

HELMHOLTZ

Open Science

3. Helmholtz Open Science Forum

Forschungssoftware

Helmholtz Open Science Briefing

Version 1.0

Impressum

Die Onlineversion dieser Publikation finden Sie unter:

<https://doi.org/10.48440/os.helmholtz.057>

Verfasser:innen

Heinz Pampel, Lea Maria Ferguson, Stephan Druskat, Uwe Konrad, Martin Hammitzsch, Christian Meeßen, David Schäfer, Robert Speck, Markus Diesmann, Achim Streit, Rene Caspart

Herausgegeben von

Helmholtz Open Science Office

Redaktion

Heinz Pampel, Roland Bertelmann, Christoph Bruch, Lea Maria Ferguson, Antonia C. Schrader, Paul Schultze-Motel, Nina Leonie Weisweiler

Kontakt

Helmholtz-Gemeinschaft
Helmholtz Open Science Office
c/o Helmholtz-Zentrum Potsdam
Deutsches GeoForschungsZentrum GFZ
Telegrafenberg, 14473 Potsdam
E-Mail: open-science@helmholtz.de

Stand

09.12.2022. Version. 1.0

Lizenz

Alle Texte dieser Veröffentlichung, ausgenommen Zitate, sind unter einem Creative Commons Attribution 4.0 International (CC BY 4.0) Lizenzvertrag lizenziert. Siehe:

<https://creativecommons.org/licenses/by/4.0>.



HELMHOLTZ

Open Science

Inhalt

Einleitung.....	3
Vorträge	5
Update zu HIFIS, Helmholtz Federated IT	5
SaQC - System for automated Quality Control	6
Citation File Format (CFF)	6
Research Software Directory (RSD)	7
Helmholtz Research Software Award	7
HiRSE_PS	7
Abschlussdiskussion.....	7
Anhang	9

HELMHOLTZ

Open Science

Abstract

Am 24. November 2022 veranstaltete das Helmholtz Forum Forschungssoftware eine Informationsveranstaltung zu aktuellen Entwicklungen im Bereich Forschungssoftware in Helmholtz. Das Helmholtz Forum Forschungssoftware wird gemeinsam von der Task Group Forschungssoftware des AK Open Science und dem HIFIS Software Cluster getragen. Die Veranstaltung wurde unter dem Titel „3. Helmholtz Open Science Forum Forschungssoftware“ vom Helmholtz Open Science Office organisiert. Eine erste Veranstaltung des Helmholtz Forum Forschungssoftware fand im Mai 2021 und eine zweite im April 2022 statt. Der vorliegende Bericht dokumentiert die erfolgreiche Veranstaltung, an der rund 90 Mitarbeitende aus Helmholtz teilgenommen haben.

HELMHOLTZ

Open Science

Einleitung

Mit der wachsenden Durchdringung der Digitalisierung steigt die Zahl an Software-Lösungen, die an wissenschaftlichen Einrichtungen entstehen. Ob zur Analyse großer Datenmengen, zur Steuerung von Instrumenten oder zur Visualisierung von Ergebnissen – Forschungssoftware ist heute ein zentrales Element im wissenschaftlichen Erkenntnisprozess. Zur Umsetzung von Open Science müssen in vielen Fachgebieten auch die Programmcodes offen zugänglich sein.¹

Dieser Herausforderung widmet sich Helmholtz unter dem Begriff Open Research Software. Seit 2020 agiert das Helmholtz Forum Forschungssoftware als Netzwerk zum Thema. Das Helmholtz Forum Forschungssoftware wird gemeinsam von der Task Group Forschungssoftware des AK Open Science² und dem HIFIS Software Cluster³ getragen.

Am 24. November 2022 von 10:00 bis 12:30 Uhr veranstaltete das Helmholtz Forum Forschungssoftware eine Informationsveranstaltung zu aktuellen Entwicklungen im Bereich Forschungssoftware in Helmholtz. Diese Veranstaltung war die dritte einer Reihe von Helmholtz Open Science Foren, die sich dem Thema Forschungssoftware widmeten. Der vorliegende Bericht dokumentiert diese Veranstaltung. Teilgenommen haben rund 90 Mitarbeitende aus Helmholtz. In der dritten Auflage des Forums wurden zentrale Entwicklungen in Helmholtz rund um den Umgang mit Forschungssoftware aufgegriffen (siehe Tab. 1).

Die genannten Akteure haben bereits im Mai 2021 ein Forum zum Thema „Policies für Forschungssoftware“ veranstaltet.⁴ Damals wurde u. a. in parallelen Sessions die Grundlage für eine Checkliste zum Thema geschaffen, die Ende 2021 veröffentlicht wurde und die Zentren bei der Policyentwicklung zum Thema unterstützt.

In einer zweiten Veranstaltung im April 2022 wurde das Thema breiter behandelt. Das damalige Forum widmete sich über die Policyentwicklung hinaus auch wegweisenden Software-Projekten, Tools und Infrastrukturen zur Entwicklung und Publikation von Forschungssoftware.⁵

Bereits 2017 wurde, aufbauend auf einem Positionspapier,⁶ mit der Entwicklung von Empfehlungen zur Implementierung von Leit- und Richtlinien zum Umgang mit wissenschaftlicher Software an den Helmholtz-Zentren begonnen, die 2019 veröffentlicht wurden.⁷ Eine Muster-Richtlinie⁸ für den nachhaltigen Umgang mit Forschungssoftware sowie eine Checkliste⁹ bieten den Zentren darüber hinaus eine Orientierung bei der Entwicklung solcher Policies.

¹ Siehe hierzu auch: <https://www.helmholtz.de/ueber-uns/unsere-werte/open-science/open-source-in-der-forschung/>

² Siehe: <https://os.helmholtz.de/open-science-in-helmholtz/arbeitskreis-open-science/>

³ <https://hifis.net>

⁴ Siehe: <https://os.helmholtz.de/veranstaltungen/foren/1-forum-forschungssoftware/>

⁵ Dokumentation unter: <https://doi.org/10.48440/os.helmholtz.045>

⁶ <https://doi.org/10.48440/os.helmholtz.042>

⁷ <https://doi.org/10.2312/os.helmholtz.008>

⁸ <https://doi.org/10.2312/os.helmholtz.007>

⁹ <https://doi.org/10.48440/os.helmholtz.031>

HELMHOLTZ

Open Science

In Kooperation mit anderen Wissenschaftsorganisationen in Deutschland wurde 2018 außerdem eine „Handreichung zum Umgang mit Forschungssoftware“¹⁰ sowie 2019 ein Positionspapier¹¹ erarbeitet.

In Zusammenarbeit mit sechs europäischen Forschungsorganisationen hat Helmholtz im G6-Netzwerk im Dezember 2021 die Bedeutung der Forschungssoftware im „G6 statement on Open Science“ als Schlüssel für die digitale Forschung prominent betont.¹²

Mit der Verabschiedung der Helmholtz Open Science Policy im September 2022 durch die Mitgliederversammlung wurde das Thema Forschungssoftware als Teil von Open Science in der Gemeinschaft definiert.¹³ Alle drei Abschnitte der Helmholtz Open Science Policy beschäftigen sich mit dem Thema. Im Abschnitt Monitoring wird festgehalten, dass alle Zentren bis 2025 Policies zum Umgang mit Forschungssoftware festlegen.

Tab. 1: Programm des 3. Helmholtz Open Science Forum: Forschungssoftware, 28.11.2022

Time	Agenda	Speaker
10:00-10:05	Welcome	Heinz Pampel, Helmholtz Open Science Office
10:05-10:25	HIFIS	Uwe Konrad, HZDR
10:25-10:40	Software-Spotlight: SaQC - System for automated Quality Control	David Schäfer, UFZ
10:40-10:55	Software-Spotlight: Citation File Format (CFF)	Stephan Druskat, DLR
10:55-11:15	Research Software Directory (RSD)	Christian Meeßen, GFZ
11:15-11:35	Helmholtz Research Software Award	Uwe Konrad, HZDR
11:35-11:55	HiRSE_PS	Robert Speck, FZJ
11:55-12:30	Further ideas and discussion	Martin Hammitzsch, GFZ

¹⁰ <https://doi.org/10.5281/zenodo.1172970>

¹¹ <https://doi.org/10.5281/zenodo.4301924>

¹² <https://os.helmholtz.de/open-science-in-helmholtz/networking/g6/>

¹³ <https://doi.org/10.48440/os.helmholtz.055>

Vorträge

Update zu HIFIS, Helmholtz Federated IT

Die Arbeit von HIFIS und den HIFIS Software Services,¹⁴ einer Plattform zur nachhaltigen Softwareentwicklung, wurde von Uwe Konrad vom Helmholtz-Zentrum Dresden-Rossendorf (HZDR) vorgestellt.

Ziel von HIFIS ist es, eine nahtlose und leistungsfähige IT-Infrastruktur zu schaffen, die Services und Wissen aus allen Helmholtz-Zentren verbindet. HIFIS bietet eine kollaborative Umgebung mit IT-Diensten, die von überall effizient zugänglich sind. Die HIFIS-Plattform unterstützt auch die Entwicklung von Forschungssoftware.

Dazu organisiert sich HIFIS in drei Kompetenzcluster, die auf verschiedene Zentren verteilt sind:

- ein Cloud Services Cluster – föderierte Plattform für Cloud Services,
- ein Backbone Services Cluster – leistungsstarke, vertrauenswürdige Netzwerkinfrastruktur mit einheitlichen Basisdiensten und
- ein Software Services Cluster – Plattform, Training und Support für qualitativ hochwertige, nachhaltige Softwareentwicklung.

Uwe Konrad berichtete über aktuelle Entwicklungen aus dem HIFIS Umfeld:

- Seit dem letzten Software Forum hat sich die Plattform sehr positiv entwickelt und wird bereits von über 12.000 Usern genutzt, darunter über 2.000 externe User von Partnereinrichtungen. Die Anzahl an Services ist deutlich größer geworden (25 produktive Cloudservices), die Angebote werden in der neu gestalteten HIFIS Webseite präsentiert.¹⁵ Auch die Anzahl an Training- und Supportaktivitäten wurde deutlich gesteigert, u.a. wurden 71 Training Courses and Events für 1.200 Teilnehmer organisiert. Über 3.000 Softwareentwickler arbeiten auf der Helmholtz Codebase zusammen.¹⁶
- Die HIFIS Plattform wurde im September von einem internationalen Gutachtergremium evaluiert. Dabei wurde bestätigt, dass HIFIS hervorragende Fortschritte beim Aufbau erster Strukturen und Dienste in seinen drei Schlüsselbereichen gemacht hat und nun von der Aufbauphase in die Betriebsphase übergehen wird. HIFIS kann ein Vorbild für ähnliche Aktivitäten in anderen Bereichen der Forschung sein. Auch ist HIFIS ein Katalysator für die Vertrauensbildung zwischen allen 18 Zentren. Die Kommission empfahl nachdrücklich, die Plattform weiterzuführen.
- Unter anderem wurde empfohlen, die Reichweite/Unterstützung der Aktivitäten im Software-Cluster auszuweiten und mit den anderen Aktivitäten noch stärker zu verbinden.

¹⁴ <https://hifis.net>

¹⁵ <https://hifis.net>

¹⁶ <https://codebase.helmholtz.cloud>

SaQC – System for automated Quality Control

David Schäfer vom Helmholtz-Zentrum für Umweltforschung – UFZ stellte SaQC – System for automated Quality Control vor.

Umweltsensornetzwerke erzeugen immer größere Mengen an Rohdaten, die in verwertbare Daten umgewandelt werden müssen, um die laufenden Umweltveränderungen zu überwachen und Entscheidungen zu unterstützen. Eine Herausforderung ist die Datenbereitstellung in Echtzeit, die eine Automatisierung der Qualitätskontrolle (QC) mit Hilfe geeigneter Software-Tools erfordert.

SaQC¹⁷ unterstützt Forschende bei der Qualitätssicherung von Daten aus Zahlenreihen. SaQC ist ein Software-Framework für die automatisierte Qualitätskontrolle von Zeitreihendaten. Die Software wird vom Forschungsdatenmanagement-Team des UFZ entwickelt und gepflegt. SaQC bietet vollautomatische Qualitätskontrollpipelines für Umweltsensordaten.¹⁸ Die Software wird auf GitHub entwickelt und liegt aktuell in der dritten Version vor.

Die Software ermöglicht Qualitätssicherungsworkflows nachvollziehbar und reproduzierbar zu machen und so FAIR-Datenströme von hoher Qualität zu fördern. Die Releases werden auch auf Zenodo veröffentlicht.¹⁹

SaQC wird auch auf den Helmholtz Software Spotlights präsentiert.²⁰

Citation File Format (CFF)

Stephan Druskat vom Deutschen Zentrum für Luft- und Raumfahrt (DLR) stellte das Citation File Format (CFF)²¹ vor.

Das CFF ermöglicht es Entwickler:innen von Forschungssoftware, Metadaten in einem Sourcecode-Repository zu hinterlegen, die für die korrekte Zitierung der Software benötigt werden. Durch die Hinterlegung des CFF wird die Zitierung von Forschungssoftware ermöglicht und gefördert. Verschiedene Dienste lesen diese Angaben automatisiert aus. CFF basiert auf den Software Citation Principles.²²

Unterstützt wird CFF durch die Heimatinstitutionen der Entwickler:innen (u. a. DLR, Netherlands eScience Center). Gefördert wurde die Entwicklung von Software Sustainability Institute (UK), campusSOURCE (DE) sowie Code for Science & Society (USA).

CFF wird auch auf den Helmholtz Software Spotlights präsentiert.²³

¹⁷ <https://rdm-software.pages.ufz.de/saqc>

¹⁸ <https://doi.org/10.2139/ssrn.4173698>

¹⁹ <https://doi.org/10.5281/zenodo.6809871>

²⁰ <https://hifis.net/spotlights/saqc>

²¹ <https://citation-file-format.github.io>

²² <https://doi.org/10.7717/peerj-cs.86>

²³ <https://hifis.net/spotlights/cff>

HELMHOLTZ

Open Science

Research Software Directory (RSD)

Christian Meeßen vom Deutschen GeoForschungsZentrum GFZ widmete sich in seinem Vortrag dem Helmholtz Research Software Directory (RSD).

Dieses Verzeichnis, das im Rahmen von HIFIS realisiert wird, setzt auf der Erhebung zu den Helmholtz Software Spotlights auf und macht Forschungssoftware aus Helmholtz auffindbar und sichtbar.

Das Helmholtz RSD baut auf dem RSD des Netherlands eScience Center auf und wurde durch HIFIS auf die Bedürfnisse der Helmholtz-Gemeinschaft angepasst. Die Entwicklung am ursprünglichen RSD wird seitens HIFIS unterstützt.²⁴

Die Einträge im RSD zeigen die Software in ihrem Forschungskontext und es können Bezüge zu teilnehmenden Organisationen, Projekten oder verwandter Software hergestellt werden. Das System extrahiert Informationen aus Software-Repositories (u. a. unter Nutzung des CFF) sowie aus ORCID und Zenodo und präsentiert zugehörige Forschungsdaten, Projekte und textuelle Publikationen. Eine Betaversion ist bereits online und wird in den kommenden Monaten weiter ergänzt.²⁵

Helmholtz Research Software Award

Über die Pläne eines Helmholtz Research Software Awards informierte Uwe Konrad vom HZDR.

Nachdem in den letzten Jahren gute Erfahrungen mit der Beteiligung des Helmholtz Open Science Office an dem campusSOURCE Award 2022 gemacht wurden,²⁶ entstand die Idee, einen Award für nachhaltige und offene Forschungssoftware in Helmholtz zu initiieren.

Aktuell wird der Entwurf des Konzeptes mit interessierten Kreisen und der Helmholtz Geschäftsstelle abgestimmt.

HiRSE_PS

Robert Speck vom Forschungszentrum Jülich widmete sich in seinem Vortrag dem Vorhaben „Helmholtz Platform for Research Software Engineering – Preparatory Study“ (HiRSE_PS).²⁷

Das Projekt wird aus dem Innovationspoolfonds des Forschungsbereichs Information sowie von den vier beteiligten Zentren (Forschungszentrum Jülich, Karlsruher Institut für Technologie, Helmholtz-Zentrum Hereon und Helmholtz-Zentrum Berlin) finanziert. Dabei kooperiert HiRSE_PS u.a. bereits mit HIFIS.

Im Kern von HiRSE_PS steht die Etablierung von RSE-Aktivitäten um strategisch wichtige Codes durch sogenannte Community Software Infrastructure (CSI) Gruppen nachhaltig betreiben zu können. In einer Preparatory Study wird HiRSE_PS die Kernelemente des HiRSE-Konzepts und deren Zusammenspiel in der Praxis über die Förderperiode von zwei Jahren evaluieren. Ein Arbeitspaket

²⁴ Zum Hintergrund: <https://blog.esciencecenter.nl/software-in-the-spotlight-launching-the-research-software-directory-59b6b5a15da5>

²⁵ <https://helmholtz.software>

²⁶ <https://ev.campussource.de/publikationen/csa2022/gewinner.html>

²⁷ <https://www.helmholtz-hirse.de>

HELMHOLTZ

Open Science

beschäftigt sich mit dem Betrieb von dezentralen CSI-Gruppen, ein zweites mit der Etablierung von Cx-Frameworks bei den Jülicher und Karlsruher HPC-Systemen sowie zentraler Beratung und Vernetzung. Ziel der vorbereitenden Studie ist die Weiterentwicklung des Konzepts, das dann auf weitere Bereiche der Gemeinschaft ausgeweitet werden kann.

Abschlussdiskussion

In der von Martin Hammitzsch, Deutsches GeoForschungsZentrum GFZ, moderierten Abschlussdiskussion wurden aktuelle Herausforderungen und aufkommende Themenfelder für den Umgang mit Forschungssoftware in Helmholtz diskutiert. Identifiziert wurden drei Aspekte

- Strukturelle Aspekte: Es gilt weiterhin, RSE-Services für Forschende auszubauen. Mit HIFIS wurde hier ein wichtiger Service geschaffen. Das Zusammenspiel von RSE-Themen auf Ebene der Zentren, der Forschungsbereiche und der gesamten Gemeinschaft gilt es im Dialog mit nationalen und internationalen Entwicklungen weiter zu fördern. Das Helmholtz Forum Forschungssoftware leistet hier bereits einen wichtigen Beitrag, der die Stakeholder zusammenbringt.
- Aspekte der Publikation und der Sichtbarkeit: Mit den Spotlights und dem RSD kann die Sichtbarkeit von Forschungssoftware gefördert werden. Hier ergeben sich auch Anknüpfungspunkte zum Publikationsmanagement in den Zentren und der Arbeit der Bibliotheken bei der Erfassung von Forschungsleistung. Auch wirken Diskussionen um die Entwicklung eines Qualitätsindikators auf die Bearbeitung des Themas. Diese Themen korrespondieren mit den Entwicklungen auf EU-Ebene rund um die Weiterentwicklung der Forschungsbewertung.
- Policy-Aspekte: Mit der Helmholtz Open Science Policy gewinnt die Entwicklung von Forschungssoftware-Policies an Relevanz. Zur Unterstützung der Zentren wurden bereits Empfehlungen, eine Muster-Richtlinie und eine Checkliste erarbeitet.²⁸ In der Zukunft gilt es, den Dialog der Zentren bei der Implementierung der Forschungssoftware-Policies weiter zu fördern.

²⁸ Siehe Abschnitt 1.

HELMHOLTZ

Open Science

Anhang

- Einleitung - Heinz Pampel
- SaQC - System for automated Quality Control - David Schäfer
- Citation File Format (CFF) - Stephan Druskat
- Research Software Directory (RSD) - Christian Meeßen²⁹
- HIRSE_PS - Markus Diesmann, Achim Streit, Robert Speck (Vortragender), Rene Caspart

²⁹ Siehe auch: <https://doi.org/10.5281/zenodo.7360385>

HELMHOLTZ

Open Science

Helmholtz Open Science Forum Forschungssoftware

November 24, 2022, 10:00 a.m. to 12:30 p.m.

Welcome

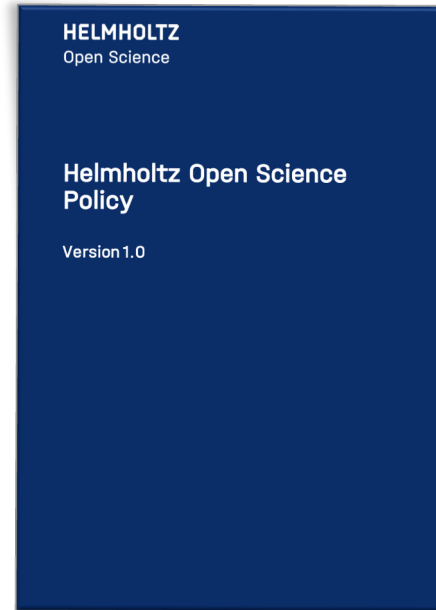
- The Helmholtz Forum Research Software is jointly organized by the Task Group Research Software of the WG Open Science and the HIFIS Software Cluster.
- The event will be hosted by the Helmholtz Open Science Office.
- The virtual forum is open to all Helmholtz employees.
- A first Helmholtz Forum Research Software event was held in May 2021 and a second in April 2022.

Open Research Software

- The topic of Open Research Software continues to gain importance; while open science calls for **reuse** and **reproducibility** of scientific results, in many research areas, this can only be ensured if **source code** is made openly accessible in addition to the corresponding research data. At the Helmholtz Centers, various research software initiatives are currently being implemented.
- Works:
 - 2017: Position Paper [Access to and re-use of research software](#)
 - 2019: [Recommendations](#) for the Implementation of Guidelines and Policies on Research Software Management at the Helmholtz Centers
 - 2019: [Model Policy on Sustainable Software at the Helmholtz Centers](#)
 - 2021: [Checklist for the Implementation of Software Policies at the Helmholtz Centers](#)
 - 2022: Contribution to the campusSOURCE Award 2022 for research software
- In 2022, the Helmholtz Association's General Assembly adopted the **Helmholtz Open Science Policy**.
- Cooperation (selection):
 - Priority Initiative „[Digital Information](#)“ of the Alliance of Science Organizations in Germany, [de-RSE](#) Gesellschaft für Forschungssoftware, [G6 Network](#)

Helmholtz Open Science Policy

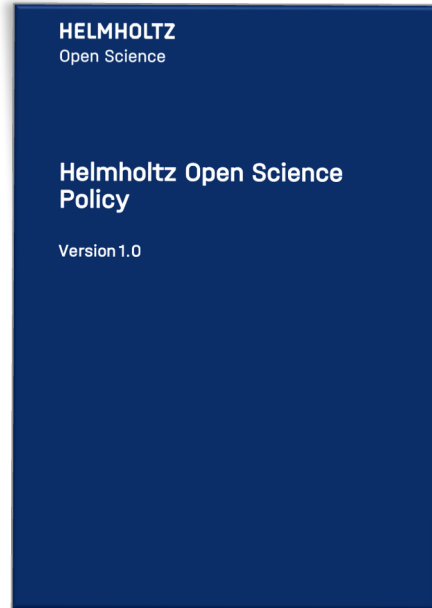
- The Helmholtz Association's General Assembly adopted an Open Science Policy in September 2022.
- The policy stipulates that scientific articles, research data and research software are to be published openly. Open Science thus becomes the standard of publication practice.
- The Helmholtz Open Science Policy is organized into three sections:
 - In the section "A. Strategic Positioning," the Helmholtz Association formulates a commitment to open science in accordance with the principle "as open as possible and as closed as necessary."
 - In the section "B. Monitoring," Helmholtz makes a commitment to open science on the basis of the three current focus areas through specific and verifiable objectives.
 - In the section "C. Implementation and Common Requirements," in line with the funding policy of the European Commission, open science practices are formulated for implementation purposes.
- Further Information:
 - <https://os.helmholtz.de/en/open-science-in-helmholtz/open-science-policy>



Helmholtz Open Science Policy

Open Research Software

- „3.1 The employees shall ensure that
 - wherever possible, the program code that is necessary for reusing and/or validating the published data is made openly accessible in a repository.
- 3.2 For implementation purposes, the Centers and their information infrastructure facilities shall ensure that
 - researchers are supported in publishing research software;
 - the metadata of the deposited software/the deposited code are published under a Creative Commons Public Domain Dedication License (CC0) or an equivalent license (insofar as legitimate interests or constraints are respected) and in accordance with the FAIR Principles (especially machine-readability), and that they include at least information about the following: program code (description, date of deposit, version, authors, repository); name, acronym, and number of the funding project; licensing conditions; persistent identifiers and, if possible, organizations and funding. If applicable, the metadata must also include persistent identifiers for related publications and other research results.“



Organisational Notes

- We would like to ask you to mute your microphone during the event so that we can enable good sound quality.
- Please signal word contributions with an exclamation mark in the chat, or use the chat function for written contributions. We will take these into consideration whenever possible.
- Slides will be made available afterwards.
- The event will be held in German. Questions and discussions are of course also possible in English.

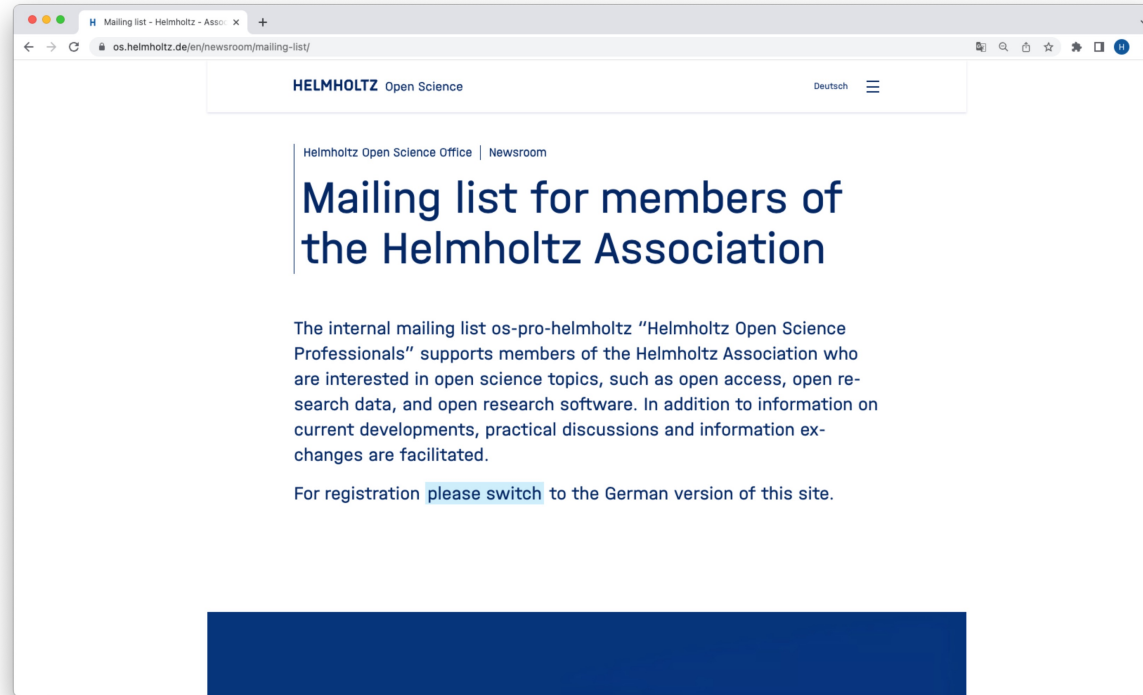
Program

Time	Agenda	Speaker
10:00-10:05	Welcome	Heinz Pampel, Helmholtz Open Science Office
10:05-10:25	HIFIS	Uwe Konrad, HZDR
10:25-10:40	Software-Spotlight: SaQC - System for automated Quality Control	David Schäfer, UFZ
10:40-10:55	Software-Spotlight: Citation File Format (CFF)	Stephan Druskat, DLR
10:55-11:15	Research Software Directory (RSD)	Christian Meeßen, GFZ
11:15-11:35	Helmholtz Research Software Award	Uwe Konrad, HZDR
11:35-11:55	HiRSE_PS	Robert Speck, FZJ
11:55-12:30	Further ideas and discussion	Martin Hammitzsch, GFZ

Upcoming Events

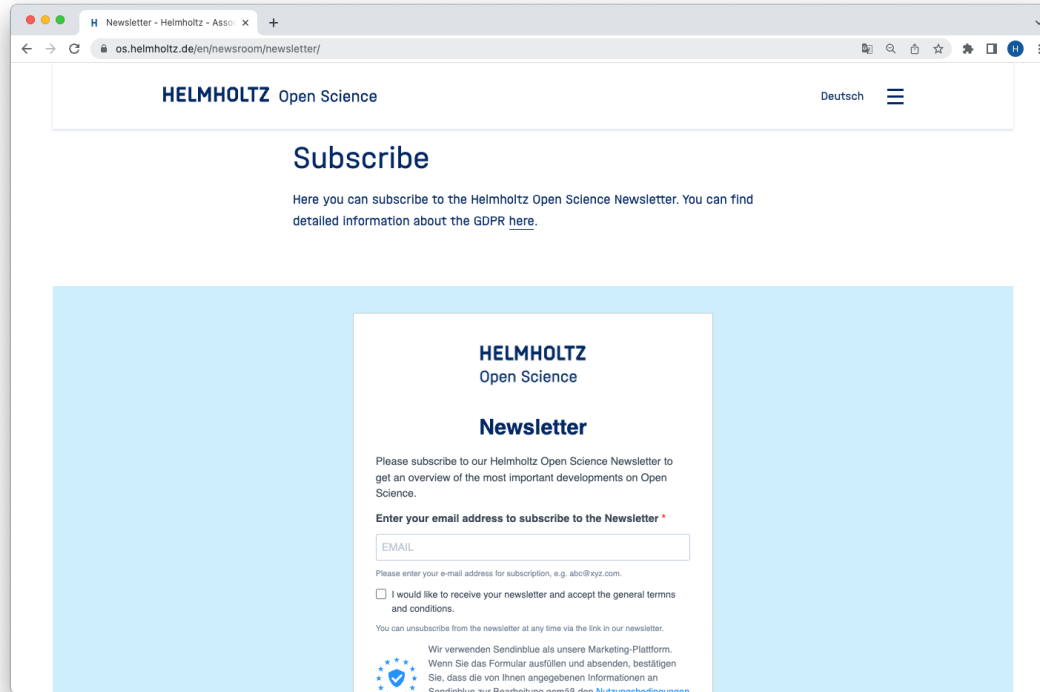
- 24.11.2022: ORCID-DE-Veranstaltung „ORCID-Implementierung in das eigene Identity-Management-System“
 - <https://www.orcid-de.org/orcid-in-idm-austauschtreffen/>
- 17.01.2023: 66th Online Seminar on Open Science Policy
- 14.02. 2023: 2nd Reproducibility Workshop (helmholtz-internal)
- 13.-17.02.2023: RDA Deutschland Tagung
 - <https://www.rda-deutschland.de/news/save-the-date-rda-de-tagung-2023>

Mailing List



[Link](#)

Newsletter



The screenshot shows a web browser window with the URL `os.helmholtz.de/en/newsroom/newsletter/`. The page header includes the logo "HELMHOLTZ Open Science" and a language selector "Deutsch" with a menu icon. The main heading is "Subscribe", followed by the text: "Here you can subscribe to the Helmholtz Open Science Newsletter. You can find detailed information about the GDPR [here](#)."

The central part of the page features a light blue background with a white box containing the following content:

- HELMHOLTZ Open Science**
- Newsletter**
- Please subscribe to our Helmholtz Open Science Newsletter to get an overview of the most important developments on Open Science.
- Enter your email address to subscribe to the Newsletter ***
-
- Please enter your e-mail address for subscription, e.g. abc@xyz.com.
- ☐ I would like to receive your newsletter and accept the general terms and conditions.
- You can unsubscribe from the newsletter at any time via the link in our newsletter.
- Wir verwenden Sendinblue als unsere Marketing-Plattform. Wenn Sie das Formular ausfüllen und absenden, bestätigen Sie, dass die von Ihnen angegebenen Informationen an Sendinblue zur Bearbeitung gemäß den [Nutzungsbedingungen](#)

[Link](#)

Keep in touch

- Email - open-science@helmholtz.de
- Twitter - [@helmholtz_os](https://twitter.com/helmholtz_os)
- Website - <https://os.helmholtz.de>
- Mailing list for members of Helmholtz - [Helmholtz Open Science Professionals](#)
- [Helmholtz Open Science Newsletter](#)



All texts in this presentation, except citations, are licensed under Attribution 4.0 International (CC BY 4.0):

<https://creativecommons.org/licenses/by/4.0/deed.de>

SaQC – System for automated Quality Control

David Schäfer, Bert Palm, Peter Lünenschloß, Lennart Schmidt,
Florian Gransee, Jan Bumberger

Leipzig, 2024-11-22

- **Timeseries:** An ordered collection of observations of random variable, comprised of several components: trend, seasonal, cyclical, irregular
- **Quality Control:** (automated) detection and annotation of anomalies in timeseries
- **Anomaly:** An observations, that deviates from other observations in a way, that is is not explainable by the timeseries' components
- **Quality Label / Flag:** An annotation marking a single observation as an anomaly

- Python module and a standalone command line application
- Integrated framework to quality control timeseries data
- Collection of methods/algorithms to analyse, process and annotate timeseries data



Why do we need SaQC?

- Bode hydrological observatory
 - 912 sensors, 42 Loggers
 - > 185. mio data values
- Forest ecosystem observatory Hohes Holz
 - 2114 sensors, 87 loggers
 - > 775 mio data values



- Python API
 - Python code
 - Feature complete
 - Integrate with the large numerical python eco system

```
from saqc import SaQC

qc = SaQC(data)

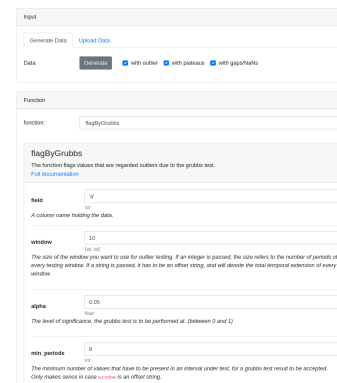
qc = (qc
      .shift("SM2", freq="15Min")
      .flagMissing("SM(1|2)+", regex=True)
      .flagRange("SM1", min=10, max=60)
      .flagRange("SM2", min=10, max=40)
      .flagMAD("SM2", window="30d", z=3.5))
```

- Configuration System
 - Plain text
 - Track changes using version control
 - Separate processes from qc configuration

```
varname      ; test
#----- ; -----
SM2          ; shift(freq="15Min")
'SM(1 | 2) + ' ; flagMissing()
SM1          ; flagRange(min=10, max=60)
SM2          ; flagRange(min=10, max=40)
SM2          ; flagMAD(window="30d", z=3.5)
```

- Dashboard
 - Webservice with a graphical user interface
 - Early stages, not yet feature complete
 - <https://webapp.ufz.de/saqc-config-app/>

SaqQC Configuration App



Input

Generate Data:

Date: ☒ with outlier ☒ with platform ☒ with gaps/NaNs

Function:

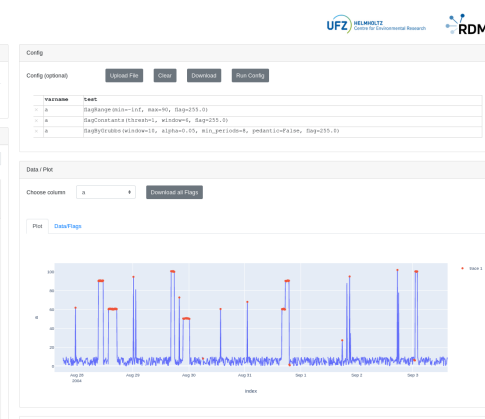
flagByGrubbs
This function flags values that are regarded outliers due to the grubbs test.
[Full documentation](#)

bold: ☒
A column name holding the data.

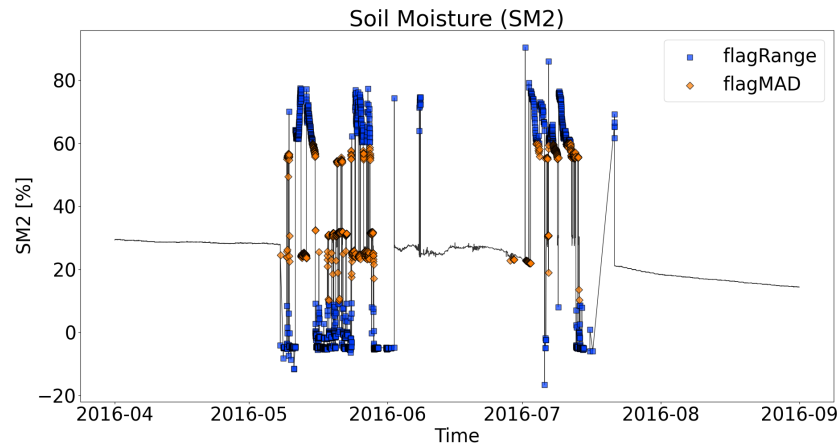
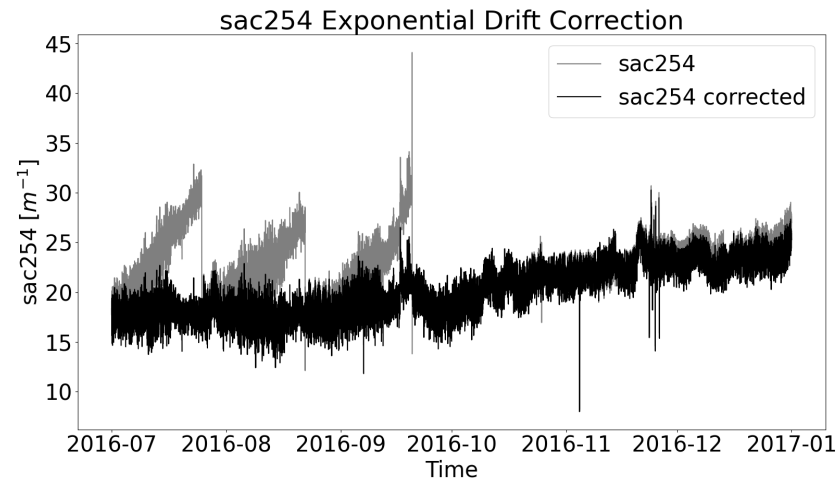
window:
The size of the window you want to use for outlier testing. If an integer is passed, the size refers to the number of periods of every testing variable. If a string is passed, it has to be an offset string, and will denote the total temporal extension of every variable.

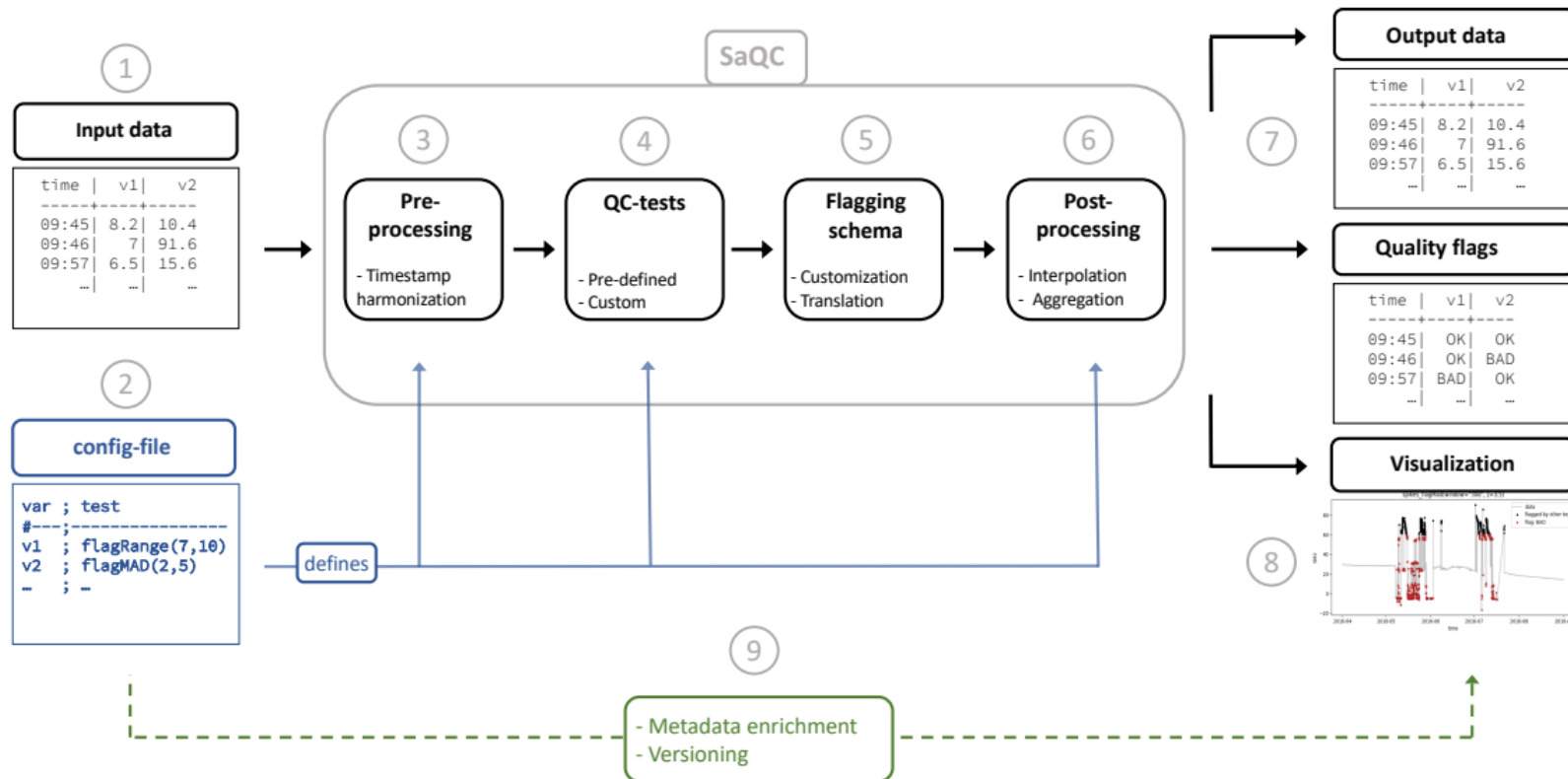
alpha:
The level of significance, the grubbs test is to be performed at. (between 0 and 1)

min_periods:
The minimum number of values that have to be present in an interval under test, for a grubbs test result to be accepted. Only makes sense in case of <string> as an offset string.



- Processing
 - Resampling, Interpolation
 - Smoothing, data products
- Quality Control
 - Basic: constants-, breaks-, outlier-detection
 - Advanced: noise-, drift-, pattern-detection
 - Generic: formulate cross variable conditions





Configuration

```
qc = SaQC(data)
qc = (qc
      .flagRange("SM2", min=10, max=60)
      .flagConstants(field="SM2", window="4h", thresh=0.001)
      .flagMAD(field="SM2", window="60min", min_residuals=1))
```

SaQC

	0	1	2
Date			
2016-07-07 22:06:32	255.0	NaN	NaN
2016-07-07 22:16:34	255.0	NaN	NaN
2016-07-07 22:26:37	255.0	NaN	NaN
2016-07-07 22:36:39	NaN	255.0	NaN
2016-07-07 22:46:41	NaN	255.0	NaN
2016-07-07 22:56:44	255.0	NaN	NaN

```
[{'func': 'flagRange',
  'args': (),
  'kwargs': {'min': 10, 'max': 60, 'dfilter': -inf, 'field': 'SM2'}},
 {'func': 'flagConstants',
  'args': (),
  'kwargs': {'window': '120min',
             'thresh': 0.1,
             'dfilter': -inf,
             'field': 'SM2'}},
 {'func': 'flagMAD',
  'args': (),
  'kwargs': {'window': '120min', 'dfilter': -inf, 'field': 'SM2'}}]
```

Quality Labels

	quality_flag	quality_comment
Date		
2016-07-07 22:06:32	BAD	flagRange
2016-07-07 22:16:34	BAD	flagRange
2016-07-07 22:26:37	BAD	flagRange
2016-07-07 22:36:39	BAD	flagConstants
2016-07-07 22:46:41	BAD	flagConstants
2016-07-07 22:56:44	BAD	flagRange

- Configurator App: <https://webapp.ufz.de/saqc-config-app/>
- Online Documentation: <https://rdm-software.pages.ufz.de/saqc/>
- Learning Material: <https://git.ufz.de/rdm/saqccourse>
- Support: saqc-support@ufz.de
- Code Repository: <https://git.ufz.de/rdm-software/saqc>
- Publications:

Schmidt et al (in review): System for Automated Quality Control (SaQC) to Enable Traceable and Reproducible Data Streams in Environmental Science. Available at SSRN: <https://ssrn.com/abstract=4173698> or <http://dx.doi.org/10.2139/ssrn.4173698>

Thank You!

Questions?



David Schäfer



david.schaefer@ufz.de



<https://www.intranet.ufz.de/rdm>



Mattermost:

<https://mm.ufz.de/ufz/channels/rdm--wkdv-news>



SOFTWARE-SPOTLIGHT: CITATION FILE FORMAT (CFF)

Stephan Druskat

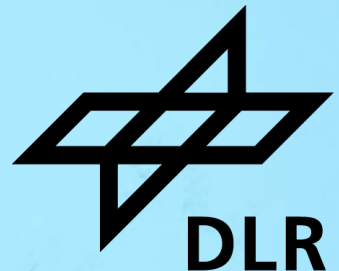
German Aerospace Center (DLR), Institute for Software Technology

DOI [10.5281/zenodo.7347066](https://doi.org/10.5281/zenodo.7347066)

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



Software
Sustainability
Institute

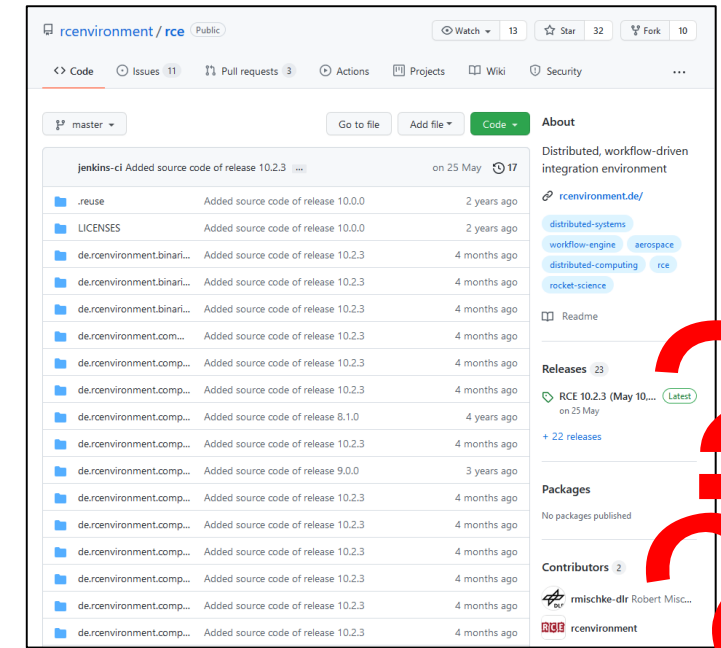
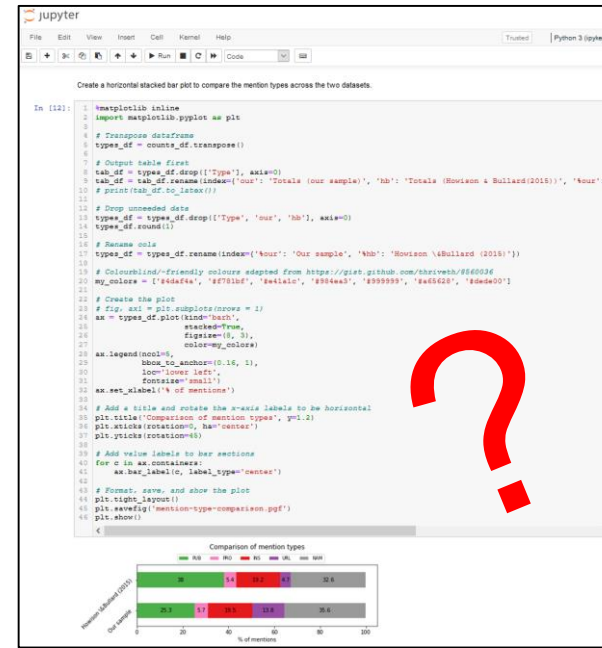
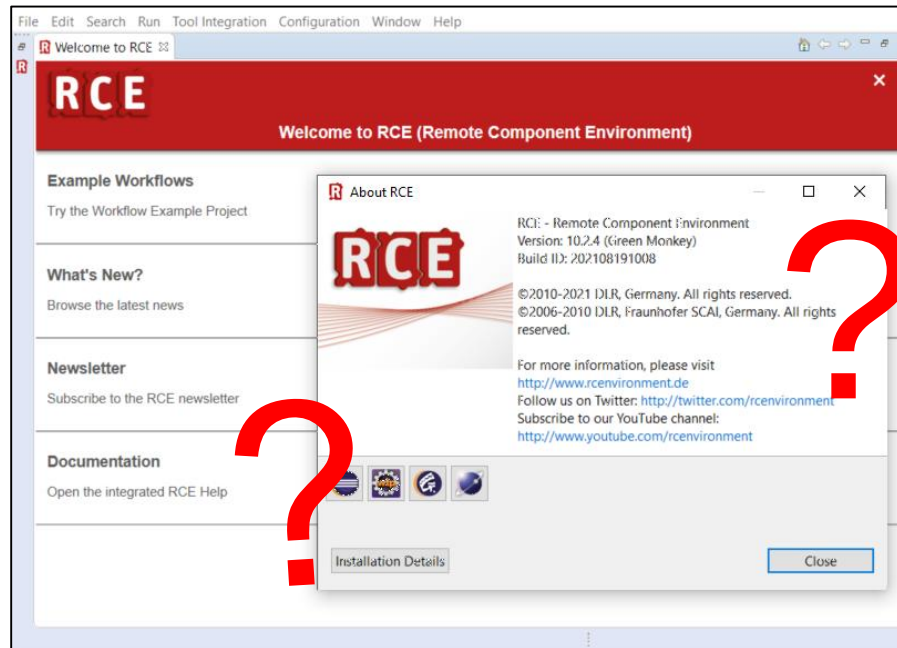


Why make research software citable?



- **Importance:** research software as part of the scholarly record
- **Credit:** it's fair, for RSEs
- **Findability, accessibility, reusability:** it's FAIR, for research software
- **Reproducibility**

The problem: metadata Software has no title page



METADATA: Name? Authors? Version? Publisher? Publication date?

Providing citation metadata with the Citation File Format



- authoritative
- controllable
- "principled":

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

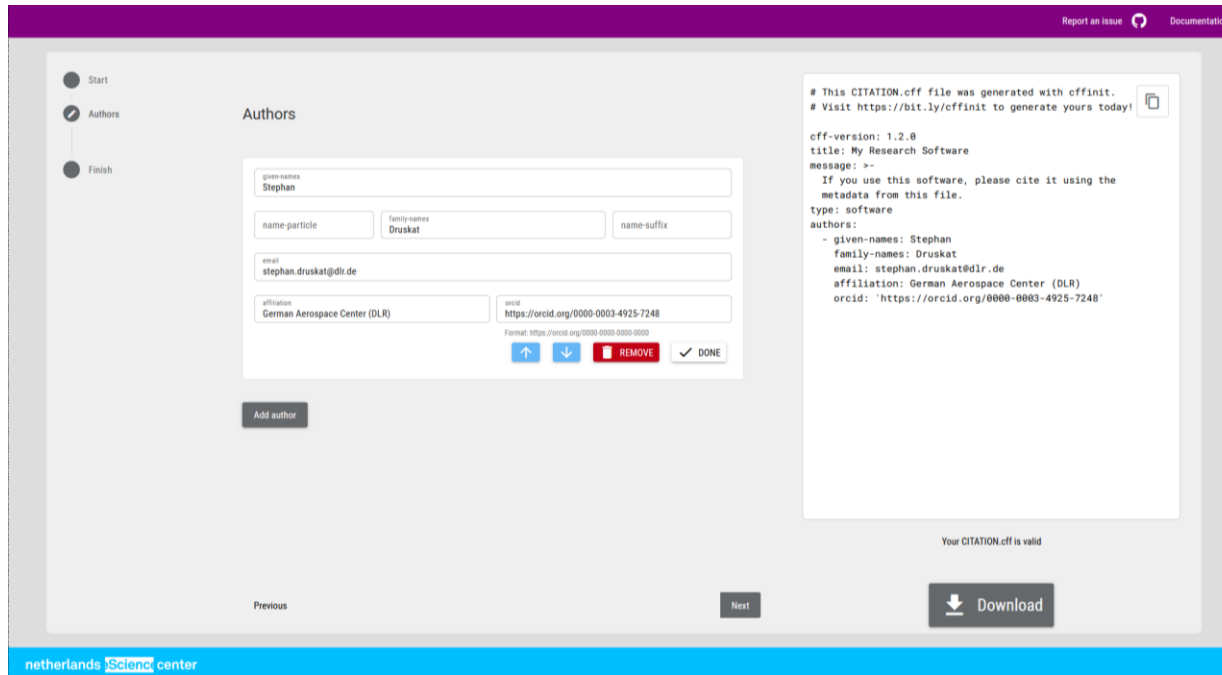
```
cff-version: 1.2.0
message: If you use this software, please cite it using these metadata.
title: My Research Software
abstract: This is my awesome research software. It does many things.
authors:
  - family-names: Druskat
    given-names: Stephan
    orcid: "https://orcid.org/0000-0003-4925-7248"
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"
```

+ references, preferred citation, dataset support (experimental)

METADATA: In a **CITATION.cff** file in the source code repository

Citation File Format: more than "just" a format

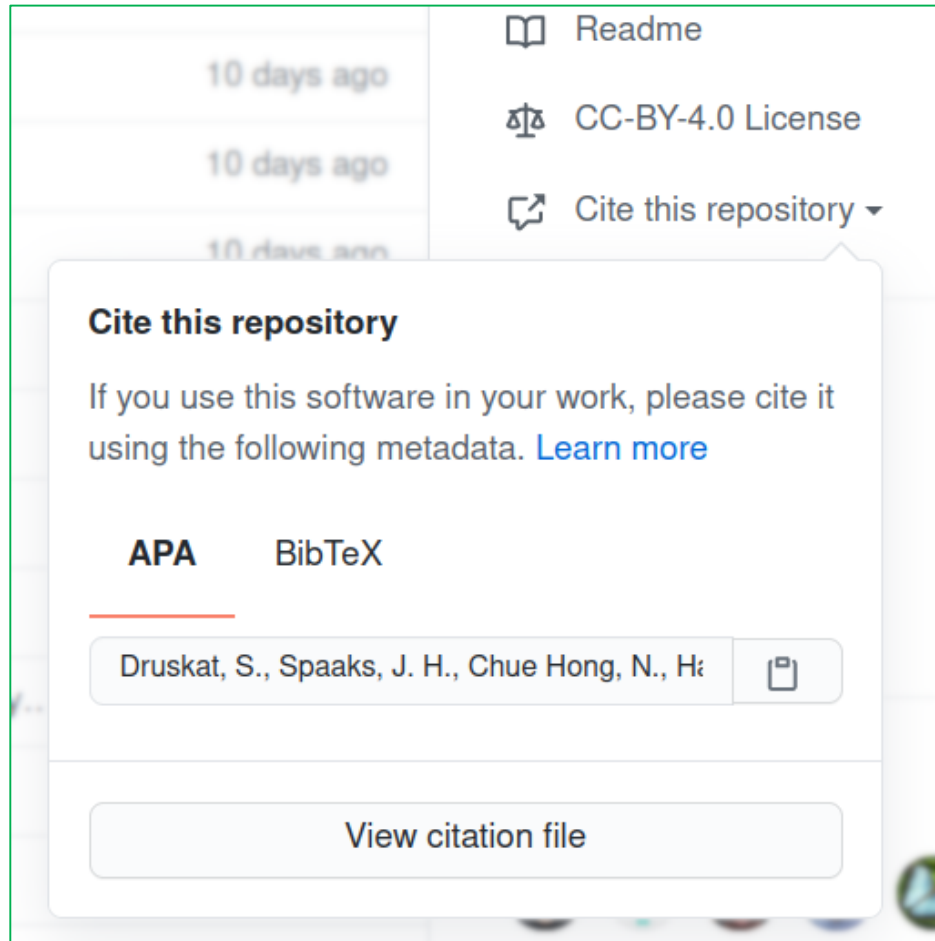
Create, edit & update, validate, convert

A screenshot of the cffinit web tool interface. On the left, a sidebar shows 'Start', 'Authors', and 'Finish' steps. The 'Authors' section is active, displaying a form with fields for 'given-names' (Stephan), 'name-particle', 'family-names' (Druskat), 'name-suffix', 'email' (stephan.druskat@dlr.de), 'affiliation' (German Aerospace Center (DLR)), and 'orcid' (https://orcid.org/0000-0003-4925-7248). Below the form is an 'Add author' button. On the right, a text area shows the generated CFF file content, including the title 'My Research Software', a message, and the author information. At the bottom, there are 'Previous', 'Next', and 'Download' buttons. The footer of the interface reads 'netherlands science center'.

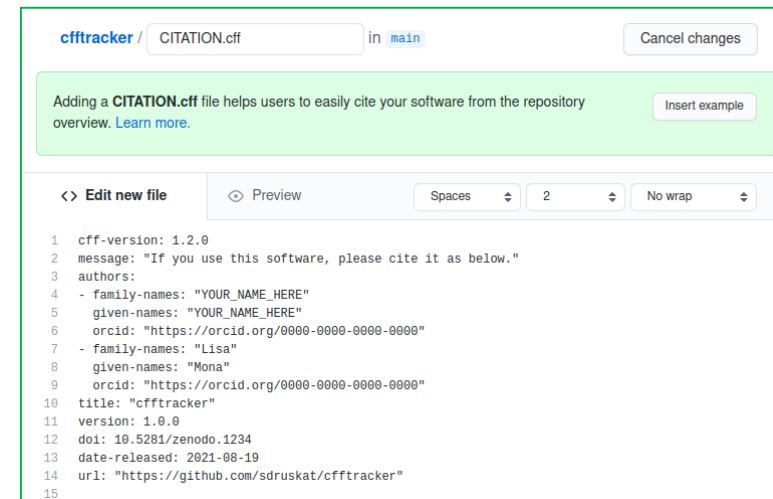
cffinit: web tool for CFF file creation - <https://bit.ly/cffinit>

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

Citation File Format ❤️ GitHub



- Detects **CITATION.cff** files
- Renders citation metadata (APA, BibTeX)
- Points to **CITATION.cff** instance
- [Documents](#) and supports creation



Citation File Format 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

```
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

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-1

- CITATION.cff
- README.md

Versions

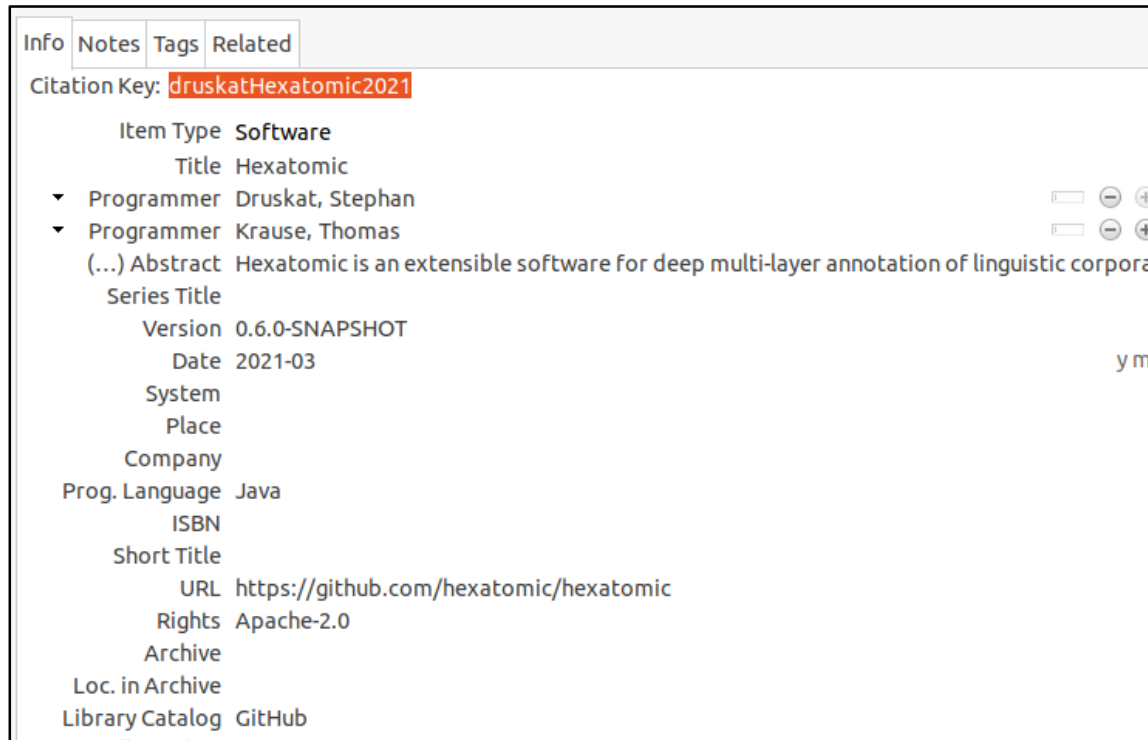
Version 0.2.0	Mar 16, 2022
10.5072/zenodo.1035737	
Version 0.1.0	Mar 16, 2022
10.5072/zenodo.1035711	

Cite all versions? You can cite all versions by using the DOI [10.5072/zenodo.1035710](https://doi.org/10.5072/zenodo.1035710). This DOI represents all versions, and will always resolve to the latest one. [Read more.](#)

Citation File Format reference managers



Import in Zotero from GitHub via browser plugin



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>

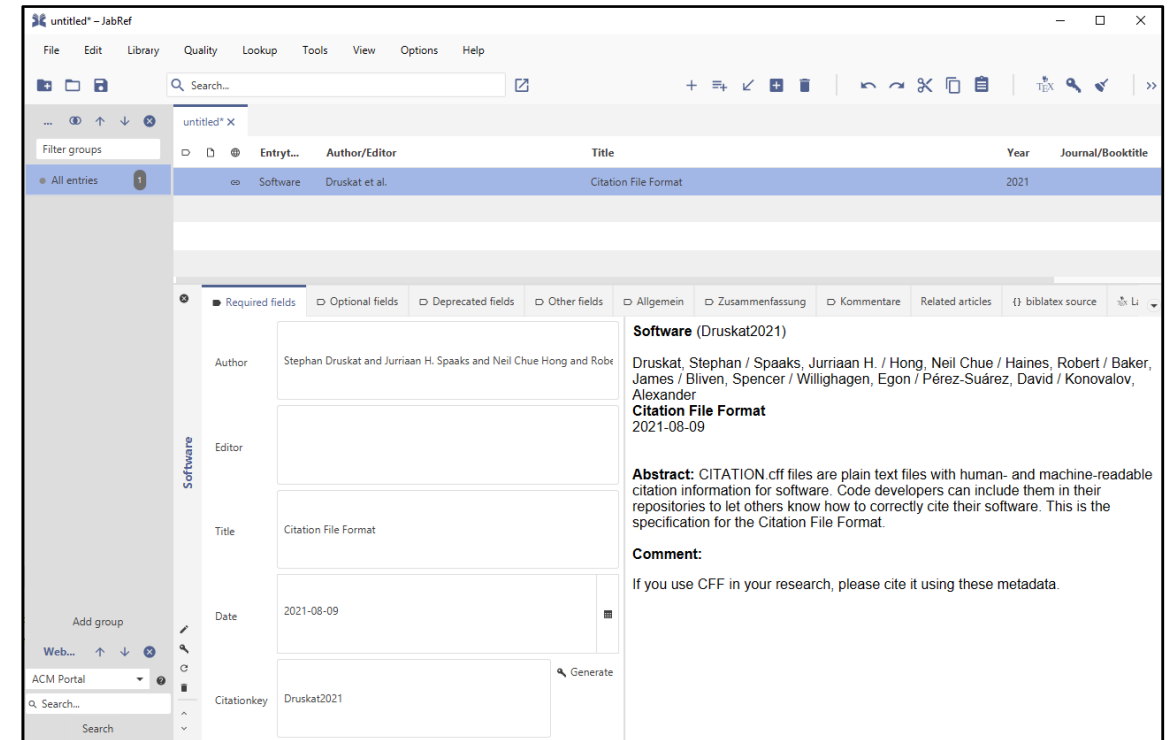
Rights Apache-2.0

Archive

Loc. in Archive

Library Catalog GitHub

Import in JabRef from **CITATION.cff** file



untitled* - JabRef

File Edit Library Quality Lookup Tools View Options Help

Search...

Filter groups

untitled* X

Entry... Author/Editor Title Year Journal/Booktitle

All entries 1

Software Druskat et al. Citation File Format 2021

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

Software (Druskat2021)

Author Stephan Druskat and Jurriaan H. Spaaks and Neil Chue Hong and Robert / Baker, James / Bliven, Spencer / Willighagen, Egon / Pérez-Suárez, David / Kononov, Alexander

Editor

Title Citation File Format

Date 2021-08-09

Citationkey Druskat2021

Generate

Citation File Format
2021-08-09

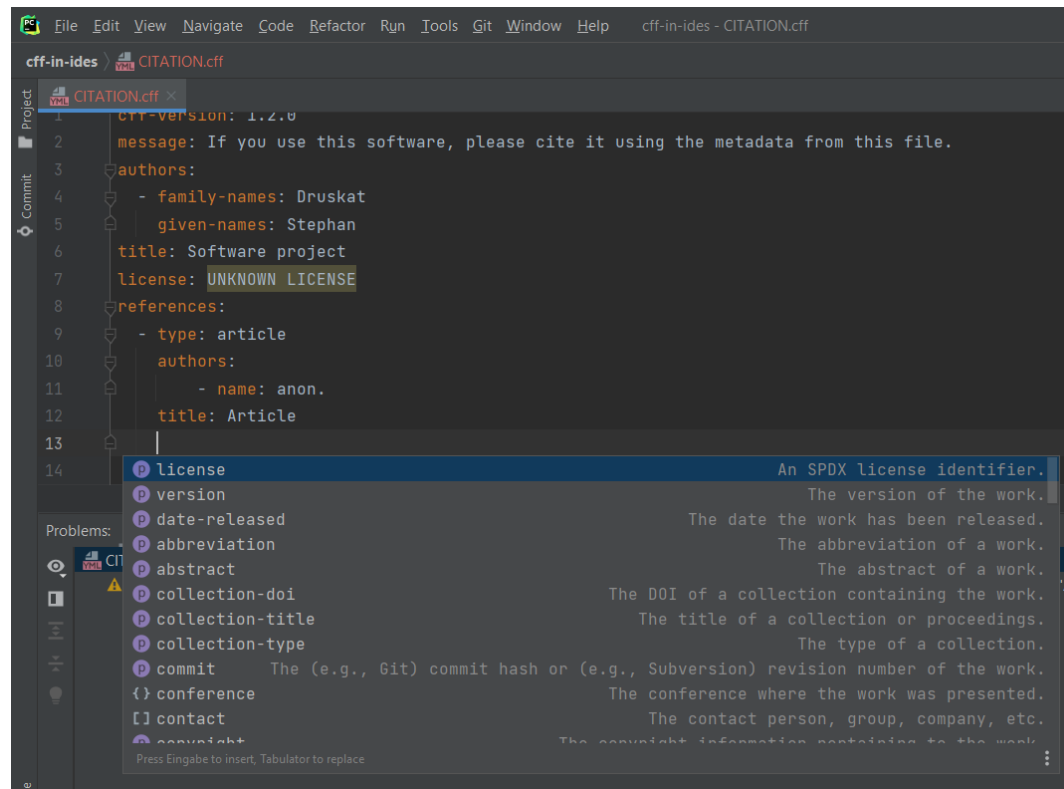
Abstract: CITATION.cff files are plain text files with human- and machine-readable citation information for software. Code developers can include them in their repositories to let others know how to correctly cite their software. This is the specification for the Citation File Format.

Comment:
If you use CFF in your research, please cite it using these metadata.

Citation File Format IDEs



- Auto-complete and validation via JSON Schema / schemastore.org

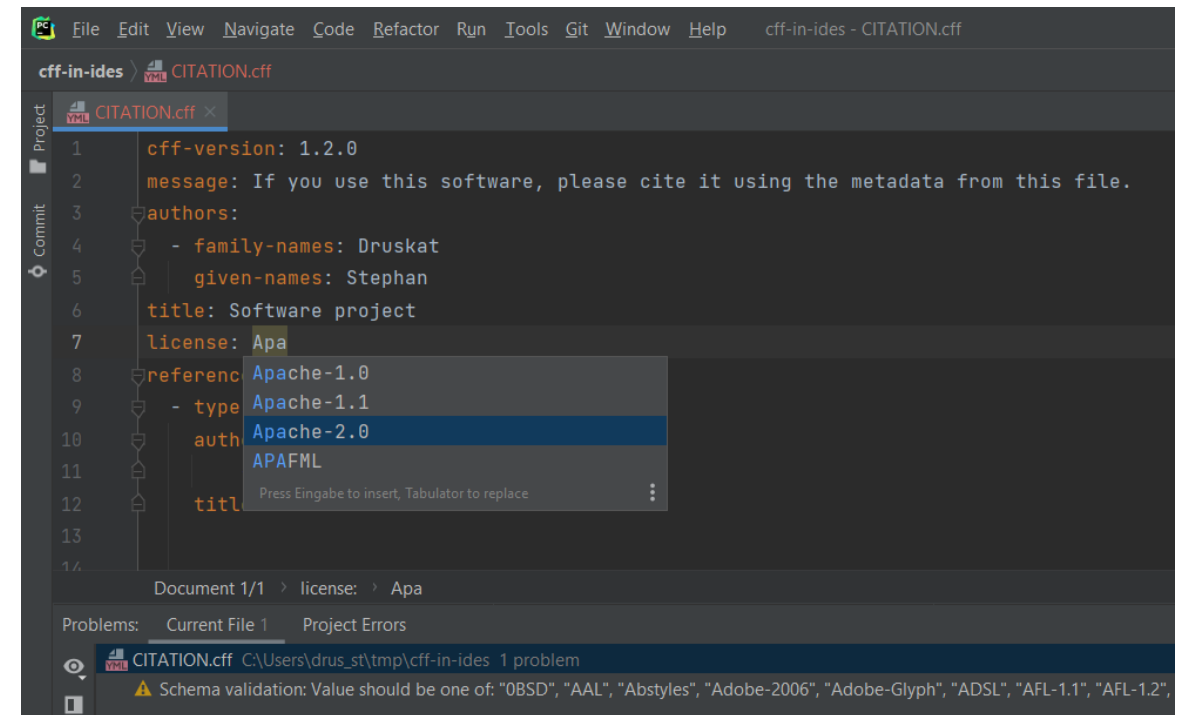


The screenshot shows an IDE window with a file named 'CITATION.cff'. The file content is as follows:

```
1 cff-version: 1.2.0
2 message: If you use this software, please cite it using the metadata from this file.
3 authors:
4   - family-names: Druskat
5     given-names: Stephan
6   title: Software project
7   license: UNKNOWN LICENSE
8 references:
9   - type: article
10     authors:
11       - name: anon.
12     title: Article
```

A dropdown menu is open for the 'license' field on line 7, showing a list of available SPDX license identifiers. The 'license' field is currently set to 'UNKNOWN LICENSE'.

Field	Description
license	An SPDX license identifier.
version	The version of the work.
date-released	The date the work has been released.
abbreviation	The abbreviation of a work.
abstract	The abstract of a work.
collection-doi	The DOI of a collection containing the work.
collection-title	The title of a collection or proceedings.
collection-type	The type of a collection.
commit	The (e.g., Git) commit hash or (e.g., Subversion) revision number of the work.
conference	The conference where the work was presented.
contact	The contact person, group, company, etc.
copyright	The copyright information pertaining to the work.



The screenshot shows the same IDE window with the 'CITATION.cff' file. The 'license' field is now set to 'Apa'. A dropdown menu is open for the 'license' field, showing a list of available SPDX license identifiers. The 'license' field is currently set to 'Apa'.

Document 1/1 > license: > Apa

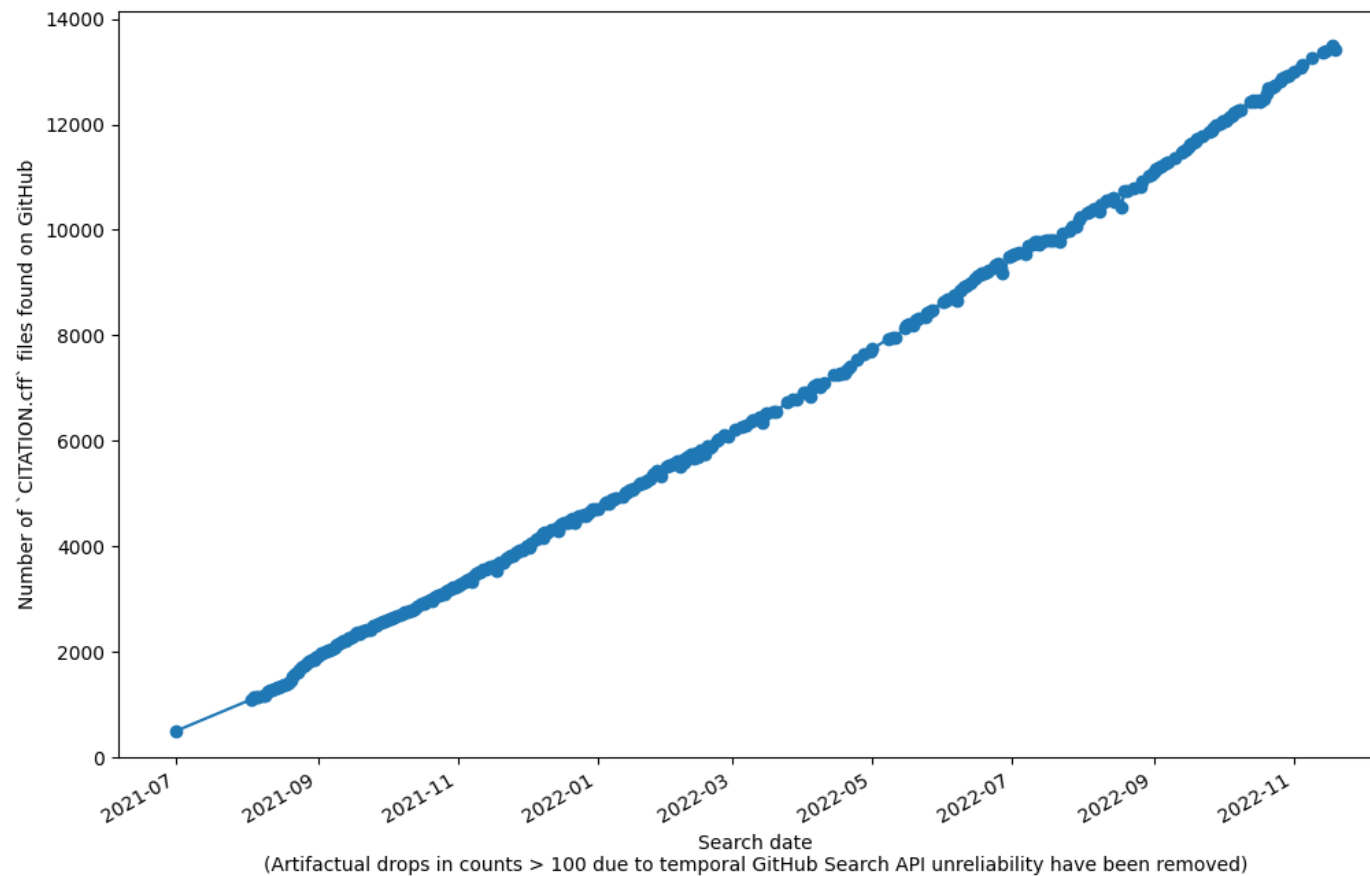
Problems: Current File 1 Project Errors

CITATION.cff C:\Users\drus_st\tmp\cff-in-ides 1 problem

Schema validation: Value should be one of: "0BSD", "AAL", "Abstyles", "Adobe-2006", "Adobe-Glyph", "ADSL", "AFL-1.1", "AFL-1.2",

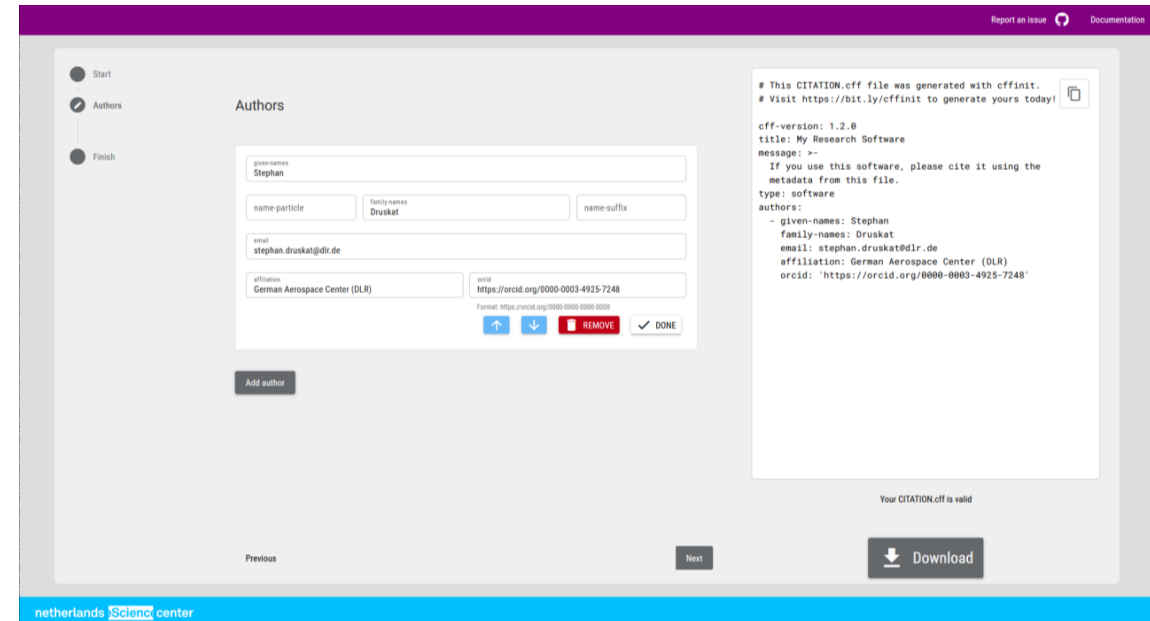
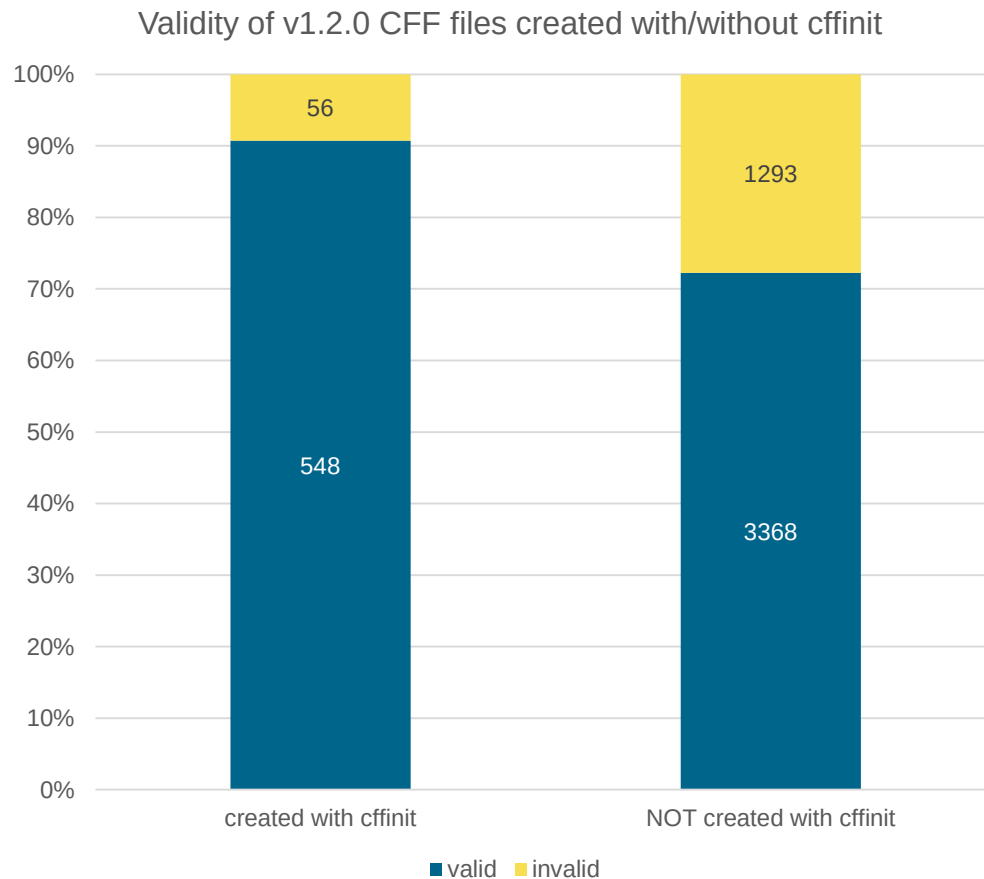
How many CFF files are there?

- On GitHub, as of yesterday (if the GitHub Search API is right): **13,419**



What do people's CFF files look like?

Does the CFF initialization website improve validity?



The screenshot shows the 'Authors' form on the cffinit website. The form includes fields for given-names, family-names, name-particle, name-suffix, email, affiliation, and orcid. Below the form are buttons for 'Add author', 'Previous', 'Next', and 'Download'. To the right, a preview of the generated CFF file is shown, including a message about the file's origin and the cff-version: 1.2.0.

```
# 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'
```

Take home message:

Use <https://bit.ly/cffinit> to create a CFF file for your software

Thanks!



Thanks to all contributors to the CFF project!

Thanks to the Software Sustainability Institute, Code for Science & Society, CampusSOURCE for funding!

Special thanks to Jurriaan H. Spaaks and Rob Haines!

Get in touch ...

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

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

(... with me)

stephan.druskat@dlr.de

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

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

Aktueller Stand des Helmholtz Research Software Directory

Christian Meeßen¹, Martin Hammitzsch¹, Uwe Konrad²

3. Helmholtz Open Science Forum: Forschungssoftware

Online, 24. November 2022

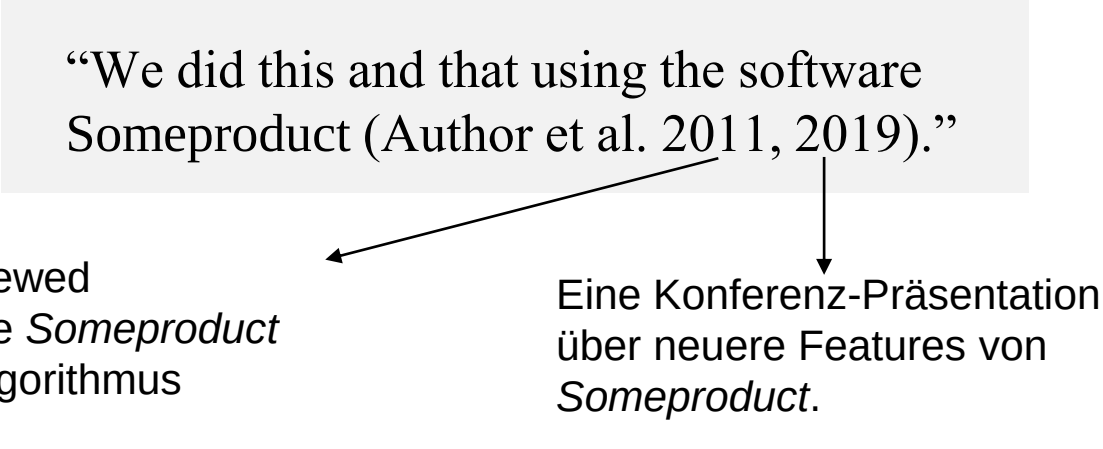
DOI [10.5281/zenodo.7360386](https://doi.org/10.5281/zenodo.7360386)

1) Helmholtz Zentrum Potsdam, Deutsches GeoForschungsZentrum Potsdam GFZ

2) Helmholtz-Zentrum Dresden-Rossendorf (HZDR)

Forschungssoftware in der Literatur

“We did this and that using the software
Someproduct (Author et al. 2011, 2019).”



```
graph TD; A["“We did this and that using the software  
Someproduct (Author et al. 2011, 2019).”"] --> B["Eine peer-reviewed  
Publikation, die Someproduct  
und dessen Algorithmus  
vorstellt."]; A --> C["Eine Konferenz-Präsentation  
über neuere Features von  
Someproduct."];
```

Eine peer-reviewed
Publikation, die *Someproduct*
und dessen Algorithmus
vorstellt.

Eine Konferenz-Präsentation
über neuere Features von
Someproduct.

FAIRe Forschungssoftware heute

Mehr Verbesserungsbedarf

- Korrekte Lizenzierung
- Dependency pinning
- Auffindbarkeit außerhalb der Forschungsliteratur
- Wertschätzung im Wissenschaftsbetrieb
- Impact measurement
- Uneinheitliche Metadaten bei DOI Providern



Lösungsansatz

Research Software Directory

The screenshot shows the Research Software Directory website. At the top, there is a navigation bar with 'Research Software Directory', a search bar, and links for 'Software', 'Projects', 'Organisations', and 'Sign in'. Below the navigation bar, the Helmholtz logo is displayed with the tagline 'Research for grand challenges.' To the right, a large banner reads 'Promote and Discover Research Software' with the subtitle 'Because software matters' and a 'Browse software' button. Below the banner, the 'Software Spotlights' section is visible, featuring a 'MassBank Record: UA002903'. The record details include the title 'Atrazine; APCI-ITFT; MS2; CE: 35%; R=30000; [M+H]+', a mass spectrum plot, and a table of metadata including accession number, record title, date, authors, license, copyright, and publication information. The MassBank logo and description are also present.

Research Software Directory

Search or jump to...

Software Projects Organisations Sign in

HELMHOLTZ
Research for grand challenges.

Promote and Discover
Research Software
Because software matters

Browse software

Software Spotlights
The latest outstanding software product developed in Helmholtz.

MassBank Search Contents Download Accession **Go** More ▼

MassBank Record: UA002903

Atrazine; APCI-ITFT; MS2; CE: 35%; R=30000; [M+H]⁺

Mass Spectrum

Abundance

m/z

MassBank

MassBank is an open source mass spectral library for the identification of small chemical molecules of metabolomics, exposomics and environmental relevance.

- Ursprünglich entwickelt am Netherlands eScience center
- Unterstützt durch HIFIS Software
- Komplette Neuentwicklung 2022
- Software Katalog für RSEs und WissenschaftlerInnen
- Netherlands eScience center:
<https://research-software-directory.org>
- Helmholtz pilot
<https://helmholtz.software>

Demo

RSD

Software

Projects

Organisations

Sign in

Search or jump to...

HELMHOLTZ

Research for grand challenges.

Promote and Discover
Research Software

Because software matters

Browse software

Software Spotlights

The latest outstanding software product developed in Helmholtz.

Load data

Peak detection

Index this frame

Index all frames

Determine unit cell

Indexing ambiguity

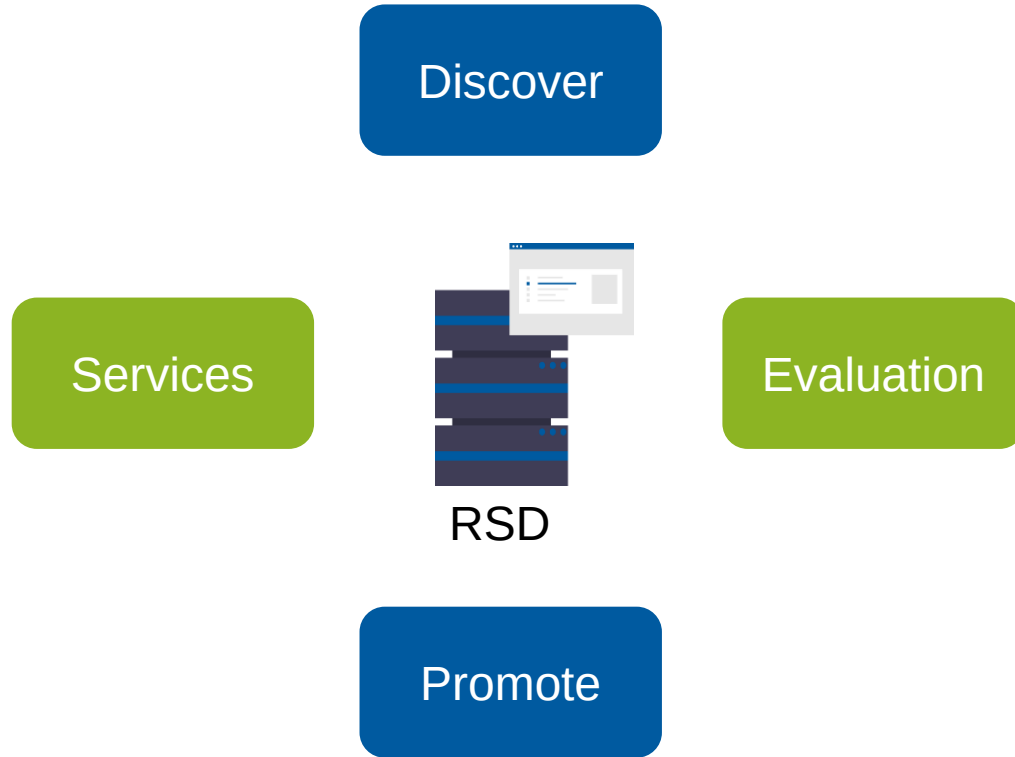
Merge

Figures of merit

Show results from: Calculations within dist

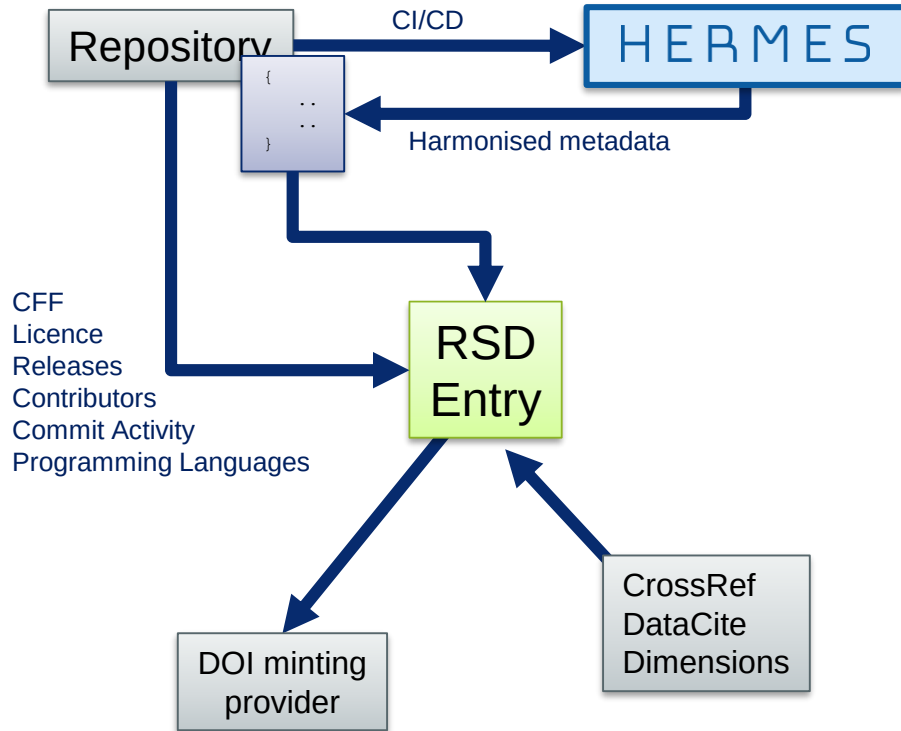
www.hifis.net

Vision



Das RSD als Service für RSEs

Metadaten Aufbereitung vor der Publikation



Ziel

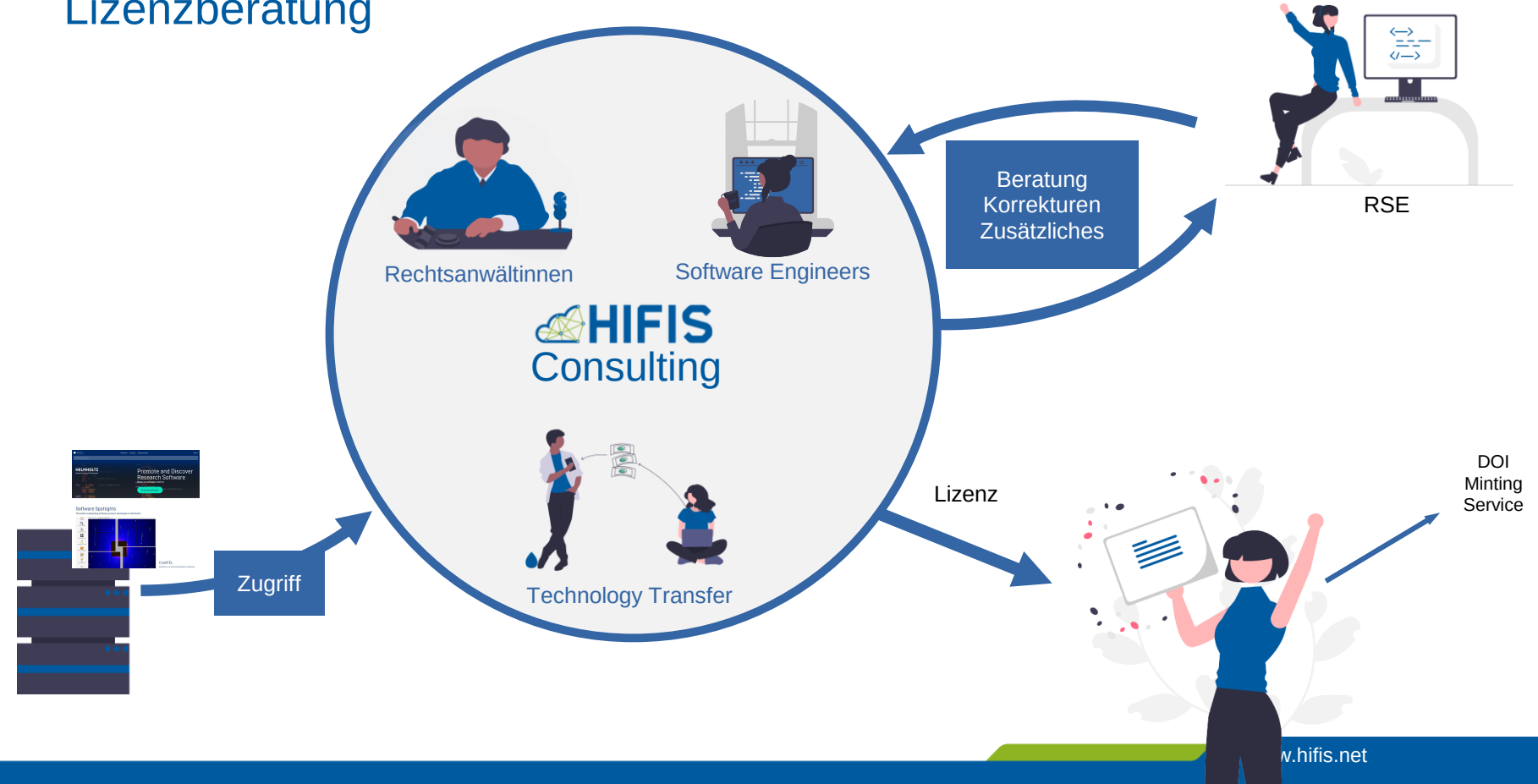
- HERMES Output als Grundlage
- Vorbereitung der Metadaten
- Weiterleitung an DOI minting provider

Nach der Veröffentlichung

- Grundlage für Impact Analyse schaffen
- Autoscraping von Zitationen
- Software Heritage

Das RSD als Service für RSEs

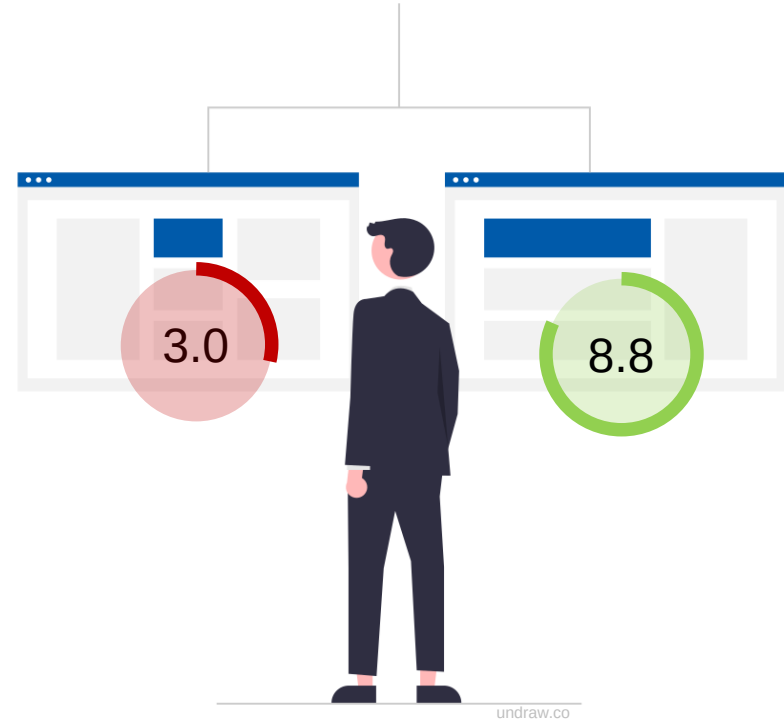
Lizenzberatung



FAIRness und Nachhaltigkeit erhöhen

Kompetitive Anreize schaffen

- Schaffung eines Qualitätsindikators
- Auswertung von
 - FAIR Aspekten, zB
 - Code Qualität
 - Vorhandensein von Tests
 - REUSE Kompatibilität
 - Art der Lizenz (FOSS vs proprietär)
 - Dokumentation
 - Zitationen
 - Projektaktivität (commits)
- WIP: [Software Quality Assurance Framework](#)



Sichtbarkeit von Forschungssoftware erhöhen

Helmholtz Software Spotlights und Awards

The screenshot shows the HIFIS (Helmholtz Federated IT Services) website. The header includes the HIFIS logo and a search icon. Below the header is a blue banner with the text 'openCARP' and a graphic of a globe. The main content area is titled 'openCARP - The open cardiac electrophysiology simulator'. It includes a paragraph describing the simulator, a diagram showing the workflow from 'Single Cell' to 'Tissue' to 'Organ' to 'Observations', and a list of keywords: 'modeling & simulation', 'computational cardiology', 'electrophysiology', and 'in silico trials'. The page also lists contributing organizations like the Karlsruhe Institute of Technology (KIT) and the Medical University of Graz.

HIFIS | HELMHOLTZ FEDERATED IT SERVICES

openCARP

◀ Back to Software Spotlights Overview ▶

openCARP - The open cardiac electrophysiology simulator

openCARP is an open cardiac electrophysiology simulator for in-silico experiments. Its source code is public and the software is freely available for academic purposes. openCARP is easy to use and offers single cell as well as multiscale simulations from ion channel to organ level. Additionally, openCARP includes a wide variety of functions for pre- and post-processing of data as well as visualization. The python-based CARPutils framework enables the user to develop and share simulation pipelines, i.e. automating in-silico experiments including all modeling/simulation steps.

Overview of typical steps in an advanced cardiac electrophysiology simulation study.

openCARP offers a wide range of functionality which enables you to

Centres
Karlsruhe Institute of Technology (KIT)

Contributing organisations
Medical University of Graz
Liryc Bordeaux
Numericor

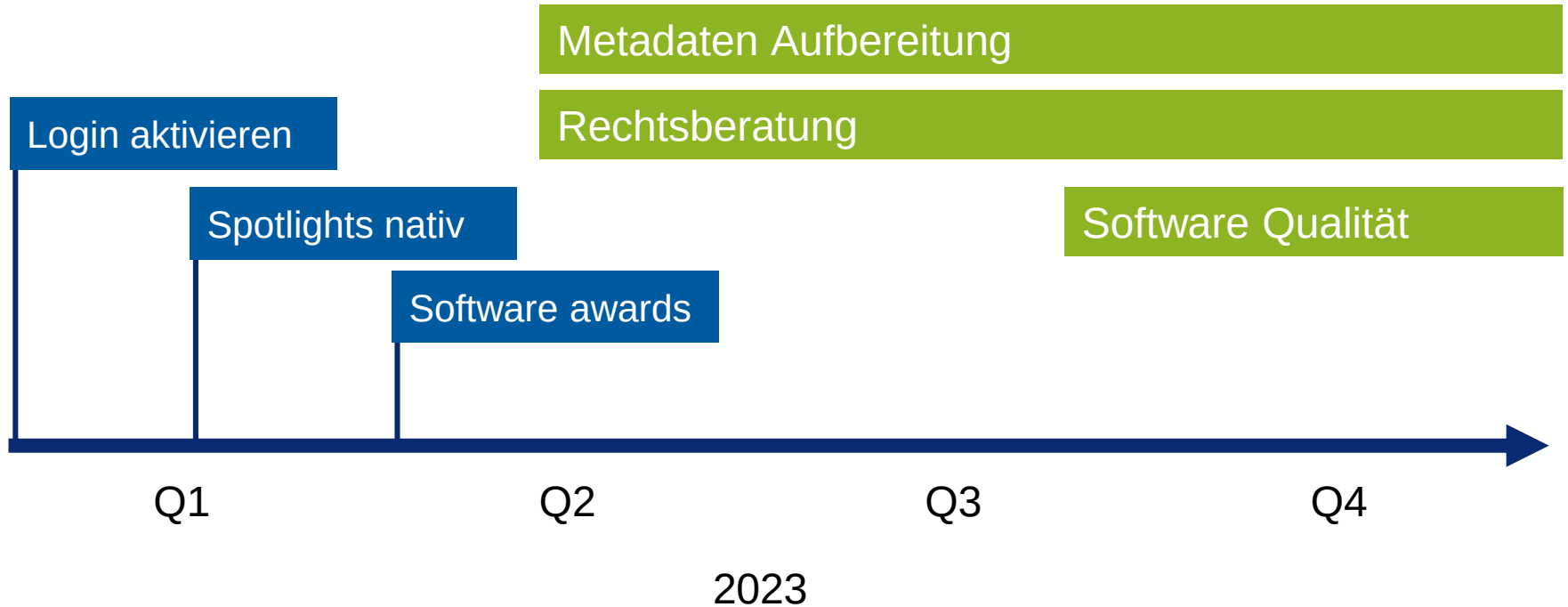
Keywords
modeling & simulation
computational cardiology
electrophysiology
in silico trials

Research field
Health

Scientific community
Computational Cardiology

- Aktuell auf <https://hifis.net/spotlights>
 - 30 online
 - Ca. 70 auf Warteliste
- RSD wird neue Plattform für Spotlights
- Auszeichnungen für herausragende Softwareprojekte

Roadmap 2023



Testnutzer gesucht!

Bei Interesse:

support@hifis.net

Helmholtz Platform for Research Software Engineering – Preparatory Study (HiRSE_PS)

Markus Diesmann, Achim Streit,
Robert Speck, René Caspart



Together with Stefan Blügel (Jülich), Markus Götz (KIT), Johannes Reuther (HZB),
Christian Cyron, Regine Willumeit-Römer & Daniel Höche (Hereon)

24.11.2022, 3. Helmholtz Open Science Forum Forschungssoftware



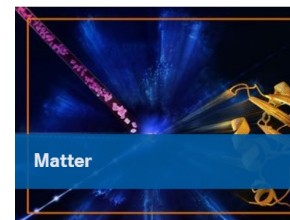
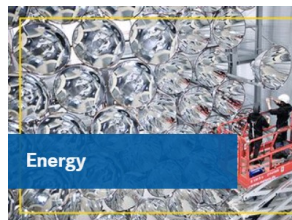
Research Software Engineering (RSE) – Why?

(Open Source) Software = 

- Key component of scientific work
- Software $\approx$ data $\approx$ devices
- Software = research infrastructure
- Valuable assets

**BETTER
SOFTWARE
BETTER
RESEARCH**

in ALL research fields of HELMHOLTZ



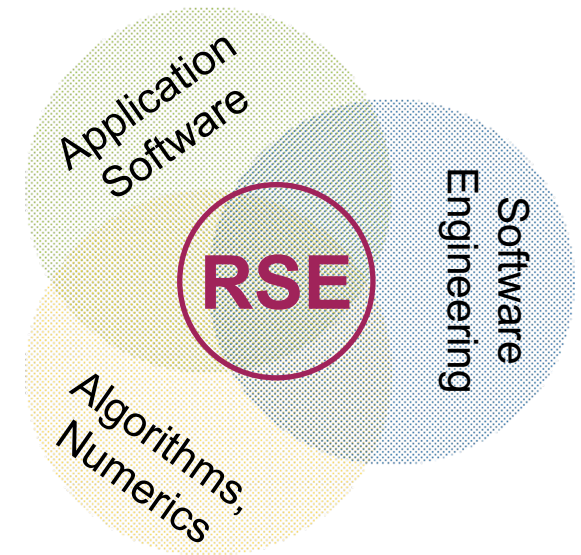
Sources: <https://www.software.ac.uk/about>, <https://www.helmholtz.de/en/research/>

- **Recommendations by the reviewers of the strategic PoF evaluation RF Information**
 - „We strongly recommend **additional activity dedicated to software tools and algorithms** ... This needs to be supported by **substantial additional funding** as it cannot be strengthened easily within the normal systematics. ... In the area of software development, **Helmholtz needs to strengthen its activities significantly.**”
- Example: Netherlands eScience Center Strategy Paper “**Raising the Profile of Research Software: Recommendations for Funding Agencies and Research Institutions**”, (Forward by Bert Meijer, panel chair strategic evaluation, [DOI: 10.5281/zenodo.3378572](https://doi.org/10.5281/zenodo.3378572))
 - “If open science is to truly lead towards better, more transparent, and reproducible research, then **research software needs to be treated in equal footing to research data and publications at the policy level and in practice.**”

- **Q4-2019:** initial feedback at strategic evaluation
- Q1-2020: eMB-meeting with Software-Breakout; afterwards first concept HiRSE developed for Prof. Marquardt (coordinator of RF Information, chairman of FZJ)
- Q2-2020: idea of a pilot pHiRSE developed for Prof. Marquardt and his meeting with BMBF
- Q3-2020: eMB-meeting with presentation of HiRSE + breakout
- Q4-2020 / Q1-2021: Definition of software survey in RF-I
- Q1-2021 / Q2-2021: Survey and its evaluation + preparation for and presentation at strategic advisory board (SAB)
- **Q3-2021:** MB-decision RSE topic as Innovation Pool Project + working on the project draft + submission to MB and BMBF

Continuous and accompanying discussions:
incubator
„Information & Data Science“,
Helmholtz
DigiStrat
Taskforce,
ThinkTank

- **Software as research infrastructure** (open, reliable, sustainable, reproduceable)
- Establish **Research Software Engineering (RSE)** as research and working area
- **Separation of Concerns:** models, numerical methods, verification, address and optimize efficiency separately
- **Sustainable paradigm shift** through $Cx = CI/CT/CD$ (Continuous Integration/Testing/Deployment)
- **Education, training and promoting young** of Research Software Engineers
- **Co-Design** through early integration of computer scientists, mathematicians and administrators in the development of a community



- **HiRSE_PS** can only provide a first impetus of a RF-wide (much larger) activity
- **3 pillars** in 2 work packages
 - Specific CSI-(Community Software Infrastructure)-groups for already existing/established Community Codes Teams
 - NEST and FLEUR (P2, FZJ), HeAT (P1, FZJ/KIT/DLR/Intel), PFFRG (P2, HZB), 4C (P3, Hereon)
 - Central support and consulting unit for RSE in program 1 of the RF-Information (FZJ, KIT)
 - Open RSE Forum/Seminar for a regular exchange of information and best practices

Innovation pool project HiRSE_PS

Our website <https://www.helmholtz-hirse.de/>



WP 1: CSI-groups (PI: Markus Diesmann)

Community Software Infrastructure Groups



■ Goals of the WP

- Establish six CSI Gruppen (topical width across all 3 programs in the RF-Information and diversity of already established and young codes)
- Fostering sustainability and long-term stability of specific codes
- Supporting the community
- Generates insights about the requirements for WP2

■ Structure and location of CSI-groups

- Expert knowledge from a scientific domain
- High community trust through in-person-responsibilities in the domain-institute

■ Tasks of the CSI-groups

- Coordination of the development, e.g. doing code reviews, generating releases, monitoring of Cx-technologies
- Taking over hard and longer-lasting development tasks (e.g. refactoring of existing codes)
- Organization of trainings and Hackathons



Goals and structure

- **Goals:**

- Establishing the technological basis for RSE
- Supporting established CSI-groups and codes in Cx-environment usage and SW-Engineering
- Taking young codes by the hand introducing modern RSE-practises

- **Usage of modern Supercomputing infrastructures**

- E.g., JUWELS (FZJ-JSC) or HoreKa (KIT-SCC)
- Future Technologies Partition (FTP) at KIT-SCC for CI on different HPC resources and architectures
- Cloud resources via OpenStack in the Helmholtz Data Federation (HDF)

- **3 sub-WPs**

1. Framework for CI/CT/CD
2. Support & consulting
3. Community building & networking

- HiRSE Seminar started in April 2022
- Broad range of topics (politics → tech)
- Every 2-4 weeks, no fixed date or time
- Online only for easy and quick access
- About 20-50 participants per talk
- Up-to-date information on <https://www.helmholtz-hirse.de/>
- Access link via Mattermost and local mailing lists/chats, now also **via de-RSE**
- Ideas for new topics/speaker? Let us know!

• December 01, 2022 - December 01, 2022 (virtual)

9th HiRSE Seminar

On December 01, 2022, 3:00pm, **Adrian Kummerländer** from [KIT/IANM](#) will continue the HiRSE Seminar with his talk on **Lattice Boltzmann Performance Engineering in OpenLB - Refactoring a legacy code to state-of-the-art performance on heterogeneous HPC clusters**.

• November 10, 2022 - November 10, 2022 (virtual)

8th HiRSE Seminar

On November 10, 2022, 3:30pm, **Heidi Seibold** will continue the HiRSE Seminar with her talk on **How to improve research quality through Research Software Engineering and Open Science**.

• October 27, 2022 - October 27, 2022 (virtual)

7th HiRSE Seminar, at FZJ OS Week

On October 27, 2022, 2pm, **Jessica Mitchell** from [FZJ/INM-6](#) will contribute to the FZJ Open Science Week with her HiRSE Seminar talk on **Docs-as-code: how to write documentation with developers**.

• October 13, 2022 - October 13, 2022 (virtual)

6th HiRSE Seminar

On October 13, 2022, 4pm, **Claire Wyatt** from the [Software Sustainability Institute](#) (and very soon [FZJ/JSC](#)) will continue the HiRSE Seminar with her talk on **2012 to 2022 - Celebrating 10 years of the RSE movement**.

- HiRSE_PS builds the foundation for digitalization of scientific flow of work
- The results of HiRSE_PS can be used to optimize the HiRSE concept concerning content and resource distribution and need
- At the end of the 2 years of HiRSE_PS: workshop to evaluate the experience of the central part (WP2) and the code-specific part (WP1) as well as their interaction
 - ➔ Revision of the HiRSE concept and recommendations for funding
- **Outlook**
 - HiRSE_PS as ramp-up start of a much longer funding of the HiRSE concept, also in order to reduce international overtaking
 - Establish RSE as a topic in RF Information as soon as possible

HELMHOLTZ

Open Science