

Softwarezitierung und das Citation File Format (CFF)

Stephan Druskat

SC-IVS, Sustainable Software Engineering Group

SC TechTalk

2022-06-09



DOI [10.5281/zenodo.6627410](https://doi.org/10.5281/zenodo.6627410)

License [CC-BY-4.0 International](#)



Überblick

1. Problemstellung: Forschungssoftware zitieren und zitierbar machen
2. Das Citation File Format (CFF)
3. Tools
4. Integrationen
5. Das CFF-Projekt
6. Herausforderungen und Ausblick

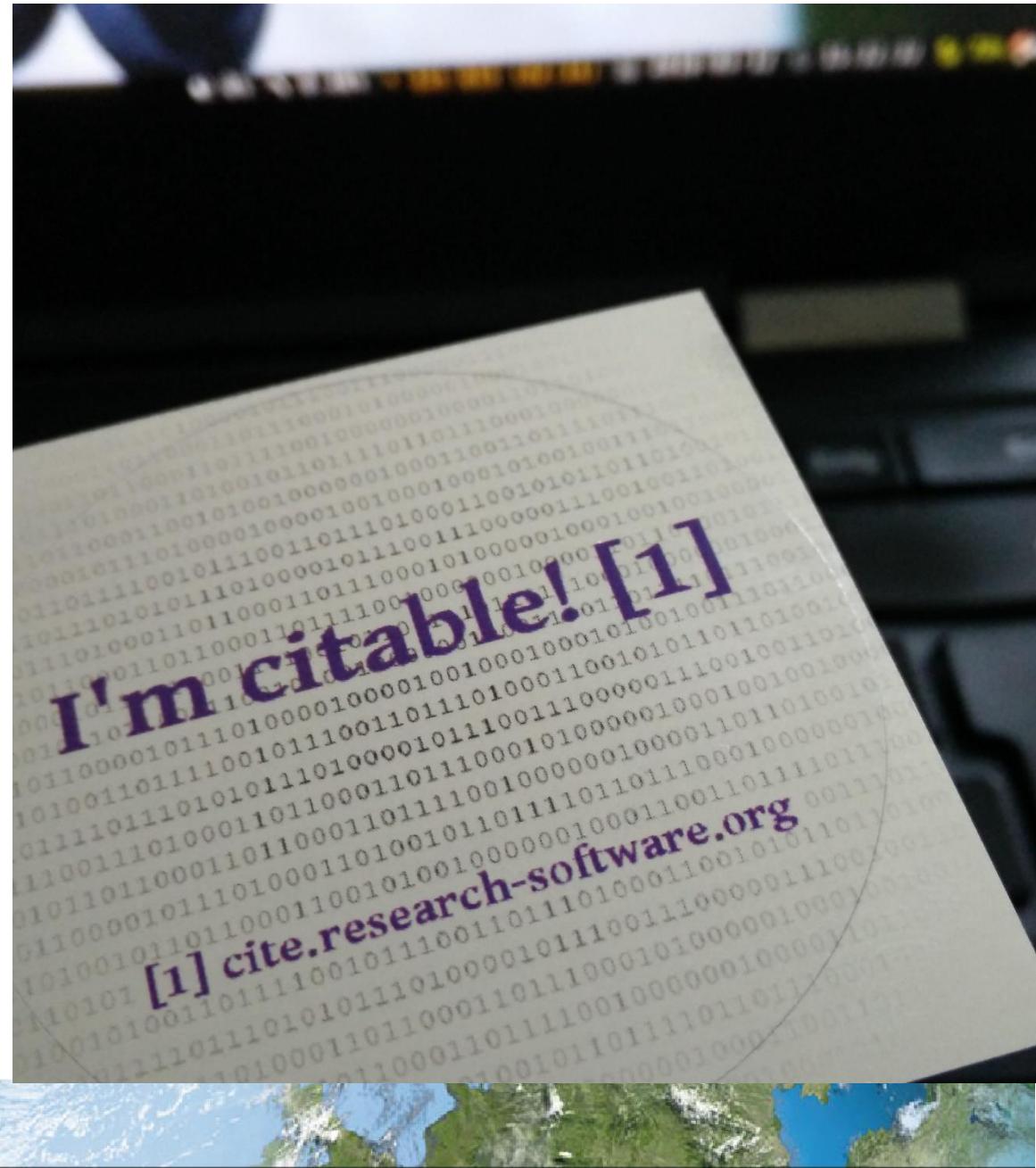


Software zitierbar machen!

- Forschungssoftware als Teil des Scholarly Record

Warum?

- Fair (für Research Software Engineers)
- FAIR
- Reproduzierbarkeit



Prinzipien der Softwarezitierung

The screenshot shows the first page of a PeerJ Computer Science article. The title is "Software citation principles". The authors listed are Arfon M. Smith^{1,2*}, Daniel S. Katz^{2*}, Kyle E. Niemeyer^{3,4} and FORCE11 Software Citation Working Group. The abstract discusses the importance of software in research and the need for citation standards. The article is freely available under a CC-BY 4.0 license.

Software citation principles
Arfon M. Smith^{1,2*}, Daniel S. Katz^{2*}, Kyle E. Niemeyer^{3,4} 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,
Corvallis, Oregon, United States
⁴These 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.

*These authors contributed equally to this work.

Submitted 24 June 2016
Accepted 23 August 2016
Published 19 September 2016
Corresponding author:
Daniel S. Katz, d.katz@eecs.org
Academic editor:
Silvia Peroni
DOI 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 Computer Sci. 2:e86 DOI 10.7717/peerj.cs.86

1. Importance

Software wird genauso zitiert wie ein Paper.

2. Credit and attribution

3. Unique identification

4. Persistence

5. Accessibility

Zitierung ermöglicht den Zugang zu Software und ihren Metadaten.

6. Specificity

Zitation identifiziert die verwendete Softwareversion.

Smith AM, Katz DS, Niemeyer KE, FORCE11 Software Citation Working Group. 2016. Software citation principles. PeerJ Computer Science 2:e86 <https://doi.org/10.7717/peerj.cs.86>

Zitierung: das können wir doch!

Software citation principles

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

¹ GitHub, Inc., San Francisco, California, United States
² National Center for Supercomputing Applications & Electrical and Computer Engineering Department & School of Information Sciences, University of Illinois at Urbana-Champaign, Urbana, Illinois, United States
³ School of Mechanical, Industrial, and Manufacturing Engineering, Oregon State University, Corvallis, Oregon, United States
* These authors contributed equally to this work.

Submitted 24 June 2016
Accepted 23 August 2016
Published 19 September 2016
Corresponding author
Daniel S. Katz, d.katz@iccc.org
Academic editor
Silvio 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

PeerJ Computer Science

Software citation principles

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

¹ GitHub, Inc., San Francisco, California, United States
² National Center for Supercomputing Applications & Electrical and Computer Engineering Department & School of Information Sciences, University of Illinois at Urbana-Champaign, Urbana, Illinois, United States
³ School of Mechanical, Industrial, and Manufacturing Engineering, Oregon State University, Corvallis, 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, review 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

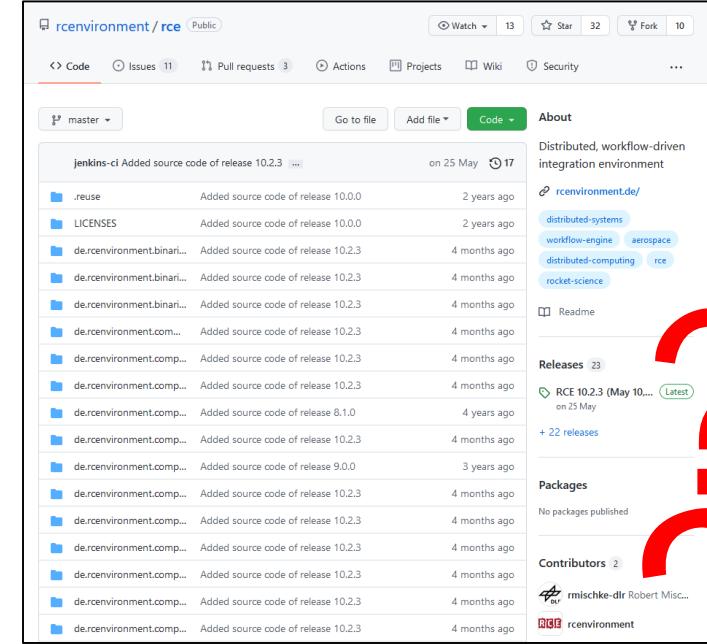
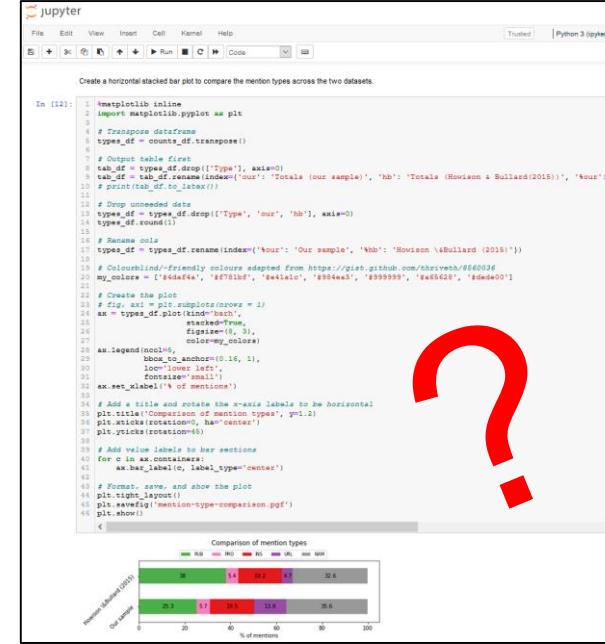
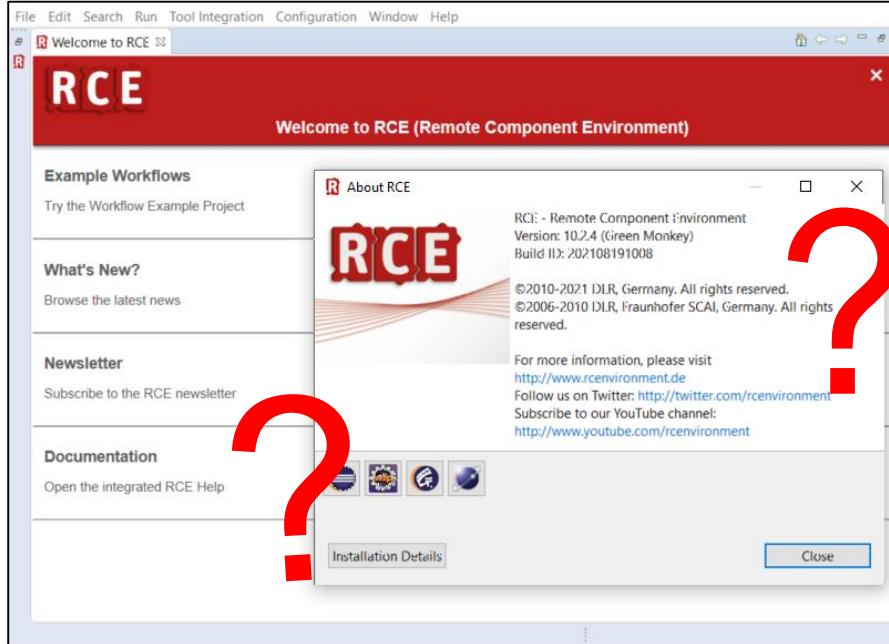
SOFTWARE CITATION PRINCIPLES
The main contribution of this document are the software citation principles, written fairly concisely in this section and discussed further later in the document (see Discussion). In addition, we also motivate the creation of these principles (see Motivation), describe the 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.

How to cite this article Smith et al. (2016), Software citation principles. PeerJ Comput. Sci. 2:e86; DOI [10.7717/peerj-cs.86](https://doi.org/10.7717/peerj-cs.86)

How to cite this article Smith et al. (2016), Software citation principles. PeerJ Comput. Sci. 2:e86; DOI [10.7717/peerj-cs.86](https://doi.org/10.7717/peerj-cs.86)

Softwarezitierung: können wir das?



METADATEN: Name? Autor*innen? Version? Verlag? Veröffentlichungsdatum?

Citation File Format (CFF): ein Format für Softwarezitationsmetadaten

- autoritativ
- kontrollierbar
- „prinzipientreu“:

1. Importance ✓
2. Credit and attribution ✓
3. Unique identification* ✓
4. Persistence* ✓
5. Accessibility** ✓
6. Specificity** ✓

```
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"
```

METADATEN: In einer **CITATION.cff**-Datei im Source Code Repository

CFF: ein einigermaßen mächtiges Format für Softwarezitationsmetadaten

Aber im wissenschaftlichen Reporting zählen nur Papers!

```
preferred-citation:  
authors:  
- family-names: Druskat  
given-names: Stephan  
title: "Software paper about My Research Software"  
type: article  
title: "My Research Software"
```

Kann ich CFF auch für Datensätze verwenden?

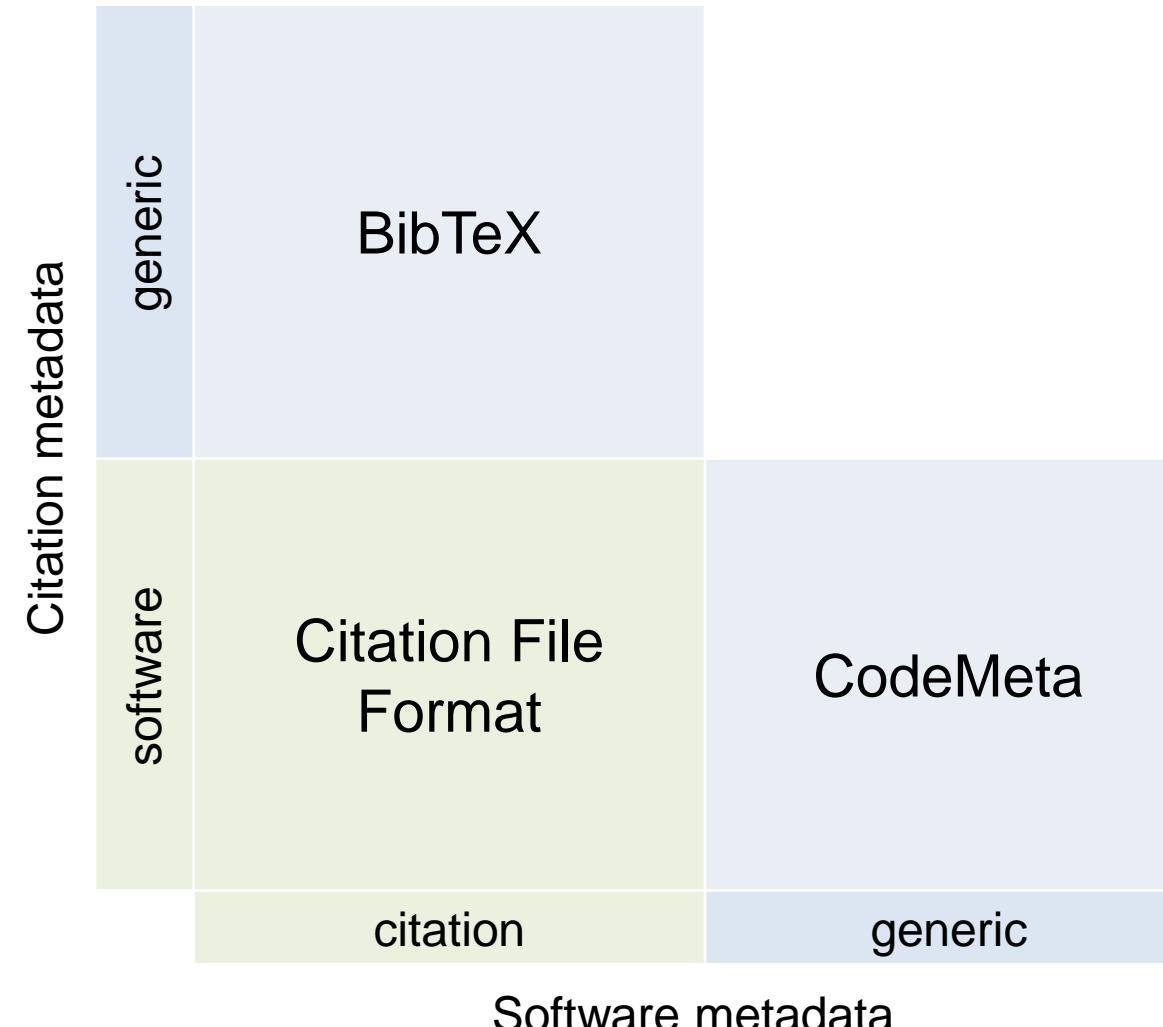
type: software

type: dataset

Software wird genauso zitiert wie ein Paper ...
... auch von Software?

```
references:  
- authors:  
- family-names: Spaaks  
given-names: "Jurriaan H."  
title: "The foundation of Research Software"  
type: software  
- authors:  
- family-names: Haines  
given-names: Robert  
title: "Ruby CFF Library"  
type: software  
version: 1.0
```

Citation File Format: Verortung



Citation File Format: Tooling

The screenshot shows the 'Authors' section of the cffinit web application. On the left, there's a form for entering author details: given-name (Stephan), family-name (Druskat), email (stephan.druskat@dlr.de), and affiliation (German Aerospace Center (DLR)). Below this is an ORCID field with the value <https://orcid.org/0000-0003-4925-7248>. On the right, the generated CFF file content is displayed:

```
# This CITATION.cff file was generated withcffinit.  
# 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'
```

At the bottom, there are buttons for 'Add author', 'Next', 'Download', and 'Done'.

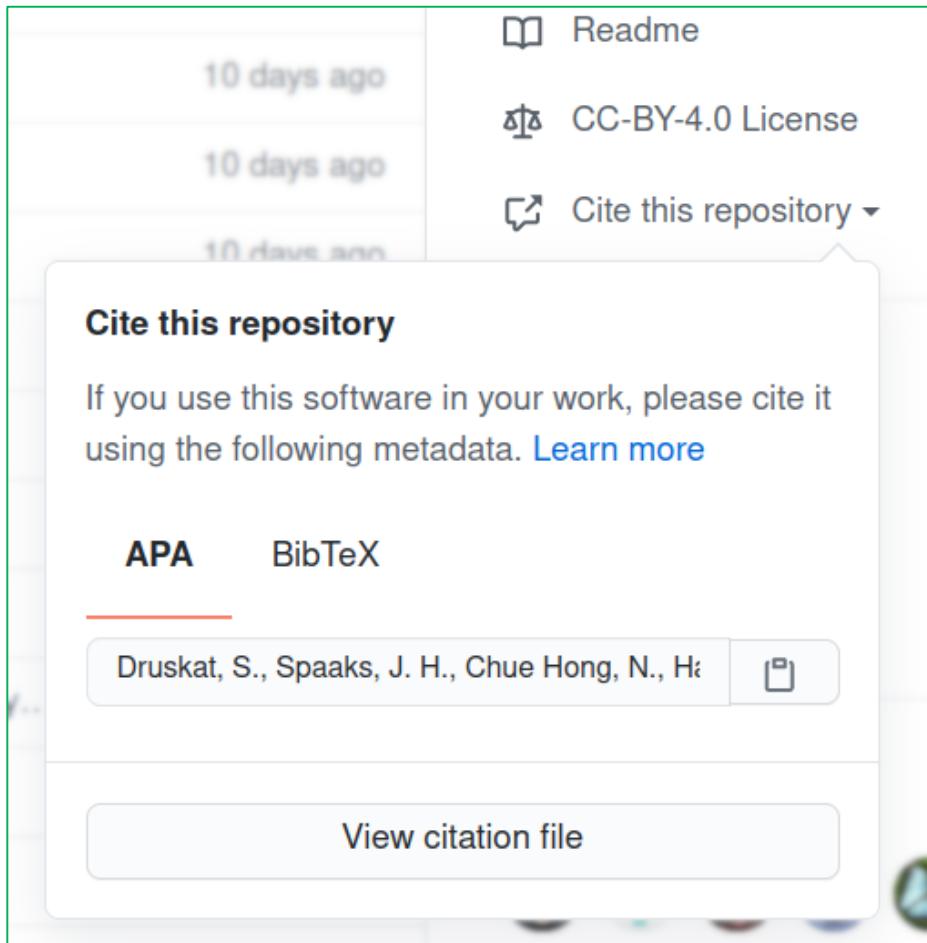
cffinit: Webtool für die Erstellung von CFF-Dateien
<https://bit.ly/cffinit>

- Bearbeiten & aktualisieren
- Validieren
- Konvertieren



Übersicht: <https://bit.ly/cff-tools-list>

Citation File Format: Integrationen I - GitHub



- Erfasst **CITATION.cff**-Dateien
- Rendert Zitationsmetadaten
- Verweist auf konkrete **CITATION.cff**
- Dokumentiert und unterstützt Erstellung

A screenshot of a GitHub code editor window titled 'cfftracker / CITATION.cff in main'. The window shows the following CFF content:

```
cff-version: 1.2.0
message: "If you use this software, please cite it as below."
authors:
- family-names: "YOUR_NAME_HERE"
  given-names: "YOUR_NAME_HERE"
  orcid: "https://orcid.org/0000-0000-0000-0000"
- family-names: "Lisa"
  given-names: "Mona"
  orcid: "https://orcid.org/0000-0000-0000-0000"
title: "cfftracker"
version: 1.0.0
doi: 10.5281/zenodo.1234
date-released: 2021-08-19
url: "https://github.com/sdruskat/cfftracker"
```

Citation File Format: Integrationen II - Zenodo

March 16, 2022

Software Open Access

sdruskat/campussource: v0.1.0

Stephan Druskat

A release without a CFF file.

Preview

campusource-0.1.0.zip

sdrukcat-campussource-a46ecd3

README.md 49 Bytes

```
1  cff-version: 1.2.0
2  message: "If you use this software, please cite it as below."
3  authors:
4    - family-names: "Druskat"
5      given-names: "Stephan"
6      orcid: "https://orcid.org/0000-0003-4925-7248"
7  title: "CampusSource Example Deposit"
8  version: 0.2.0
9  doi: 10.5281/zenodo.1035710
10 date-released: 2022-03-16
11 url: "https://www.campussource.de/events/e2203hagen/#Programm"
```

citation-file-format/citation-file-format

DOI: 10.5072/zenodo.896351

GitHub / Releases

1.2.0 citation-file-format/citation-file-format: 1.2.0

DOI: 10.5072/zenodo.896351

Published 7 months ago

1.2.0

Citation File Payload Metadata

Create CITATION.cff

Citation File

CITATION.cff files are plain text files with human- and machine-readable citation information for software. Code developers can include them in their repositories to let others know how to correctly cite their software.

An example of the CITATION.cff for this release can be found below:

```
cff-version: 1.1.0
message: "If you use this software, please cite it as below."
authors:
- family-names: Joe
  given-names: Johnson
  orcid: https://orcid.org/0000-0000-0000-0000
title: citation-file-format/citation-file-format: 1.2.0
version: 1.2.0
date-released: 2017-12-18
```

March 16, 2022

Software Open Access

CampusSource Example Deposit

Druskat, Stephan

This is a release WITH a CITATION.cff file :tada:.

If you use this software, please cite it as below.

Preview

campusource-0.2.0.zip

sdrukcat-campussource-1211c23

CITATION.cff 360 Bytes

README.md 49 Bytes

Citation File Format: Integrationen III - Referenzmanager

Import in Zotero aus GitHub über Browserplugin

The screenshot shows the Zotero interface with a citation entry for 'Hexatomic'. The citation key is 'druskatHexatomic2021'. The entry details are as follows:

- Item Type: Software
- Title: Hexatomic
- Programmer: Druskat, Stephan
- Programmer: Krause, Thomas
- Abstract: Hexatomic is an extensible software for deep multi-layer annotation of linguistic corpora
- Series Title: Version 0.6.0-SNAPSHOT
- Date: 2021-03
- System
- Place
- Company
- Prog. Language: Java
- ISBN
- Short Title:
 - URL: <https://github.com/hexatomic/hexatomic>
 - Rights: Apache-2.0
 - Archive
 - Loc. in Archive
 - Library Catalog: GitHub

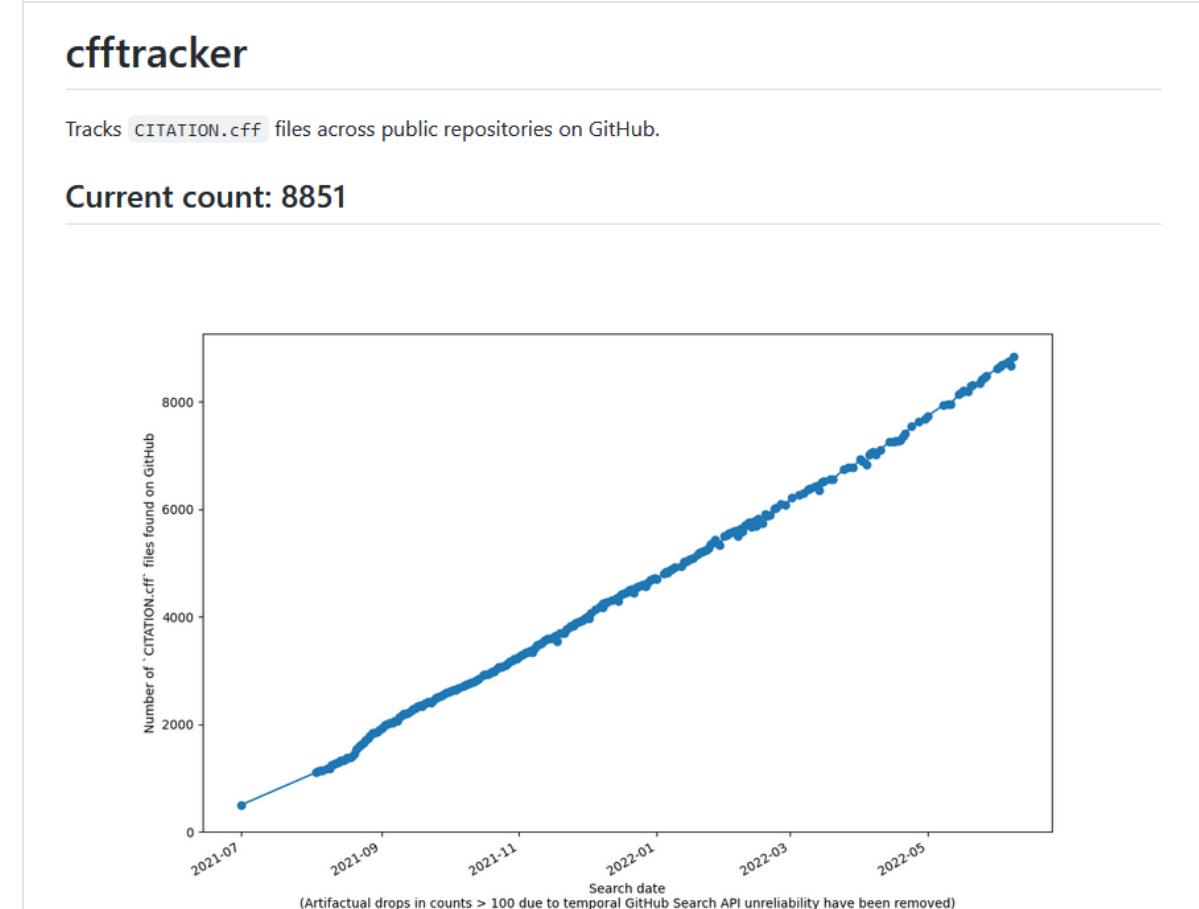
Import in JabRef aus CITATION.cff-Datei

The screenshot shows the JabRef interface with a citation entry for 'Citation File Format'. The entry details are as follows:

- Software (Druskat2021)
- Author: Stephan Druskat and Jurriaan H. Spaaks and Neil Chue Hong and Robert Baker, James Bliven, Spencer Willighagen, Egon Pérez-Suárez, David Konovalov, Alexander
- Title: Citation File Format
- Editor
- Date: 2021-08-09
- Abstract: CITATION cff files are plain text files with human- and machine-readable citation information for software. Code developers can include them in their repositories to let others know how to correctly cite their software. This is the specification for the Citation File Format.
- Comment: If you use CFF in your research, please cite it using these metadata.
- Citationkey: Druskat2021

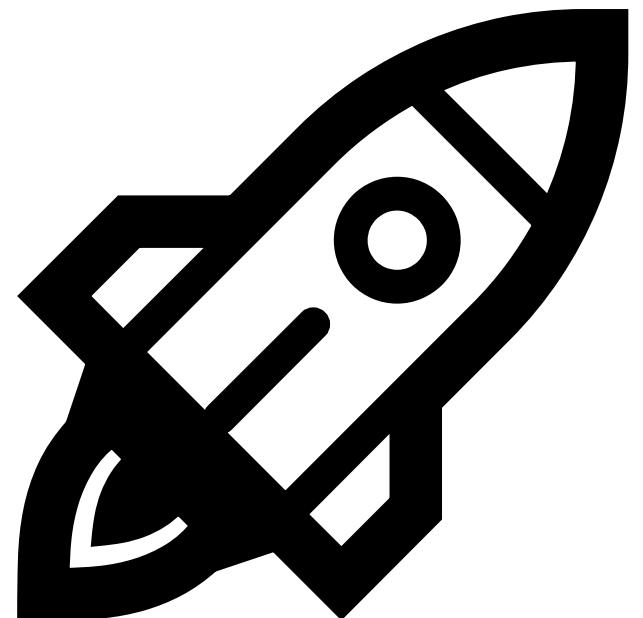
CFF als Projekt

- Community-Projekt (seit WSSSPE5.1, 2017)
- Hauptentwicklung: DLR, NLeSC, U Manchester
- Community-Events
- Offene Entwicklung:
<https://github.com/citation-file-format/>
- Durchgängig offene Lizenzen:
CC0, CC-BY, MIT, Apache 2.0
- Community-Dokumentation
- Kollaboration mit externen Projekten
- Eigenes Kapitel in [The Turing Way](#)
- Governance-Entwicklung als Teil des CS&S [Digital Infrastructure Incubator](#)
- CampusSource Award 2022



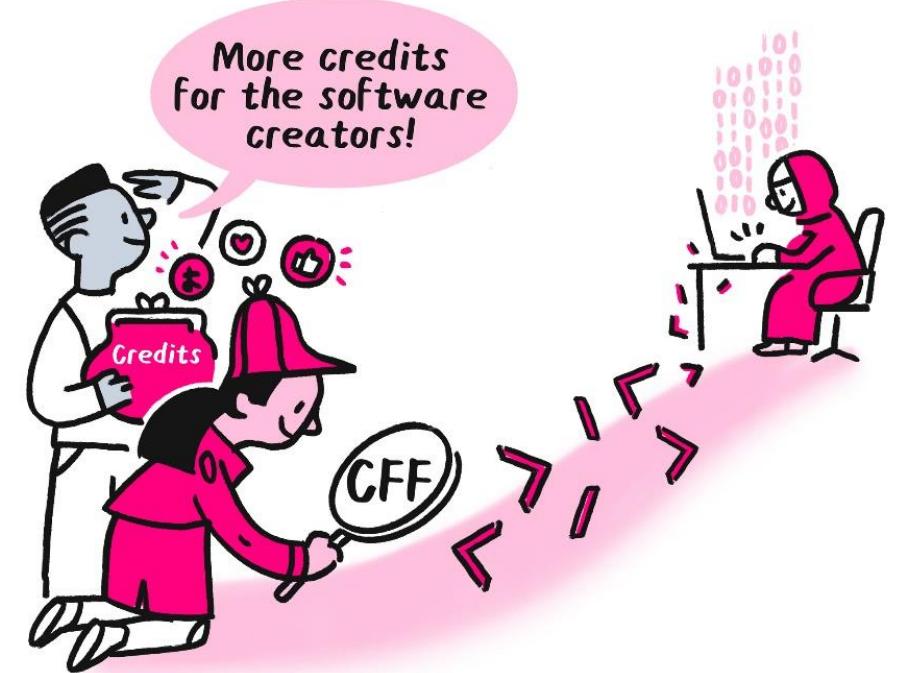
Herausforderungen und Ausblick

- Wachstumsfolgen
- Governance
- Finanzierung
- Weiterentwicklung von Schema und Tooling
- Entwicklung weiterer Integrationen
- Weiterhin Vernetzung mit anderen (Forschungs-) Projekten



Danke!

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



More credits for the software creators. The Turing Way project illustration by Scriberia. Zenodo.
<https://doi.org/10.5281/zenodo.3332807>, License: CC BY-4.0

Scriberia

Bildcopyrights und Lizenzen

- GitHub Actions Logo verwendet unter limited fair use
- Docker Logo verwendet unter limited fair use
- Go Logo © 2018 The Go Authors, verwendet unter limited fair use
- Duke Logo © Joe Palrang, lizenziert unter [New BSD license](#)
- JavaScript Logo unter Public Domain
- Julia Logo © 2012-2019 Stefan Karpinski, lizenziert unter [CC BY-NC-SA 4.0](#)
- PHP Logo von Colin Viebrock, lizenziert unter [CC BY-SA 4.0](#)
- Python Logo verwendet unter der [Python Software Foundation Trademark Usage Policy](#)
- R Logo © 2016 The R Foundation, lizenziert unter [CC BY-SA 4.0](#)
- Ruby Logo © 2006 Yukihiro Matsumoto, lizenziert unter [CC BY-SA 2.5](#)
- TypeScript Logo verwendet unter limited fair use
- Twitter Logo © Twitter Inc., lizenziert unter [Apache-2.0](#)

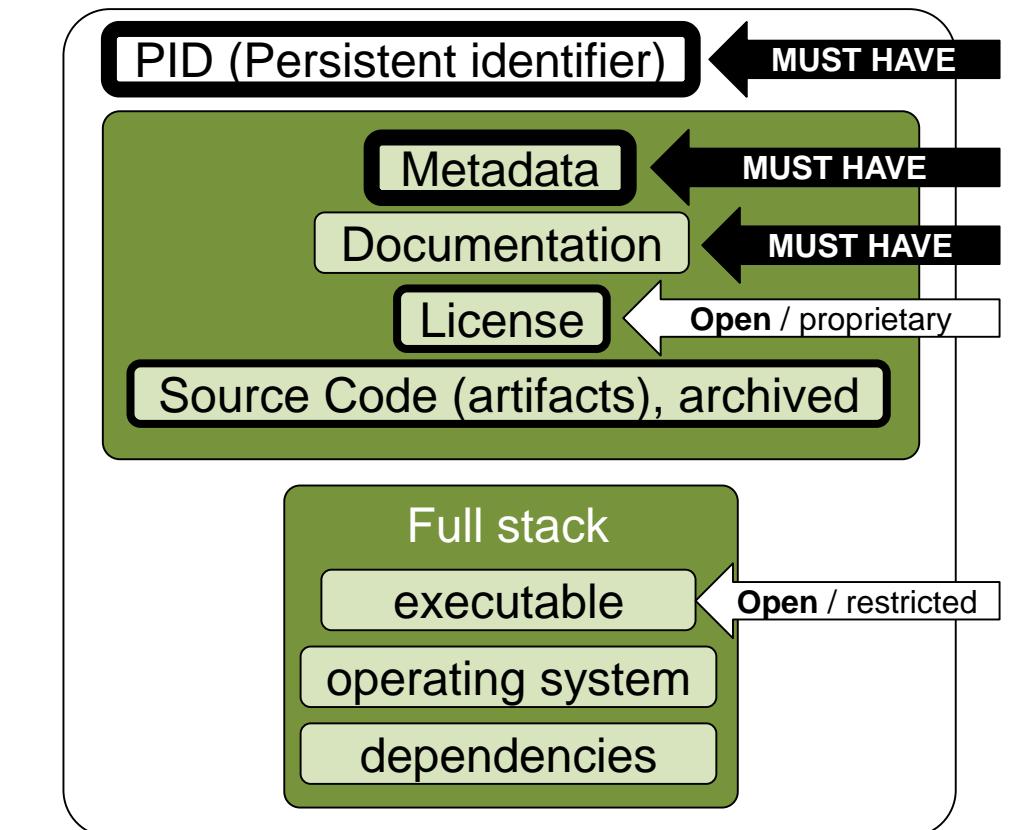


Appendix I

Publication and metadata as prerequisites for FAIR and citable software

- PID
 - FAIR software needs a persistent identifier (and PIDs greatly help citation as well!)
- Publication
 - You get a persistent identifier when you publish software (e.g., on Zenodo)
- Metadata
 - To publish software (and for FAIR), you need relevant metadata
- Citation metadata
 - Relevant metadata includes citation metadata (+ accessibility metadata + reusability metadata)

FAIR software, Open Source and Sustainable



Adapted from erratum for [1], <https://doi.org/10.1016/j.patter.2021.100267>.