

# Forschungssoftware in Deutschland

RDA DE 2021

*Frank Löffler <sup>1</sup>, Christian Busse <sup>2</sup> & Stephan Janosch <sup>3</sup>*

1 - Friedrich-Schiller-Universität Jena

2 - Deutsches Krebsforschungszentrum (DKFZ), Heidelberg

3 - Max-Planck-Institut für molekulare Zellbiologie und Genetik, Dresden



GESELLSCHAFT FÜR  
FORSCHUNGS SOFTWARE

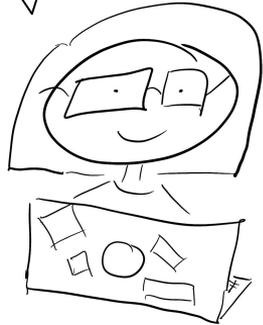
# Agenda

- Vorstellung
- Anknüpfungspunkte im Konferenzprogramm
- Software vs. Daten
- Anderer Ansatz: Handreichungen und Positionspapiere
- RSE/FDM Gruppen
- Feedback

Und was machst du so?

Ein was?!

Ich bin Research Software Engineer!



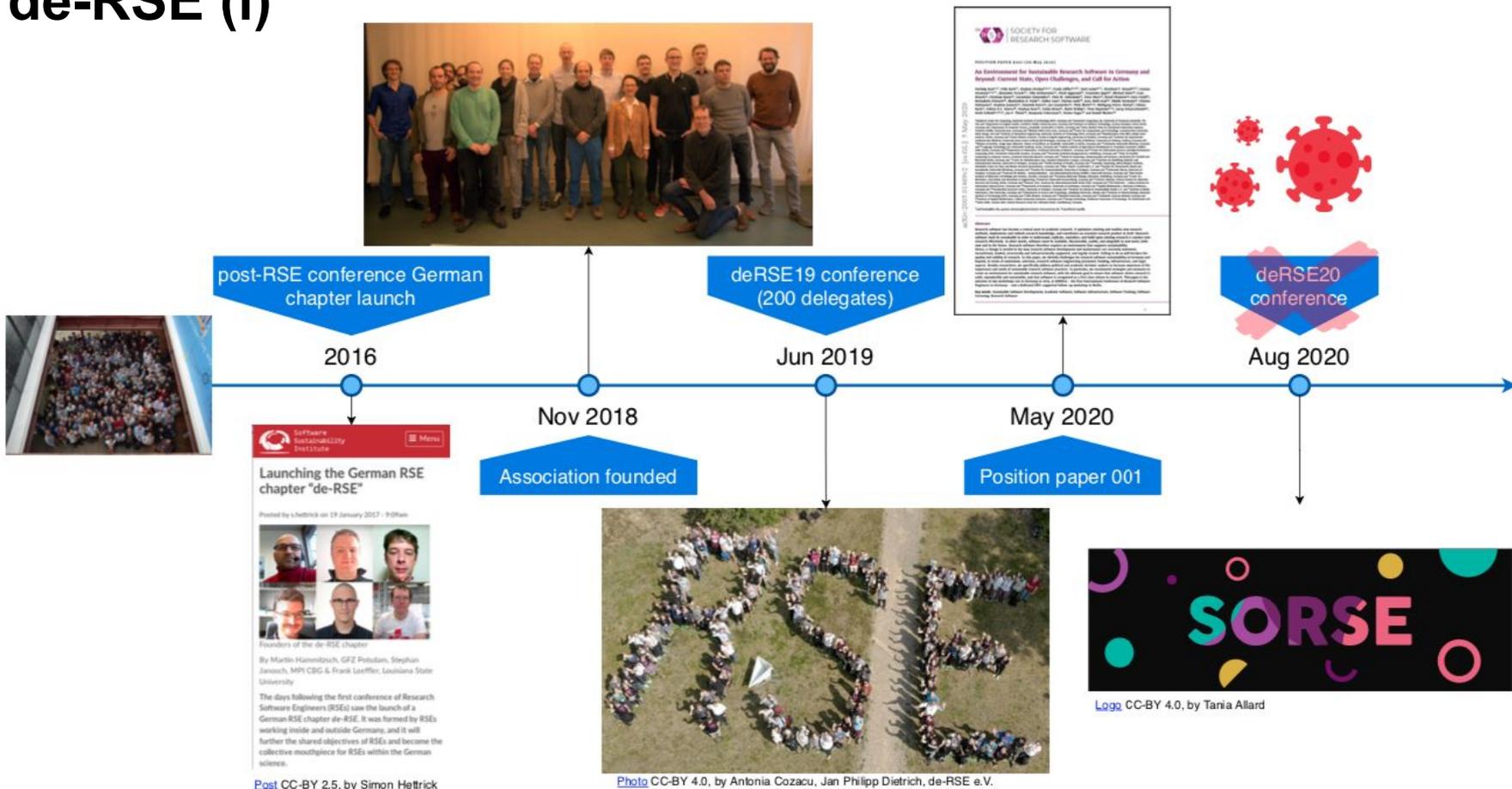
Wir sind Leute die ...

- Nützliche Werkzeuge in Software gießen
- Andere Forschende bei der Auswahl von Werkzeugen beraten und schulen
- Bei Reproduzierbarkeit durch wissenschaftliches Rechnen helfen
- Unterstützung bei Erstellung von SW liefern und gute Praxis lehren
- Ein **sehr** breites Spektrum haben

=

Research Software  
Engineers (RSEs)

# de-RSE (i)



Post CC-BY 2.5, by Simon Heitrick



Photo CC-BY 4.0, by Antonia Cozacu, Jan Philipp Dietrich, de-RSE e.V.

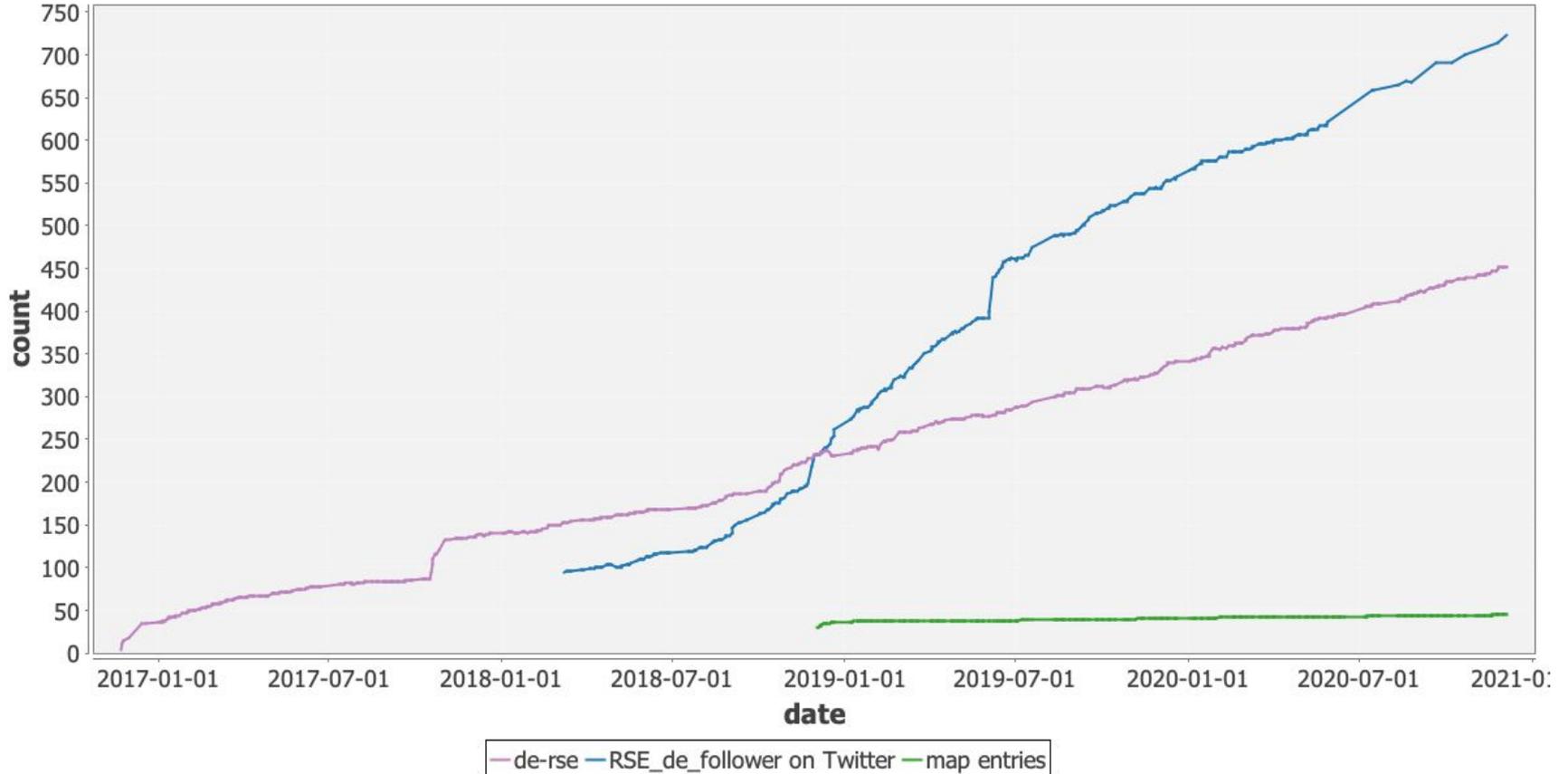


Logo CC-BY 4.0, by Tania Allard

# de-RSE (ii)

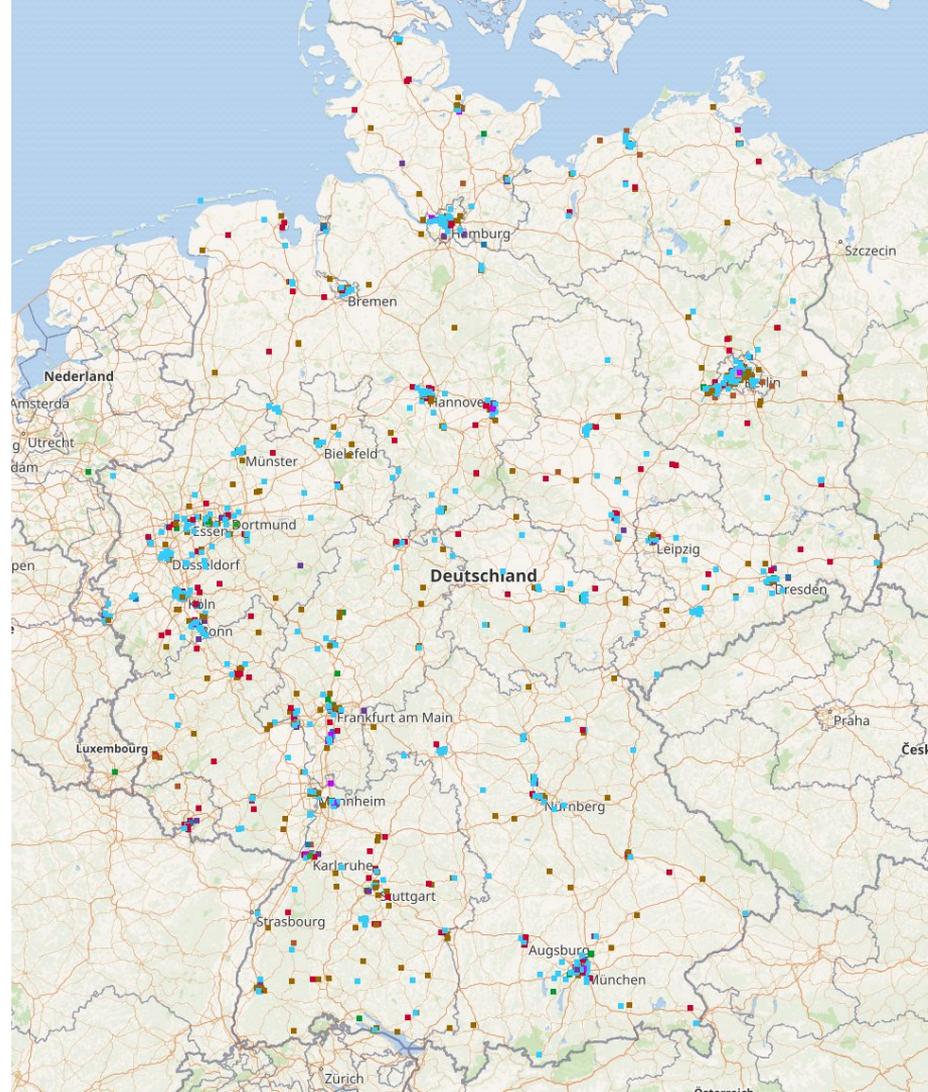
Email: [liste@de-rse.org](mailto:liste@de-rse.org)

Twitter: [@RSE\\_de](https://twitter.com/RSE_de)



# 2019 Orte mit DFG PID

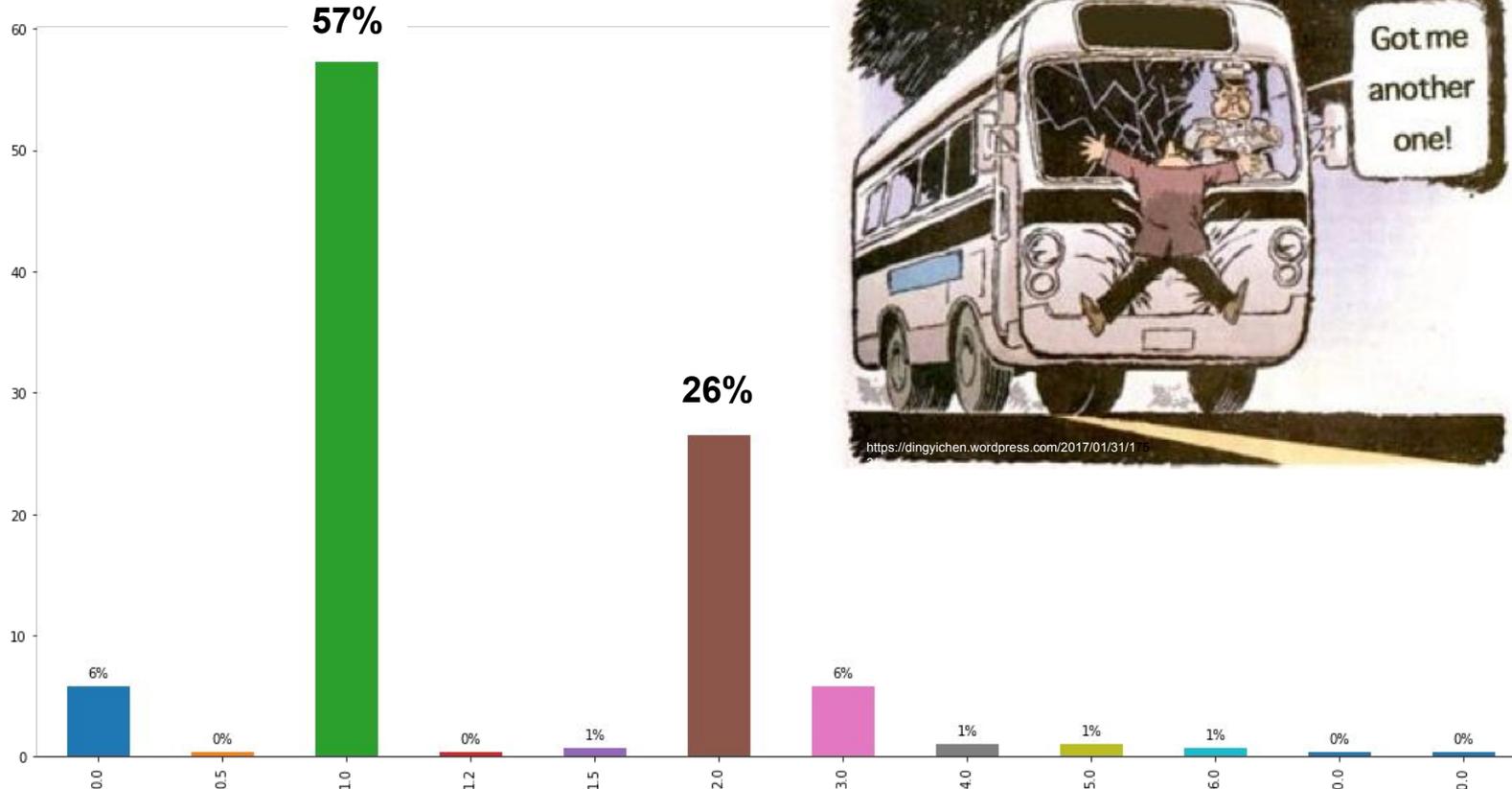
Einrichtungsart	Anzahl
Akademie d. Wiss.	15
Bundesforschung	49
FhG	84
HGF	21
Int. wiss. Org.	29
Landesforschung	247
MPG	91
Musik/Kunst-HS	57
Sonst. Forschung	952
Universitäten	116
WGL	98
Weitere HS/FS	260



# Wir... sind nicht allein!

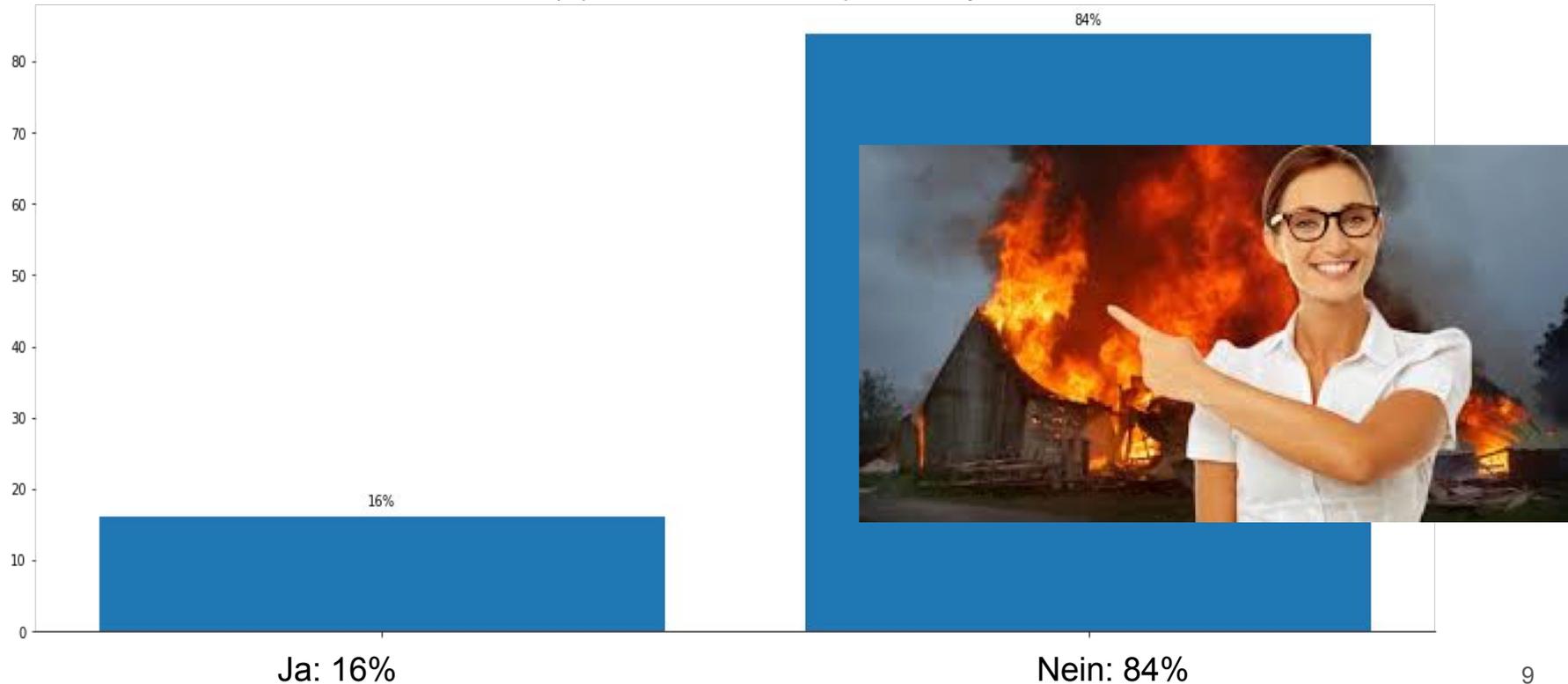
- be-RSE, NL-RSE, Nordic RSE, RSE-AUNZ, UK-RSE, US-RSE Menschen
  - International Council of RSE Associations
- UK-SSI (de-SSI?) Institution
- Force11 Gremien
  - ~~Software citation WG~~, Software citation impl. WG
- RDA & ReSA (seit 2019)
  - z.B. FAIR4RS
- Carpentries (Software, Data, Library) Training
- [German Reproducibility Network](#) (GRN) Qualität

# Nachhaltigkeit: Schlüsselpersonal



# Nachhaltigkeit: Übergabepläne

Current proportion of Presence of a transition plan for Germany



# Programm RDA-DE 2021

09:00	09:00 <b>Mapping the landscape of visualization tools (EcoDM - Schulungsveranstaltung)</b>	09:00 <b>RDA4Newbies</b> Online Event	09:00 <b>Förderung der europäischen Digitalwirtschaft: Welche Ziele sollen realisiert werden? Wie ist eine praktische Umsetzung möglich?</b>	09:00 <b>Horizon Europe - Anforderungen an Open Science und FAIR Data</b> Online Event
11:00	11:00 <b>Begrüßung, Neues von RDA / Datenstrategie der Bundesregierung</b> Online Event	11:00 <b>Covid-19 – RDA Guidelines und offene Verwaltungsdaten</b> Online Event	11:00 <b>Update vom NFDI Direktorat</b> Online Event	11:00 <b>Offene Verwaltungsdaten des Bundes</b> Online Event
13:30	13:30 <b>Research Software Engineering in Deutschland</b> Online Event	13:30 <b>Der DFG-Kodex "Leitlinien zur Sicherung guter wissenschaftlicher Praxis" im Kontext Forschungsdaten</b>	13:30 <b>Dateninfrastrukturen in Deutschland – „Wer macht was?“</b> Online Event	13:30 <b>Keynote: Big Data und sensible Daten</b> Online Event
15:30	15:30 <b>Vorstellung des Entwurfs des Datennutzungsgesetzes</b> Online Event	15:30 <b>EOSC</b> Online Event		

# Rechtliche Aspekte

09:00	09:00 <b>Mapping the landscape of visualization tools (EcoDM - Schulungsveranstaltung)</b>	09:00 <b>RDA4Newbies</b> Online Event	09:00 <b>Förderung der europäischen Digitalwirtschaft: Welche Ziele sollen realisiert werden? Wie ist eine praktische Umsetzung möglich?</b>	09:00 <b>Horizon Europe - Anforderungen an Open Science und FAIR Data</b> Online Event
11:00	11:00 <b>Begrüßung, Neues von RDA / Datenstrategie der Bundesregierung</b> Online Event	11:00 <b>Covid-19 – RDA Guidelines und offene Verwaltungsdaten</b> Online Event	11:00 <b>Update vom NFDI Direktorat</b> Online Event	11:00 <b>Offene Verwaltungsdaten des Bundes</b> Online Event
13:30	13:30 <b>Research Software Engineering in Deutschland</b> Online Event	13:30 <b>Der DFG-Kodex "Leitlinien zur Sicherung guter wissenschaftlicher Praxis" im Kontext Forschungsdaten</b>	13:30 <b>Dateninfrastrukturen in Deutschland – „Wer macht was?“</b> Online Event	13:30 <b>Keynote: Big Data und sensible Daten</b> Online Event
15:30	15:30 <b>Vorstellung des Entwurfs des Datennutzungsgesetzes</b> Online Event	15:30 <b>EOSC</b> Online Event		

# Recht in der Praxis

Darf man Reverse Engineering (RE) der Mess-Software benutzen, um an die Rohdaten ranzukommen?

# Recht in der Praxis

Darf man Reverse Engineering (RE) der Mess-Software benutzen, um an die Rohdaten ranzukommen?

*RDA-Tagung. Thema ansprechen. Umsteuern von Neugeräten schnell (Reaktionszeit von 10-20 Jahren, s.o.). very top -> down Richtlinie geben: "NFDI steht für FAIRe Daten", deswegen müssen Neugeräte Rohdaten in offenen Formaten liefern*

# Recht in der Praxis

Darf man Reverse Engineering (RE) der Mess-Software benutzen, um an die Rohdaten ranzukommen?

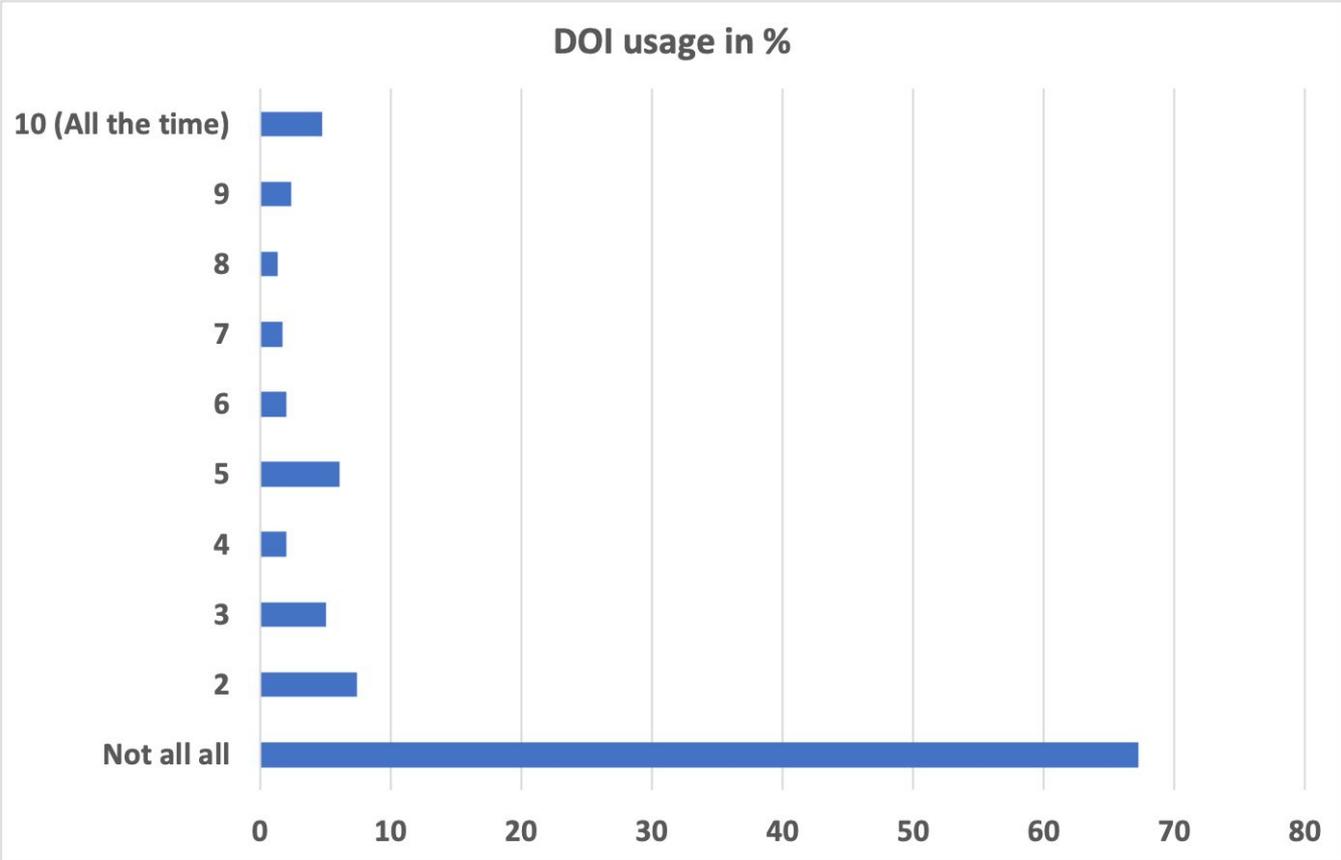
*RDA-Tagung. Thema ansprechen. Umsteuern von Neugeräten schnell (Reaktionszeit von 10-20 Jahren, s.o.). very top -> down Richtlinie geben: "NFDI steht für FAIRe Daten", deswegen müssen Neugeräte Rohdaten in offenen Formaten*

2019 ... GeschGehG (Gesetz zum Schutz von Geschäftsgeheimnissen) geändert.  
"untersuchen, rückzubauen oder zu testen, um darin enthaltenes Know-How zu erkennen"... erlaubt. Nicht für Prototypen.

# “Die” Richtlinie

09:00	09:00 <b>Mapping the landscape of visualization tools (EcoDM - Schulungsveranstaltung)</b>	09:00 <b>RDA4Newbies</b> Online Event	09:00 <b>Förderung der europäischen Digitalwirtschaft: Welche Ziele sollen realisiert werden? Wie ist eine praktische Umsetzung möglich?</b>	09:00 <b>Horizon Europe - Anforderungen an Open Science und FAIR Data</b> Online Event
11:00	11:00 <b>Begrüßung, Neues von RDA / Datenstrategie der Bundesregierung</b> Online Event	11:00 <b>Covid-19 – RDA Guidelines und offene Verwaltungsdaten</b> Online Event	11:00 <b>Update vom NFDI Direktorat</b> Online Event	11:00 <b>Offene Verwaltungsdaten des Bundes</b> Online Event
13:00	Bis 2022			
13:30	13:30 <b>Research Software Engineering in Deutschland</b> Online Event	13:30 <b>Der DFG-Kodex "Leitlinien zur Sicherung guter wissenschaftlicher Praxis" im Kontext Forschungsdaten</b>	13:30 <b>Dateninfrastrukturen in Deutschland – „Wer macht was?“</b> Online Event	13:30 <b>Keynote: Big Data und sensible Daten</b> Online Event
15:30	15:30 <b>Vorstellung des Entwurfs des Datennutzungsgesetzes</b> Online Event	15:30 <b>EOSC</b> Online Event		

# DOIs



# Europäische Infrastruktur

09:00 <b>09:00 Mapping the landscape of visualization tools (EcoDM - Schulungsveranstaltung)</b>	09:00 <b>09:00 RDA4Newbies</b> Online Event	09:00 <b>09:00 Förderung der europäischen Digitalwirtschaft: Welche Ziele sollen realisiert werden? Wie ist eine praktische Umsetzung möglich?</b>	09:00 <b>09:00 Horizon Europe - Anforderungen an Open Science und FAIR Data</b> Online Event
11:00 <b>11:00 Begrüßung, Neues von RDA / Datenstrategie der Bundesregierung</b> Online Event	11:00 <b>11:00 Covid-19 – RDA Guidelines und offene Verwaltungsdaten</b> Online Event	11:00 <b>11:00 Update</b> Online	<div data-bbox="1078 449 1924 1064" data-label="Complex-Block"> <p style="text-align: center;"><b>Strategic Research and Innovation Agenda (SRIA)</b> of the <b>European Open Science Cloud (EOSC)</b> Version 0.9 16 November 2020</p>  <p style="text-align: right;"><b>EUROPEAN OPEN SCIENCE CLOUD</b></p> </div>
13:30 <b>13:30 Research Software Engineering in Deutschland</b> Online Event	13:30 <b>13:30 Der DFG-Kodex "Leitlinien zur Sicherung guter wissenschaftlicher Praxis" im Kontext Forschungsdaten</b>	13:30 <b>13:30 Daten</b> „Wer“ Online	
15:30 <b>15:30 Vorstellung des Entwurfs des Datennutzungsgesetzes</b> Online Event	15:30 <b>15:30 EOSC</b> Online Event		

# Querschnittsthemen

NFDI4



09:00	09:00	09:00	09:00
Ma vis Sc		Förderung der europäischen Digitalwirtschaft: Welche Ziele sollen realisiert werden? Wie ist eine praktische Umsetzung möglich?	Horizon Europe - Anforderungen an Open Science und FAIR Data Online Event
10:00			
11:00		11:00	11:00
Be Da On		Update vom NFDI Direktorat Online Event	Offene Verwaltungsdaten des Bundes Online Event
12:00			
13:00			
13:00			
13:00			
13:00		13:30	13:30
Re De On		Dateninfrastrukturen in Deutschland – „Wer macht was?“ Online Event	Keynote: Big Data und sensible Daten Online Event
14:00			
15:00			
15:00			
15:00			
15:00			
15:00			
15:00			
15:00			
15:00			
17:00			



**Stephan Janosch**  
@StephanJanosch

Cross cutting topics at #nfdi

- \* Data quality
- \* Data literacy and training
- \* Software (inkl Web Apps and curation)
- \* Open Standards
- \* Governance
- \* Interoperability and transfer
- \* Use of existing infrastructures
- \* Form consortia based on data types

[Tweet übersetzen](#)

9:35 vorm. · 14. Mai 2019 · TweetDeck

||| [Tweet-Aktivität anzeigen](#)

12 Retweets 1 Tweet zitieren 15 „Gefällt mir“-Angaben

# FOSS als Produkt - ?!

09:00	09:00 <b>Mapping the landscape of visualization tools (EcoDM - Schulungsveranstaltung)</b>	09:00 <b>RDA4Newbies</b> Online Event	09:00 <b>Förderung der europäischen Digitalwirtschaft: Welche Ziele sollen realisiert werden? Wie ist eine praktische Umsetzung möglich?</b>	09:00 <b>Horizon Europe - Anforderungen an Open Science und FAIR Data</b> Online Event
10:00				
11:00	11:00 <b>Begrüßung, Neues von RDA / Datenstrategie der Bundesregierung</b> Online Event	11:00 <b>Covid-19 – RDA Guidelines und offene Verwaltungsdaten</b> Online Event	11:00 <b>Update vom NFDI Direktorat</b> Online Event	11:00 <b>Offene Verwaltungsdaten des Bundes</b> Online Event
12:00				
13:00				
14:00	13:30 <b>Research Software Engineering in Deutschland</b> Online Event	13:30 <b>Der DFG-Kodex "Leitlinien zur Sicherung guter wissenschaftlicher Praxis" im Kontext Forschungsdaten</b>	13:30 <b>Dateninfrastrukturen in Deutschland – „Wer macht was?“</b> Online Event	13:30 <b>Keynote: Big Data und sensible Daten</b> Online Event
15:00				
16:00	15:30 <b>Vorstellung des Entwurfs des Datennutzungsgesetzes</b> Online Event	15:30 <b>EOSC</b> Online Event		
17:00				

# Software-Daten-Dualismus

- Software: nur Nebenprodukt der Datenerhebung und -analyse?
- Trotz zunehmender Komplexität und Abhängigkeit von Software:  
wenig Professionalisierung des wissenschaftlichen Software-Engineerings
- Daten-zentrische Sichtweise auf Forschungsprodukt “Software”
- Software kann (und sollte) als FAIR Digital Object (FDO) abgebildet werden,  
**aber...**

# Software-Daten-Dualismus

... die **ausschließliche** Darstellung von Software als (Daten-)FDO wird vielen ihrer Aspekte nicht gerecht. Software

- 1) benötigt kontinuierliche Pflege für fortgesetzte Verwendbarkeit
- 2) ist selten statisch, sondern wird ständig weiterentwickelt
- 3) erlaubt die Beteiligung von Dritten bei der (Weiter-)Entwicklung
- 4) hat andere Anforderungen an die Datenmanagement-Infrastruktur
- 5) weist Abhängigkeit von externen Funktionsbibliotheken auf
- 6) kann als Quell- oder Binärcode zur Verfügung gestellt werden

# Software ist eine eigene Klasse von FDO

- Software  $\Rightarrow$  Daten ist eine valide Transformation, aber man muss sich der Implikationen bewusst sein
  - Wenn Software wie Daten behandelt wird, verbleiben nur die gemeinsamen Eigenschaften der Metadaten
- Software sollte als eigene Klasse der FDO erkannt werden

# FAIRe Software?

- Können wir das ursprüngliche FAIR Paradigma <sup>[1]</sup> auf Software anwenden?
- Sollten wir:
  - Die FAIR Kriterien ändern,
  - sie erweitern (mehr “Buchstaben”) oder
  - sie als (zu erweiterndes) Minimalkriterium ansehen?
- "Towards FAIR principles for research software" <sup>[2]</sup>
  - Neuinterpretation der FAIR Kriterien für Software
- [FAIR4RS](#) (gem. Arbeitsgruppe RDA/ReSA/Force11)
- Entwicklung ist ein Prozess, aber FAIR beschreibt einen Zustand
- Natürlich ist "FAIR" nicht gleich "FLOSS"

[ 1 ] Wilkinson MD et al., Sci Data 3:160018 (2016)  
DOI:[10.1038/sdata.2016.18](https://doi.org/10.1038/sdata.2016.18)

[ 2 ] Lamprecht AL et al., Data Science 3:37 (2019)  
DOI:[10.3233/DS-190026](https://doi.org/10.3233/DS-190026)

# FAIR als Minimal Kriterium für Software

- FAIR, FLOSS und nachhaltige Archivierung sind komplementär zueinander
- FAIR ist eine Grundvoraussetzung für **alle** digitalen Objekte (sprich: FDO)
- Software Objekte sollten auch andere Kriterien erfüllen
- Dieses Prinzip wird bereits für wiss. Publikationen verwendet (FAIR + OA)





# allianzinitiative.de



Wissenschaftliches  
Publikationssystem



Digitale Werkzeuge –  
Software und Dienste



Digitale  
Datensammlungen



Förderieren von IT-  
Infrastrukturen



Digitales Lernen,  
Lehren und Vernetzen



Digital qualifiziertes  
Personal



Recht für Wissenschaft  
im digitalen Zeitalter



Wissenschaftspraxis

Herausforderung: Sichtbarkeit und Umsetzung

# Digitale Werkzeuge - Software und Dienste



2018: <http://doi.org/10.5281/zenodo.1172970>

# Digitale Werkzeuge - Software und Dienste



2018: <http://doi.org/10.5281/zenodo.1172970>

2020: <http://doi.org/10.5281/zenodo.4301924>

# Quansight Labs

Quansight Labs is a public-benefit division of Quansight created to provide a home for a "PyData Core Team" which consists of developers, community managers, designers, and writers who create and maintain open-source technology around all aspects of scientific and data science workflows.

## Sustaining the future of Open Source

Projects like NumPy, SciPy, Pandas, Jupyter, Matplotlib, scikit-learn, Dask, Conda, and many others serve as a foundation for data science workflows and a launchpad for AI. We have had and continue to have the privilege of participating in these communities and understand personally the need for people to be funded to work on these projects. We also understand the needs of companies that use these projects, and aim to connect communities and companies to guarantee the health and accelerate development of the projects and technology that tens of millions of users rely on.

See our [Mission](#), [Vision](#) and [Model](#) for more.

## Sponsors



S

# Geschäftsmodelle für Forschungssoftware

# Geschäftsmodell (II)



## Condensed History

- 1987:** Graphics task force at NCSA began work on architecture-independent format and library, **HDF**.




---

- 1990:** NSF provided funding to improve documentation, testing, and user support.



---

- 1994:** NASA selected HDF as standard format for Earth Observing System.




---

- 1996–1998:** DOE tri-labs and NCSA, with additional support from NASA, developed **HDF5**, initially called “BigHDF”.







---

- 2005:** NASA funded development of netCDF-4, a new version of netCDF that uses the HDF5 file format.




---

- 2006:** **The HDF Group**, a non-profit corporation, spun off from NCSA and the University of Illinois.



# F

# Forschungssoftware-Gruppen in Deutschland

- Wo gibt es sie?
- Wie groß sind sie?
- Wie funktionieren diese (Geschäftsmodelle)?
- Was sind die Hauptprobleme?
- (Wie) werden sie evaluiert?

## de-RSE-Umfrage:

- Gruppenliste: <https://bit.ly/3knejh7> (gern mit eintragen)
- Fragenliste: <https://bit.ly/2NmwoWS>

# Nachhaltigkeit

- Investition in die Zukunft
  - Hohe sichtbare Grenzkosten von Hardware (€€€€)
  - Vergleich zu Software (€)
- Personal
  - Karriere
  - Anerkennung
  - Kuratoren und Manager
- Ort
  - Verzeichnis (swMath arbeitet mit 100 000 €/a)



<https://www.reading.ac.uk/RES/rdm/about/res-rdm-lifecycle.aspx>

# de-RSE im Jahr 2021

- Stärkung der Gemeinschaft
  - keine Konferenz, dafür dezentrale Veranstaltungen
  - Partnerschaften mit Fachgesellschaften: Vernetzung
  - Fellowships
- NFDI Mitarbeit
- Plattform für offene AGs, z.Bsp.:
  - RSE-(Doppel-)rolle im dt. System
  - Evaluation / Metriken
  - de-SSI?



Danke!

Fragen!



# Fragen an Sie

- Wer hat FDM-Stellen geschaffen oder plant dies zu tun?
- Wer schafft RSE-Stellen?
- Wer hat diese Handreichung schon mal gesehen?
- Wurde die Handreichung “von oben” durch die Institution geschickt?
- Ausbrechen nicht möglich
  - Evaluation
  - Anerkennung (recognition), Zitation
  - Software vs. Daten

