

Das Citation File Format (CFF): Forschungssoftware zitierbar machen

Stephan Druskat

Deutsches Zentrum für Luft- und Raumfahrt (DLR)

Institut für Softwaretechnologie

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

DOI [10.5281/zenodo.6364654](https://doi.org/10.5281/zenodo.6364654)



Wissen für Morgen



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
7. CFF und der CampusSOURCE Award

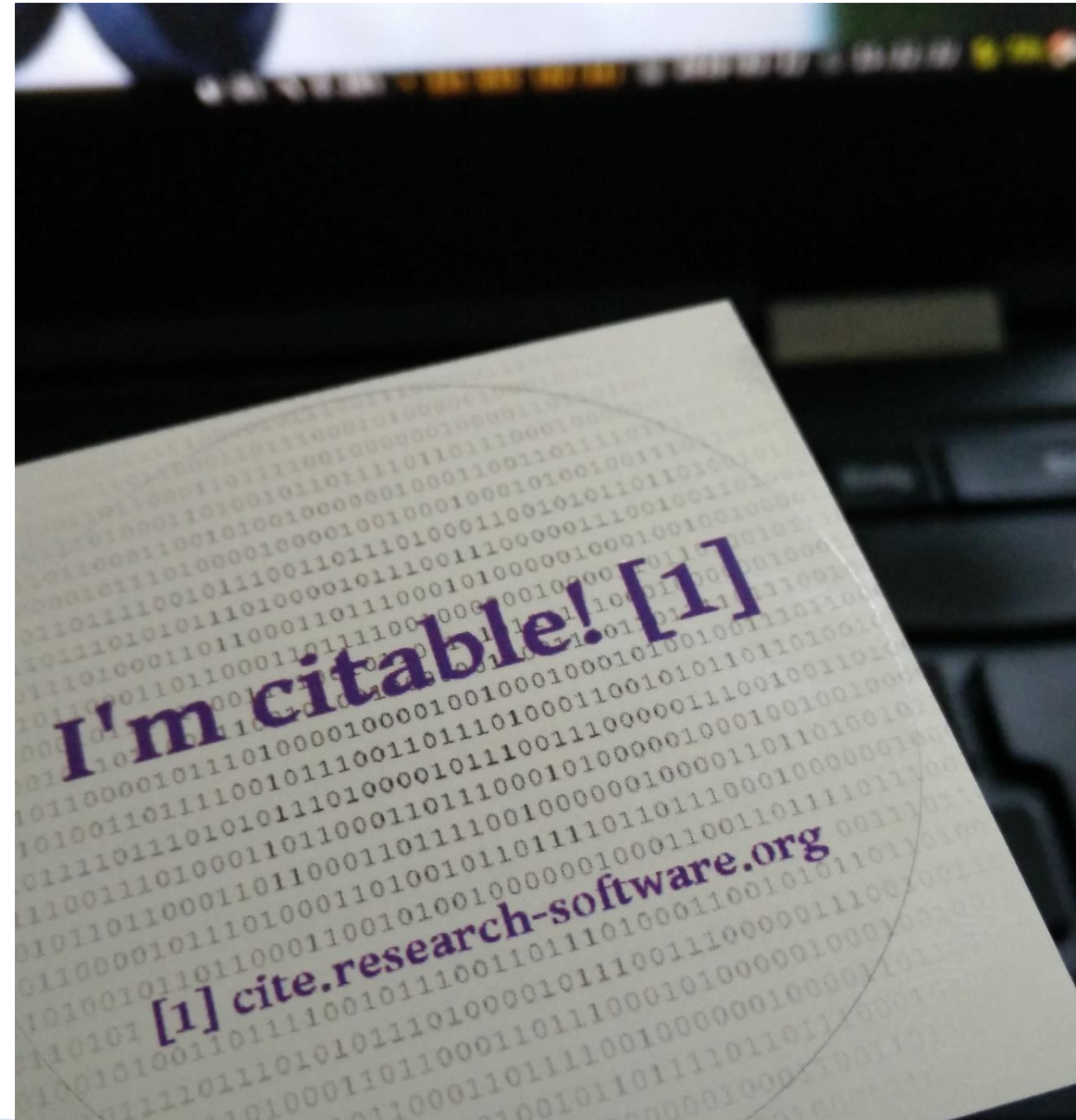


Software zitierbar machen!


- Forschungssoftware als Teil des Scholarly Record

Warum?

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



Prinzipien der Softwarezitierung



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, 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@rice.edu
 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

[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)

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

© Copyright
2016 Smith et al.

Distributed under
Creative Commons CC-BY 4.0

OPEN ACCESS



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, 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.

[How to cite this article](#) Smith et al. (2016), Software citation principles. PeerJ Comput. Sci. 2:e86; DOI 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

Softwarezitierung: können wir das?

File Edit Search Run Tool Integration Configuration Window Help

Welcome to RCE

Welcome to RCE (Remote Component Environment)

Example Workflows

Try the Workflow Example Project

What's New?

Browse the latest news

Newsletter

Subscribe to the RCE newsletter

Documentation

Open the integrated RCE Help

About RCE

RCE: - Remote Component Environment
Version: 10.2.4 (Green Monkey)
Build ID: 202108191008

©2010-2021 DLR, Germany. All rights reserved.
©2006-2010 DLR, Fraunhofer SCAI, Germany. All rights reserved.

For more information, please visit
<http://www.rcenvironment.de>
Follow us on Twitter: <http://twitter.com/rcenvironment>
Subscribe to our YouTube channel:
<http://www.youtube.com/rcenvironment>

Installation Details

Close

```
In [12]: 1 import matplotlib
2 import matplotlib.pyplot as plt
3
4 # Transpose dataframe
5 types_df = counts_df.transpose()
6
7 # Output table first
8 tab_df = types_df.drop(['Type'], axis=0)
9 tab_df = tab_df.rename(index={'our': 'Totals (our example)', 'hb': 'Totals (Houston & Bullard(2016))', 'house':
10 # print(tab_df.to_latex())
11
12
13 # Drop unneeded data
14 types_df = types_df.drop(['Type', 'our', 'hb'], axis=0)
15 types_df.rouse()
16
17 # Rename cols
18 types_df = types_df.rename(index={'our': 'Our example', 'hb': 'Houston & Bullard (2016)'})
19
20 # Colorblind-friendly colours adapted from https://gist.github.com/chriveth/866034
21 my_colors = ['#66a6c1', '#f79646', '#a65628', '#c44e52', '#808080', '#4d4d4d']
22
23 # Create the plot
24 # Fig. ax1 = plt.subplots(nrows = 2)
25 ax = types_df.plot(kind='bar',
26                 stacked=True,
27                 figsize=(8, 3),
28                 color=my_colors)
29 ax.legend(bbox=(
30     bbox_to_anchor=(0.15, 1),
31     loc='lower left',
32     fontsize='small')
33 ax.set_xlabel('% of mentions')
34
35 # Add a title and rotate the x-axis labels to be horizontal
36 plt.title('Comparison of mention types', y=1.2)
37 plt.xticks(rotation=45)
38 plt.yticks(rotation=45)
39
40 # Add y-axis labels to bar sections
41 for c in ax.containers:
42     ax.bar_label(c, label_type='percent')
43
44 # Format, save, and show the plot
45 plt.tight_layout()
46 plt.savefig('mention-type-comparison.pdf')
47 plt.show()
```

Create a horizontal stacked bar plot to compare the mention types across the two datasets.

Comparison of mention types

Mention Type	our	hb
Person	14	14
Organization	11	11
Location	11	11
Product	11	11
Event	11	11
Other	11	11

rcenvironment / rcenvironment Public

<> Code Issues 11 Pull requests 3 Actions Projects Wiki Security

master

Go to file Add file Code About

Distributed, workflow-driven integration environment

rcenvironment.de

- distributed-systems
- workflow-engine aerospace
- distributed-computing rce
- rocket-science

Readme

Releases 23

RCE 10.2.3 (May 10... Latest) on 23 May

+ 22 releases

Packages

No packages published

Contributors 2

- mischke-dir Robert Misc...
- rcenvironment

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
```

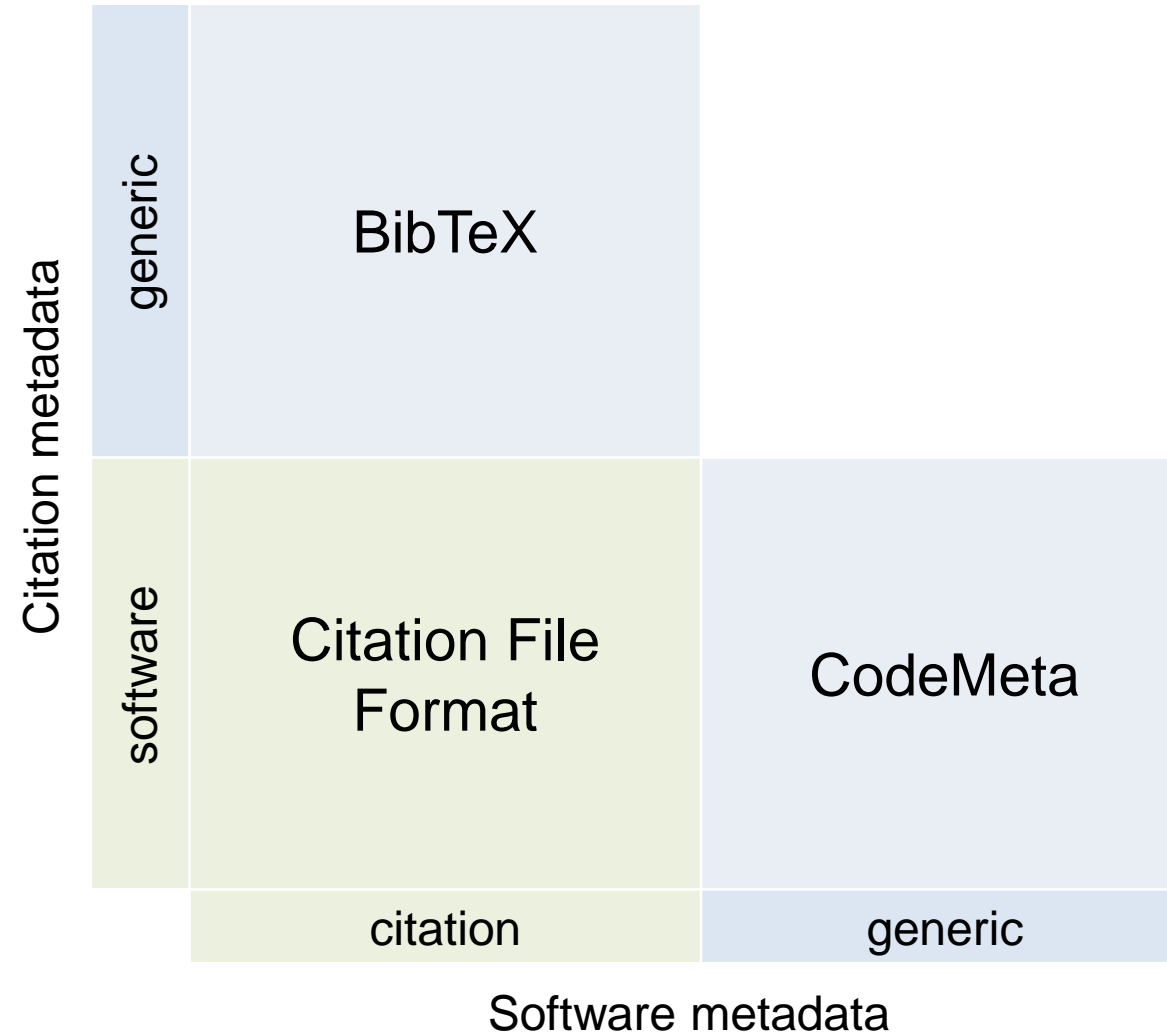
```
type: dataset
```

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

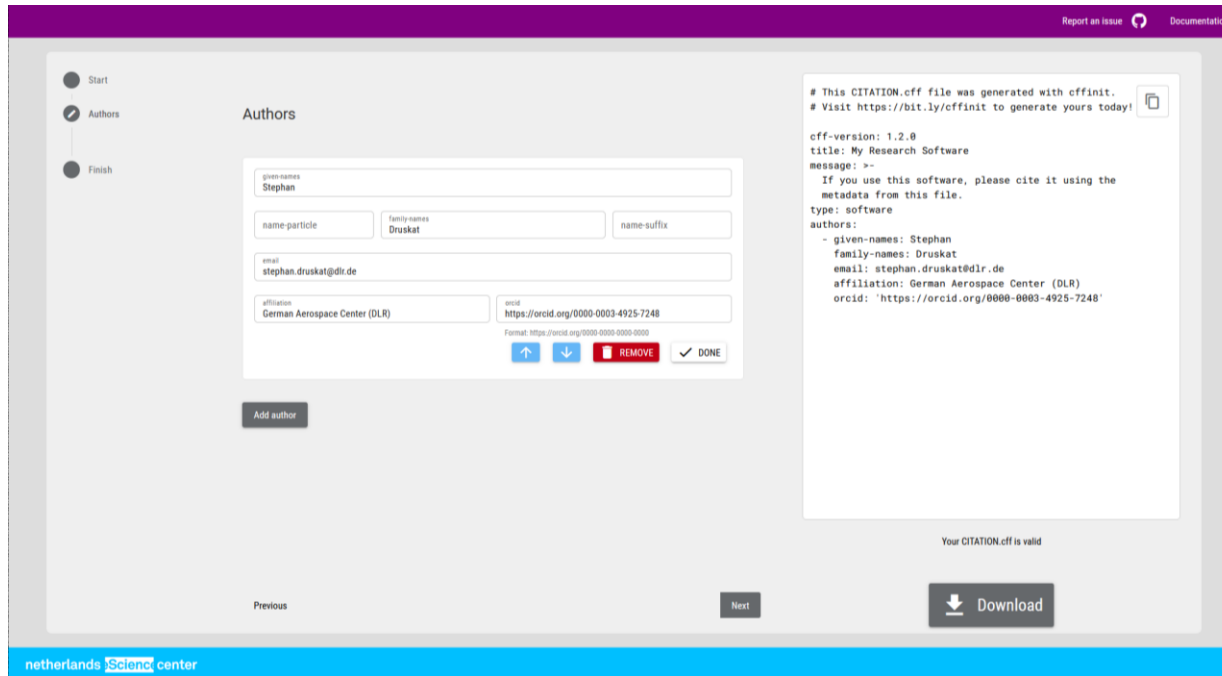
```
references:  
  - authors:  
    - family-names: Spaaks  
      given-names: "Jurriaan H."  
    title: "The foundation of Research Software"  
    type: software  
  - authors:  
    - family-names: Haines  
      given-names: Robert  
    title: "Ruby CFF Library"  
    type: software  
    version: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



Citation File Format: Verortung



Citation File Format: Tooling



cffinit: Webtool für die Erstellung von CFF-Dateien

<https://bit.ly/cffinit>

- Bearbeiten & aktualisieren
- Validieren
- Konvertieren



Citation File Format: Integrationen I - GitHub

10 days ago

10 days ago

10 days ago

Readme

CC-BY-4.0 License

Cite this repository ▾

Cite this repository

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

APA BibTeX

Druskat, S., Spaaks, J. H., Chue Hong, N., H...

View citation file

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

cfftracker / CITATION.cff in main Cancel changes

Adding a **CITATION.cff** file helps users to easily cite your software from the repository overview. [Learn more](#) Insert example

<> Edit new file Preview Spaces 2 No wrap

```

1 cff-version: 1.2.0
2 message: "If you use this software, please cite it as below."
3 authors:
4 - family-names: "YOUR_NAME_HERE"
5   given-names: "YOUR_NAME_HERE"
6   orcid: "https://orcid.org/0000-0000-0000-0000"
7 - family-names: "Lisa"
8   given-names: "Mona"
9   orcid: "https://orcid.org/0000-0000-0000-0000"
10 title: "cfftracker"
11 version: 1.0.0
12 doi: 10.5281/zenodo.1234
13 date-released: 2021-08-19
14 url: "https://github.com/sdruskat/cfftracker"
15

```

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

- campussource-0.1.0.zip
- sdruskat-campussource-a46ecd3
 - README.md 49 Bytes

citation-file-format/citation-file-format ON

DOI: 10.5072/zenodo.896351

GitHub / Releases Create release ...

1.2.0 citation-file-format/citation-file-format: 1.2.0 Published
DOI: 10.5072/zenodo.896351 7 months ago

1.2.0

Citation File Create CITATION.cff

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
```

```
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"
```

March 16, 2022 Software Open Access

CampusSource Example Deposit

id Druskat, Stephan

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

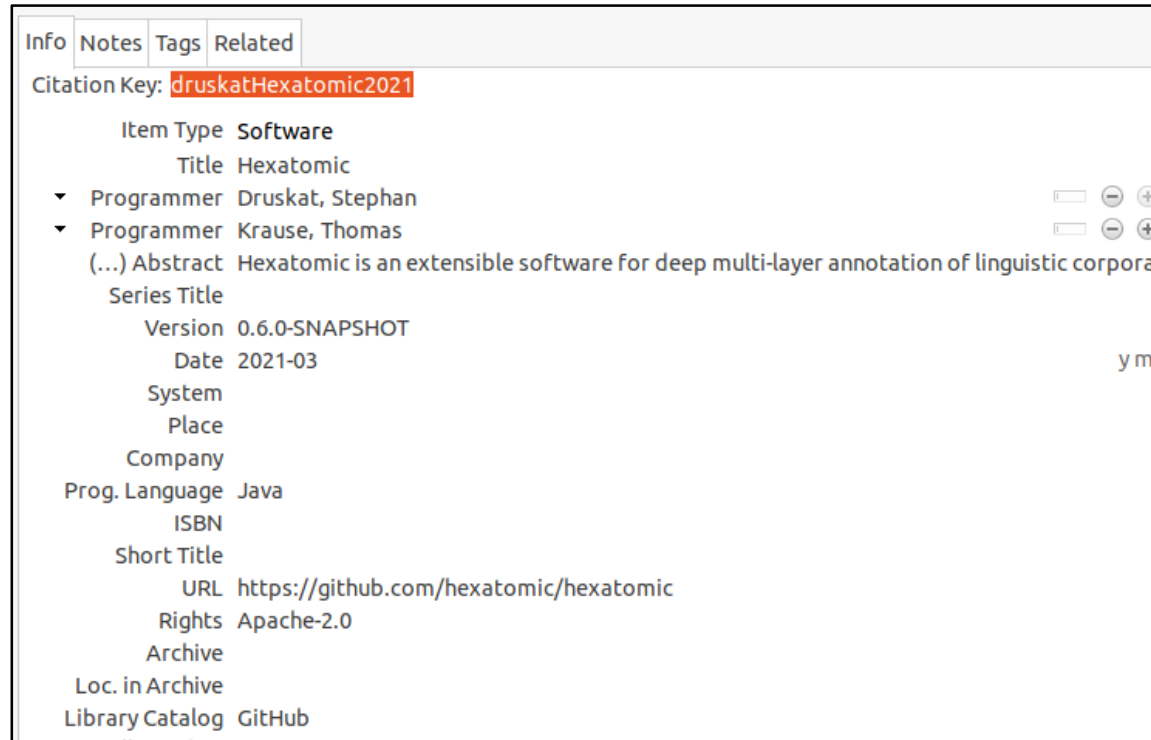
If you use this software, please cite it as below.

Preview

- campussource-0.2.0.zip
- sdruskat-campussource-1211c23
 - CITATION.cff 360 Bytes
 - README.md 49 Bytes

Citation File Format: Integrationen III - Referenzmanager

Import in Zotero aus GitHub über Browserplugin



Info Notes Tags Related

Citation Key: **druskatHexatomic2021**

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>

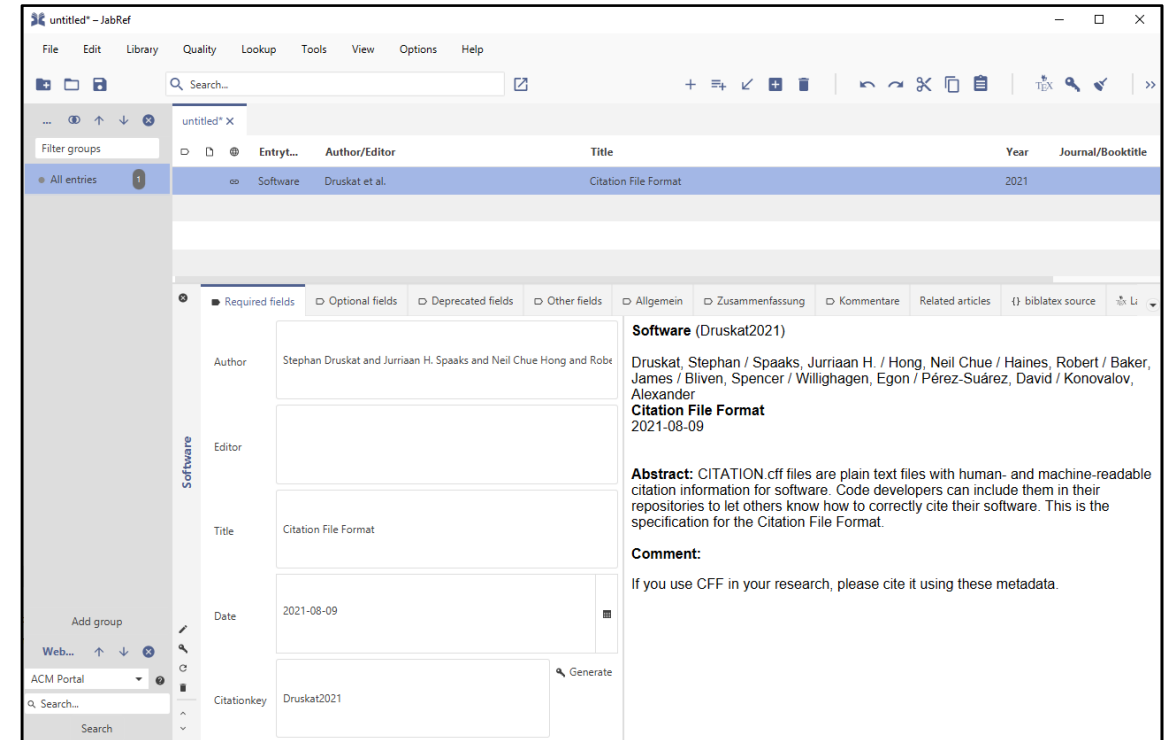
Rights Apache-2.0

Archive

Loc. in Archive

Library Catalog GitHub

Import in JabRef aus CITATION.cff-Datei



untitled* - JabRef

File Edit Library Quality Lookup Tools View Options Help

Search...

Filter groups	Entry...	Author/Editor	Title	Year	Journal/Booktitle
All entries	Software	Druskat et al.	Citation File Format	2021	

Required fields Optional fields Deprecated fields Other fields Allgemein Zusammenfassung Kommentare Related articles biblatex source

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

Editor

Title Citation File Format

Date 2021-08-09

Citationkey Druskat2021

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.

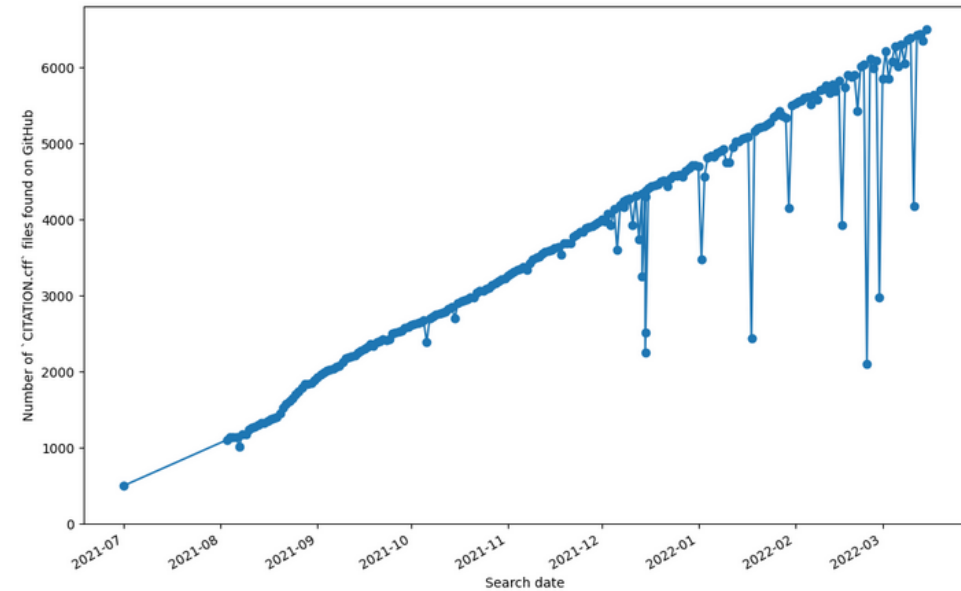
CFF als Projekt

- 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](#)

cfftracker

Tracks CITATION.cff files across public repositories on GitHub.

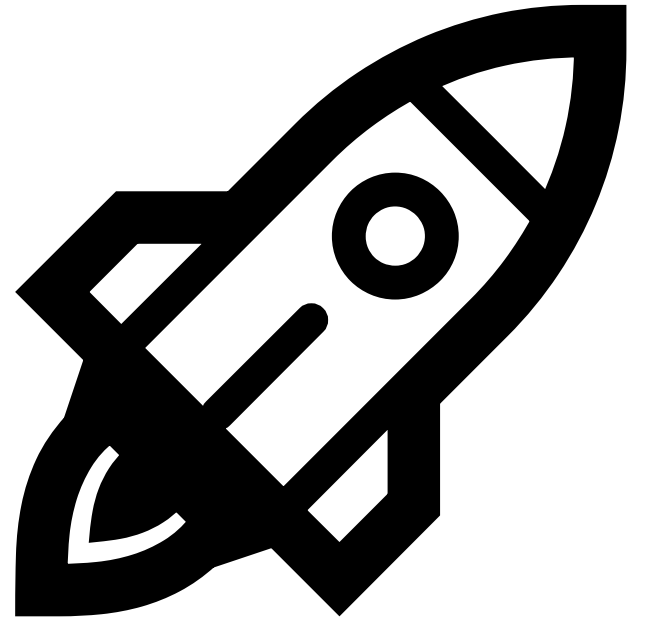
Current count: 6500



Herausforderungen und Ausblick

- Wachstumsfolgen
- Governance
- Finanzierung

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




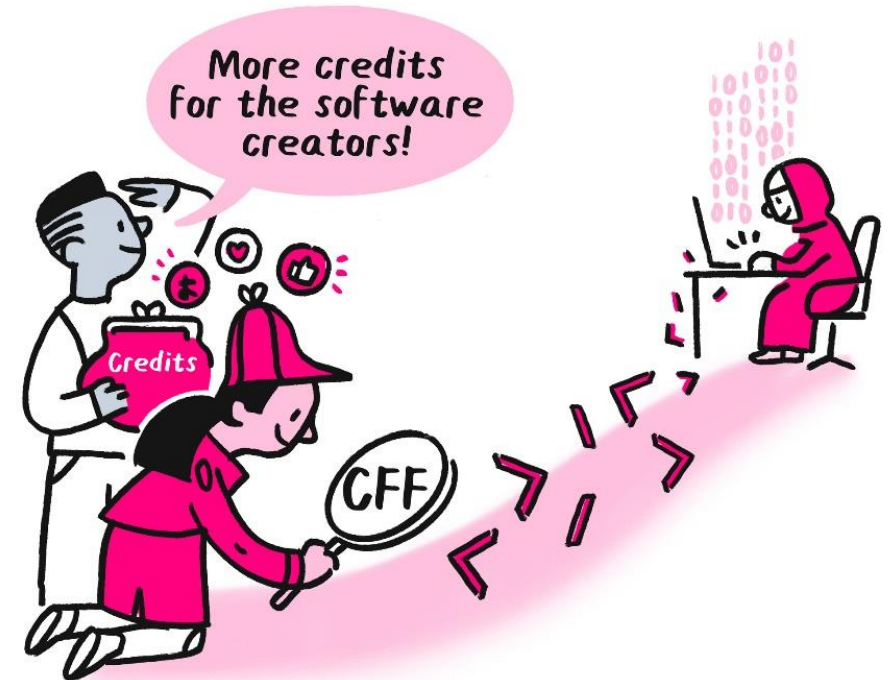
CFF und der CampusSOURCE Award

- Verbesserung der Sichtbarkeit von Forschungssoftware, ihrer Entwickler*innen/Maintainer*innen
- Verbesserung der Bedingungen für die Entwicklung von Forschungssoftware
- Unterstützung von nachhaltiger Forschungssoftware
- Verbesserung der Qualität von Forschungssoftware



Danke!

- <https://citation-file-format.github.io>
- <https://github.com/citation-file-format/citation-file-format/>
- stephan.druskat@dlr.de
- <https://sdruskat.net>
-  [@stdruskat](https://twitter.com/stdruskat)



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](#)

