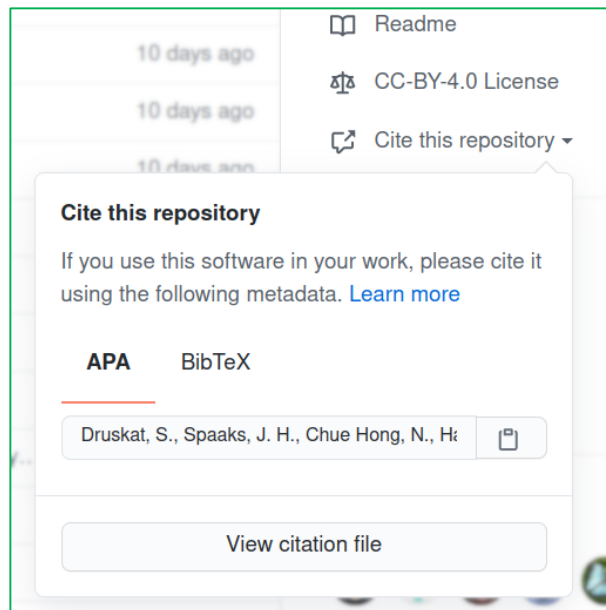
 Below the preview, a checkmark indicates 'Your CITATION.cff is valid' and a 'Download' button is present. The footer shows 'netherlands science center' and 'Version 2.2.0'.





# Citation File Format Integrations

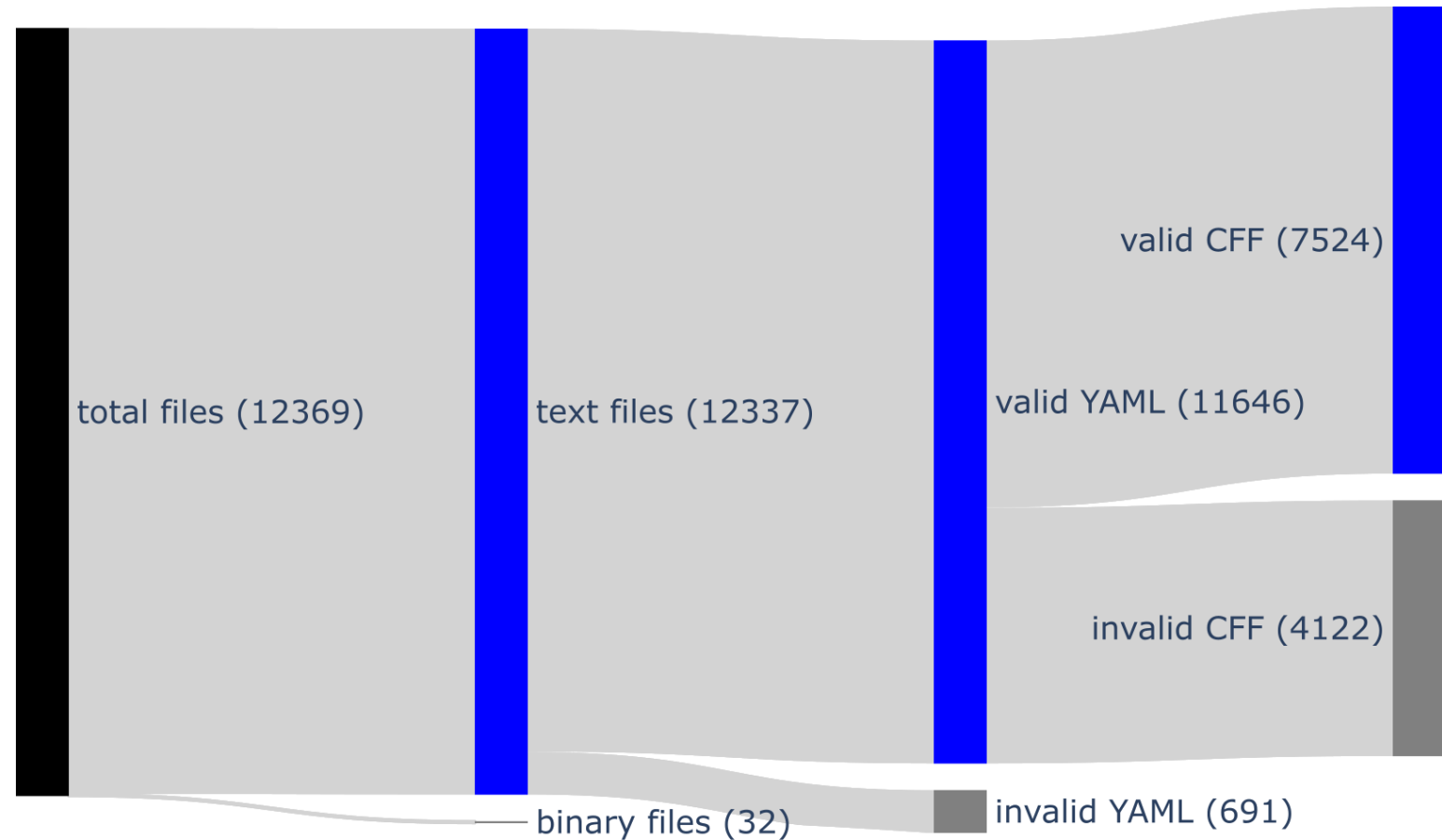
- GitHub
  - render citation information
- Zenodo
  - import metadata from `CITATION.cff`
- Zotero, JabRef reference managers
  - import metadata/file
- IDEs via [schemastore.org](https://schemastore.org)
  - Auto-completion, validation



# Citation File Format Usage

Preliminary study of 12,000+ CITATION.cff files on GitHub:

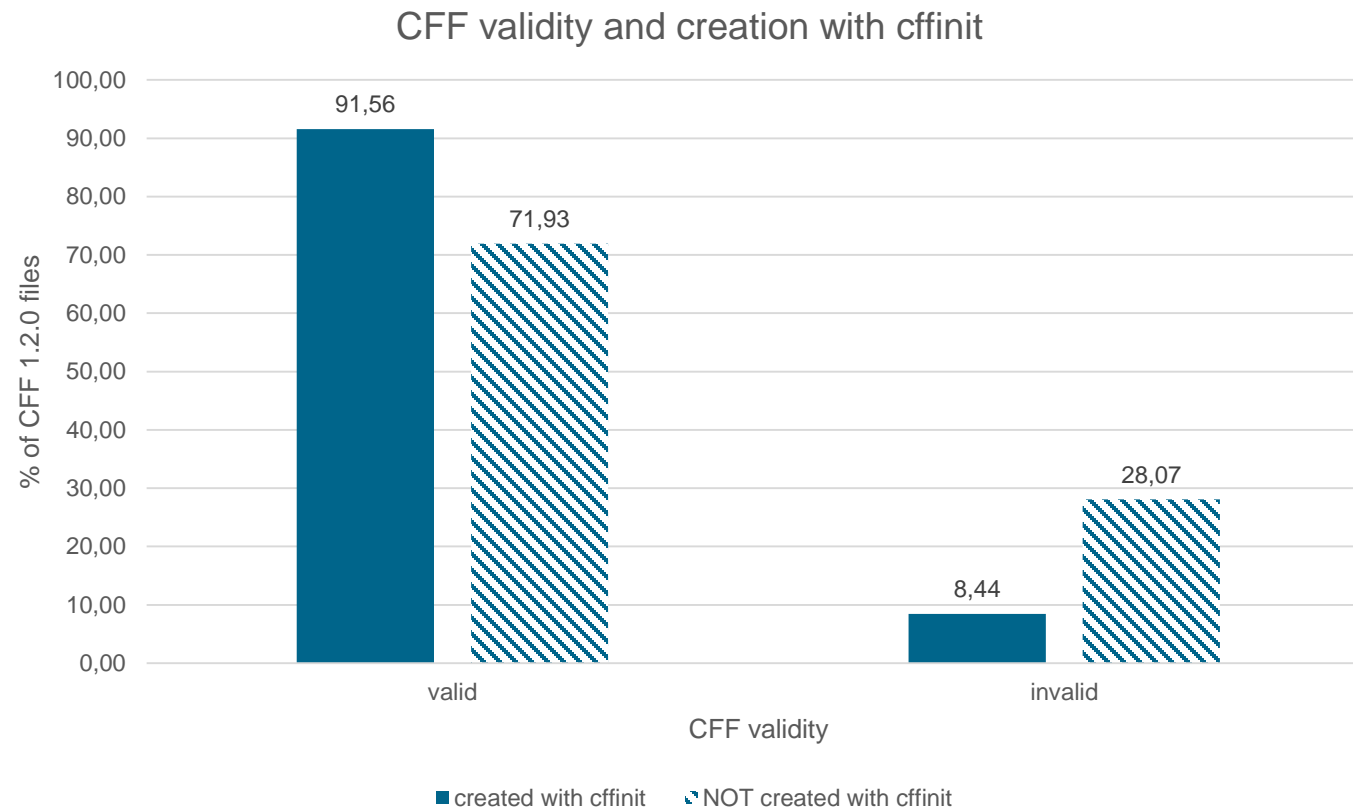
- 60.8% of **all files** are valid CFF files
- 64.6% of **valid YAML files** are valid CFF files



# Citation File Format

## How to improve validity?

- Preliminary study of 12,000+ CITATION.cff files on GitHub





# Citation File Format

## Key points



1. Create a `CITATION.cff` file for your research software
2. Use `cffinit` to create/edit/validate CFF files: [bit.ly/cffinit](https://bit.ly/cffinit)

### References:

- [1] Deutsche Forschungsgemeinschaft. (2022). Guidelines for Safeguarding Good Research Practice. Code of Conduct. <https://doi.org/10.5281/zenodo.6472827>
- [2] Druskat, Stephan, Spaaks, Jurriaan H., Chue Hong, Neil, Haines, Robert, Baker, James, Bliven, Spencer, Willighagen, Egon, Pérez-Suárez, David, & Konovalov, Alexander. (2021). Citation File Format (1.2.0). Zenodo. <https://doi.org/10.5281/zenodo.5171937>

# Thank you



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



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

Contact: [stephan.druskat@dlr.de](mailto:stephan.druskat@dlr.de) | <https://orcid.org/0000-0003-4925-7248>



[@sdruskat@scholar.social](https://scholar.social/@sdruskat)

Funding: Software Sustainability Institute (UK),   
Code for Science & Society (USA),   
CampusSOURCE (DE) 

German Aerospace Center (DE),   
Netherlands eScience Center (NL) 