Common Infrastructures

Nationale Forschungsdaten Infrastruktur

How to improve the visibility and added value of RSE(s) in NFDI

Florian Thiery (CAA-DE e.V. & LEIZA)

Bernd Flemisch (University of Stuttgart)

Jan Bernoth (University of Potsdam)

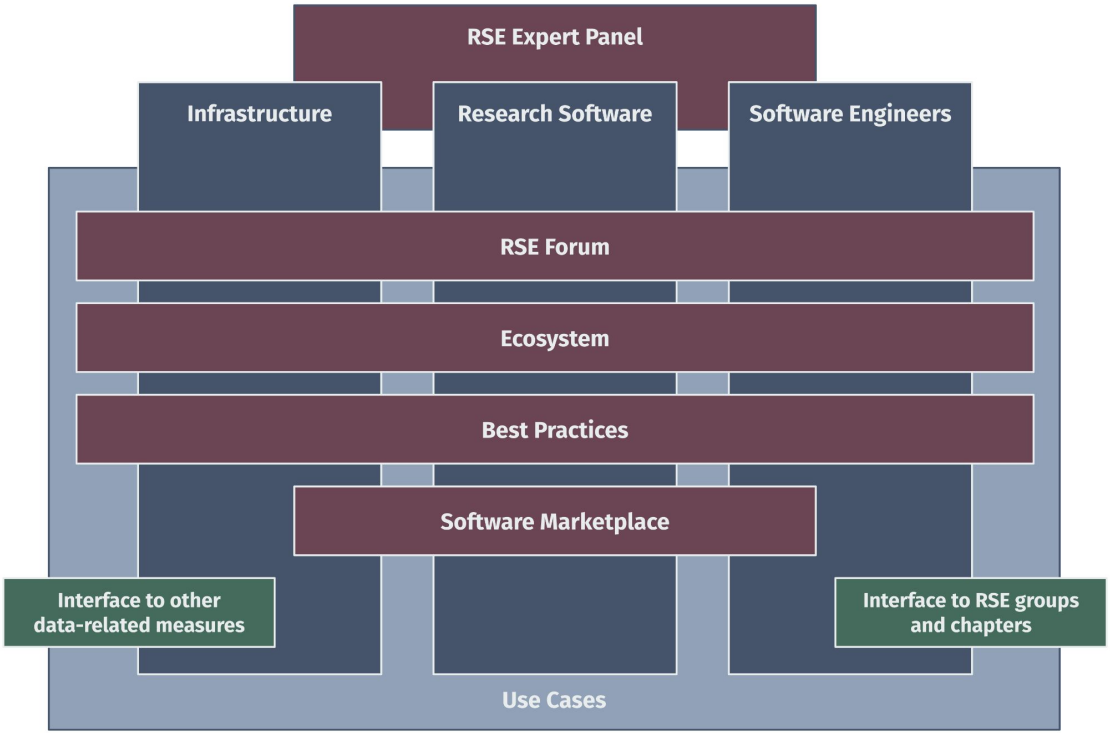
Corinna Seiwerth (FAU Erlangen-Nürnberg)

Jan Linxweiler (TU Braunschweig)

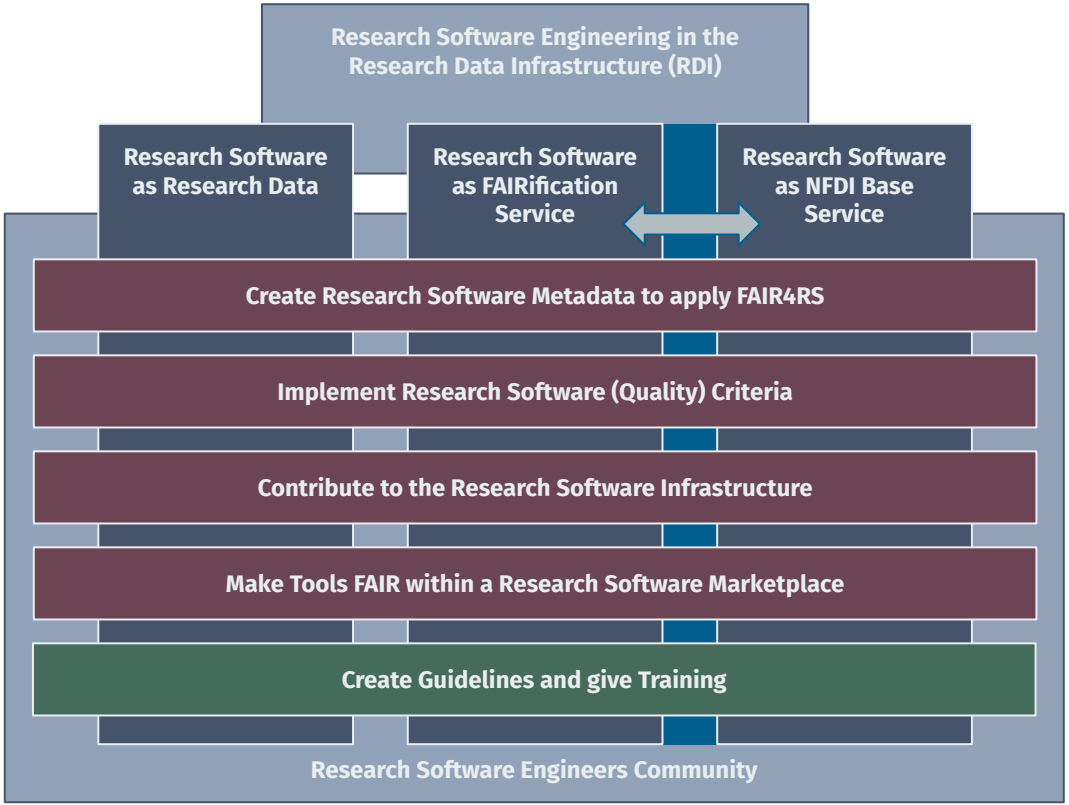
Linnaea Söhn (Academy of Sciences and Literature | Mainz)

deRSE25, 27.02.2025, KIT: Karlsruhe, Session: Meet-Up

Research Software Engineering Layer Cake in the NFDI



Roles of RSEs and RS in the RDM-RDI NFDI-Context



RSE @ Research Data Infrastructure

- RS in the RDM/NFDI context is community-driven and interdisciplinary
- RS can act as Research Data itself, as FAIRification Service within the RDM process and can become a NFDI Base Service
- This needs common RS metadata and quality criteria, to create an interlinked and distributed RS infrastructure and RS Marketplace
- Within this RSE in RDM-RDI infrastructure RSE Training for the RSE Community is needed

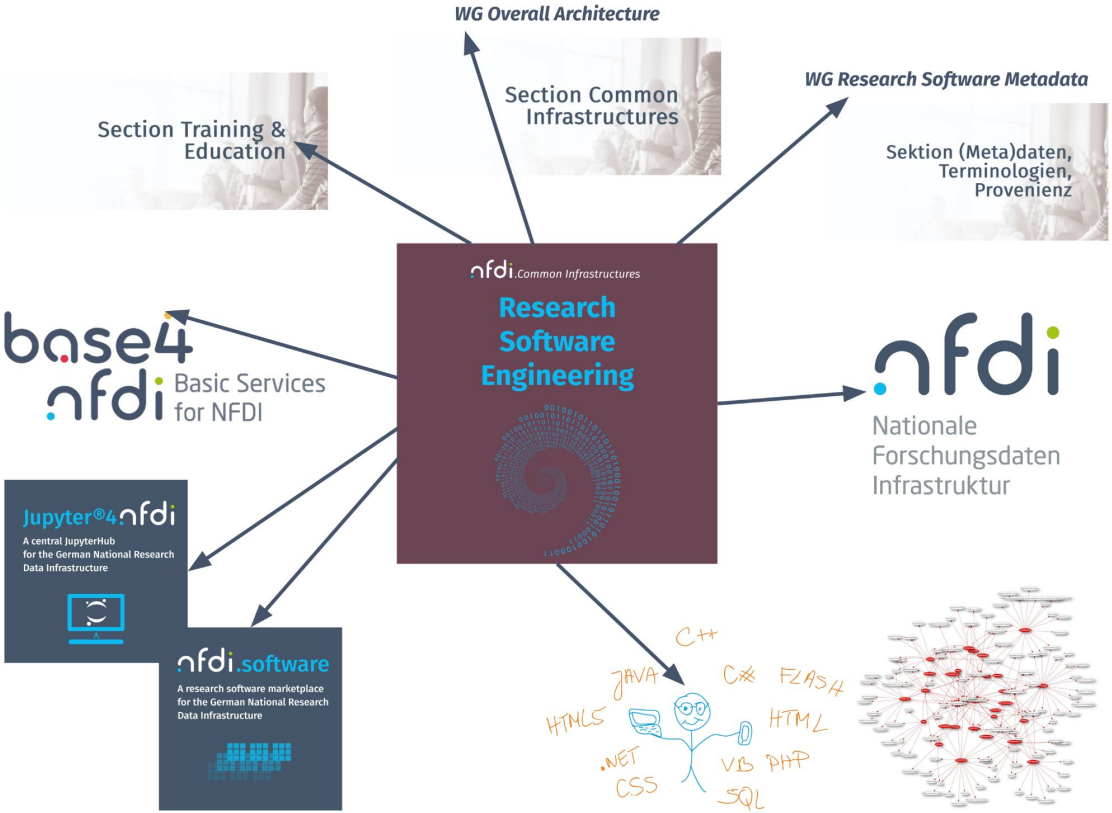
RSE @ NFDI

Research Software Engineering is fundamental to the NFDI.

Within the NFDI e.V., several "sections" are dealing with

- the *Section Common Infrastructures*
  with its working groups on
  - **WG Research Software Engineering (RSE)**
- the *Section (Meta)Data, Terminologies, Provenance*
  - **WG Research Software Metadata**
- the *Section Education & Training*
  - *Working Group on RSE missing?*

Connections of the WG RSE to other NFDI entities



Active Tasks in 2022-25

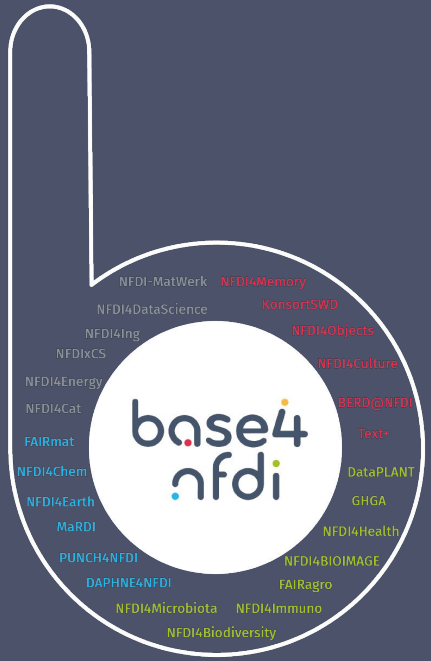
- **INFRA-RSE-T01: Jupyter**
- INFRA-RSE-T02: Training Materials
- **INFRA-RSE-T03: Software Marketplace**
- **INFRA-RSE-T04: Mission Statement "Software Engineering in the NFDI"**
- **INFRA-RSE-T05: Status Quo Survey and Report**
- INFRA-RSE-T06: Monitoring of other overarching initiatives (EOSC, GAIA-X, ...)
- INFRA-RSE-T07: NFDI Software Ecosystem
- INFRA-RSE-T08: Identification of Use Cases
- **INFRA-RSE-T09: Quality Criteria for Research Software**
- INFRA-RSE-T10: Criteria for NFDI Software Components

How to Join the WG RSE at INFRA

- Work on tasks in small subgroups of 1-4 persons
- Report and reflect at monthly meetings of the WG
- Sign up for the mailing list by sending a “Subscribe” to section-infra-wg-rse-join@lists.nfdi.de
- Join the monthly meetings
  - currently on each month’s last Monday at 13:00 CET

Base4NFDI


DAPHNE4NFDI
NFDI4Microbiota
DataPLANT
FAIRagro
KonsortSWD
NFDI4Immuno
NFDI4BIOIMAGE
NFDI4Culture
Text+
NFDI4Chem
Data Usage
Data Production
PUNCH4NFDI
NFDI4Health
BERD@NFDI
NFDI4Earth
NFDI4Objects
GHGA
MaRDI
NFDIxCS
NFDI4Cat
FAIRmat
NFDI4Energy
NFDI-MatWerk
NFDI4DataScience
NFDI4Biodiversity
NFDI4Memory
NFDI4Ing
Data Provision
Data Science & AI
Training Modules
META(DATA), TERMINOLOGY, PROVENANCE
Multi Cloud
FAIR Digital Objects
Data Protection & Research Ethics
Research Software Engineering
Electronic Lab Notebooks
Identity & Access Management
EDUCATION & TRAINING
Persistent Identifiers
INDUSTRY ENGAGEMENT
Licensing & Copyright
Data Integration
Metadata Schemata
Terminology Harmonisation
COMMON INFRASTRUCTURES
Knowledge Graphs
Anonymising & Pseudonymising
Long Term Archiving
Error Culture
ETHICAL, LEGAL & SOCIAL ASPECTS
Data Management Plans
Search & Harvesting


NFDI-MatWerk
NFDI4Memory
NFDI4DataScience
KonsortSWD
NFDI4Ing
NFDI4Objects
NFDIxCS
NFDI4Culture
NFDI4Energy
BERD@NFDI
NFDI4Cat
Text+
FAIRmat
NFDI4Chem
DataPLANT
NFDI4Earth
GHGA
MaRDI
NFDI4Health
PUNCH4NFDI
NFDI4BIOIMAGE
DAPHNE4NFDI
FAIRagro
NFDI4Microbiota
NFDI4Immuno
NFDI4Biodiversity
base4 nfdi

Base4NFDI meets INFRA-WG-RSE

Jupyter®4nfdi

A central JupyterHub
for the German National Research
Data Infrastructure

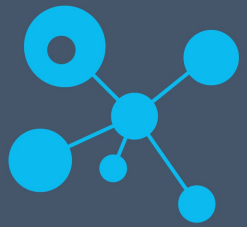
nfdi.software

A research software marketplace
for the German National Research
Data Infrastructure

INFRA-WG-RSE meets Interconnectivity Services

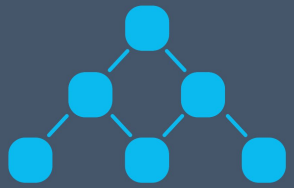
KGI4 nfdi

Knowledge Graph Infrastructure
for the German National Research
Data Infrastructure



TS4 nfdi

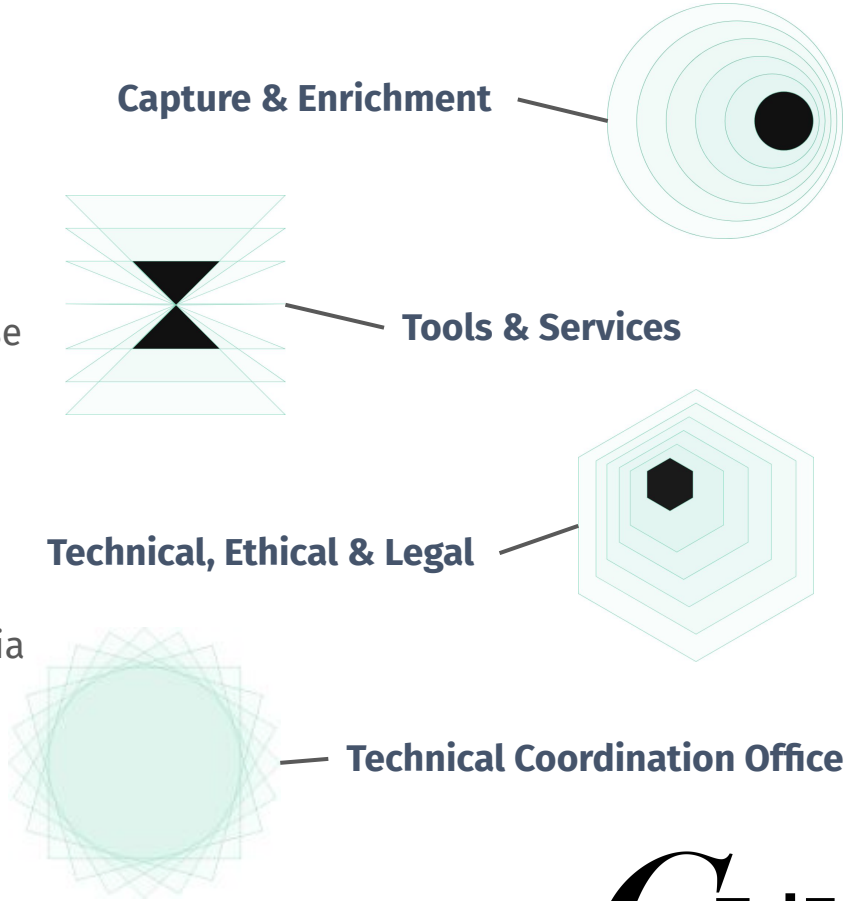
Terminology Services
for the German National Research
Data Infrastructure



RSE(s) within the NFDI

RSE @ NFDI4Culture

- Development of **FOSS Services** such as Semantic Kompakkt and Wikibase4Research
- **Fundings** for research software development via Flex Funds
- **Guideline** for the sustainable development and use of research software
- **Visibility** of research software via the Culture Registry for tools and services
- **Interchange** about RSE in the Expert Forum for sustainable software development
- **Capturing and integration** of research information into european infrastructures, such as the EOSC, via the Culture Information Portal
- **Connection** of research data via the Culture Knowledge Graph and the NFDIcore Ontology
- **Coordination and implementation** of an overarching technical infrastructure for the entire consortium



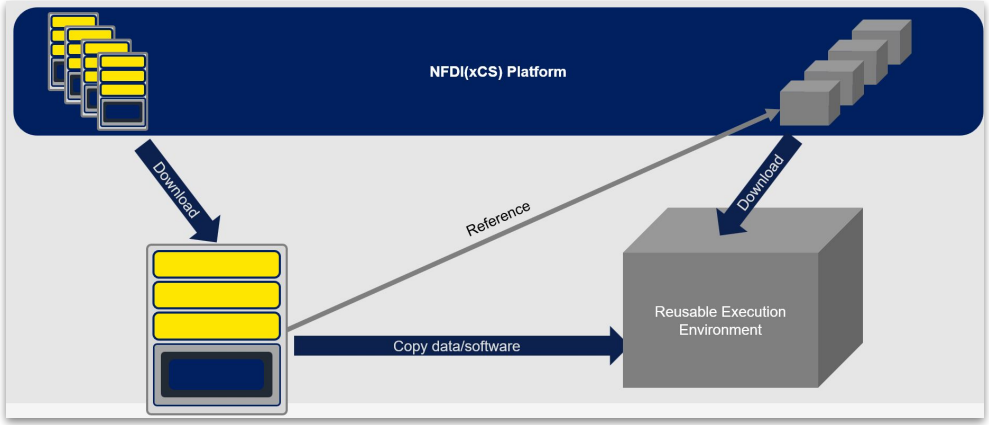
Need help in RSE in the NFDI4Culture domains? Contact the NFDI4Culture Helpdesk

RSE @ NFDIxCS

NFDIxCS

Information about legal aspects -> NFDIxCS.org/blog

Research Data Management Container and the Reusable Execution Environment



Coffeetablebook about Research Data and Software Managment -> https://nfdixcs.org/publications/book

RSE @ NFDI4Energy

nfdi4 energy

Main Aspects of Task Area 5 in NFDI4Energy

Improve Findability

Software Registry

- Links to implementations of simulation techniques
- Test cases and other resources
- Providing guidance services for quick access and suitability finding

Energy Simulation Software Ontology

- Structured overview of different modeling approaches and guide for researchers
- Allowing experts to add details for their specific areas of expertise

Improve Interoperability

Simulation-as-a-Service

- Distributed simulation for the combination of existing models and running a comprehensive simulation
- Providing easy access to simulation middleware that enables different types of distributed simulation

Distributed Simulation Frameworks

mosaik

VILLASframework

daceDS

Improve Reusability

Scenario Ontology

- Integration of semantics and domain knowledge in the process of planning, execution, and evaluation of simulations
- Integrate specialized hardware-in-the-loop (HIL) and laboratory testing in power system simulation scenarios

Information Model

- Formalizing relationships and properties of simulation models and components
- Including references to external model and component registries and the domain-specific ontology

RSE @ NFDI4Ing

NFDI4ING

RESEARCH DATA MANAGEMENT
IN ENGINEERING SCIENCES

"Hello, I'm Betty. I'm an engineer and self-taught programmer that develops research software. Very often, this software represents a computational model for the simulation of an engineering application. For validating such a model, I have to compare my results with data such as other simulation data or experimental observations. For this and other purposes, I also write code for analysing and converting research data. My software usually has a lot of dependencies in form of the operating system and third-party libraries. While I'm very keen on guaranteeing the reproducibility of my computational results, I can't dedicate too much working time to achieve this. My professional background can be located in any engineering discipline."

RSE @ NFDI4Objects

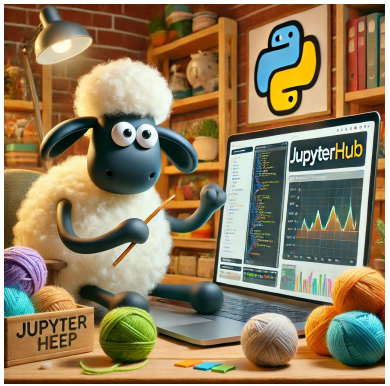
Florian Thiery CC BY 4.0, created using OpenAI's DALL·E

rse.i | FAIR4RS

Florian Thiery CC BY 4.0, created using OpenAI's DALL·E

rse.ii | n4o.software

Florian Thiery CC BY 4.0, created using OpenAI's DALL·E



rse.iii | jupyter4objects

NFDI4Objects

Community Cluster
Research Software Engineering (RSE)

Planned Temporary Working Groups |

FAIR4RS | n4o.software | jupyter4objects

Mailing-List |
https://www.listserv.dfn.de/sympa/info/n4o_cc_rse

Advices given in the survey

You are not alone!
Engage with the Research Software Community
- Join RSE societies (e.g., de-RSE)
- Contribute to open-source projects
- Attend conferences & workshops

Make your software visible and available to the community, as soon (but not earlier) as it is stable and useful

Everything you do is relevant to the community! Share it!

Study science deeply and broadly.

Stay tuned! Don't be discouraged!
Be FAIR(4RS)!

I'm new in the field myself. My advice would be: if you see a specific need or lack that many researchers share for tooling or a dataset or infrastructure, then fill this gap and publish it. If it does serve a real need, researchers will adopt it.

Software development in research differs from that in industry. It often (but not always) appears worse. Your goal is to identify what needs improvement and work to make it better.

Advices given in the survey

– Start to **do networking** as early as possible, and do not hesitate to talk to the perceived masters of our field out of fear.
– Count any **research about technologies**, learning of languages for programming or data encoding, careful planning of your work as time you **work**ed for your/their project, 100% on par with the time needed for actually implementing the thing.
– **Do never feel bad** if you take an hour to (re-)familiarize yourself with a detail of a specification, something you did in the past, or a **“silly” bug** which you think you should have found sooner.
– Do not ever start on anything without having some **reliable version control and backup plan** set up for it.
– Familiarize yourself with a **project management system** which suits your needs and you deem fun to use.
– Learn about **Evidence-Based Scheduling**.
– Read and impersonate **Joel Spolsky’s “10 Steps to Better Code”**. Actually, read a majority of the articles at joelonsoftware.com (and do not forget to count this as paid time spent for work).
– The more an **end user** sees of something, the longer it probably takes to implement and get it into a stable state.
– Never say, and never let anyone assume, that a specific part of your code will not need any **alteration or maintenance** in the future.
– Do not start to work on something if you do not get **a stable build** of it working on your own machine, setup and network intended for development, which is 100% separate from the build currently used in production.

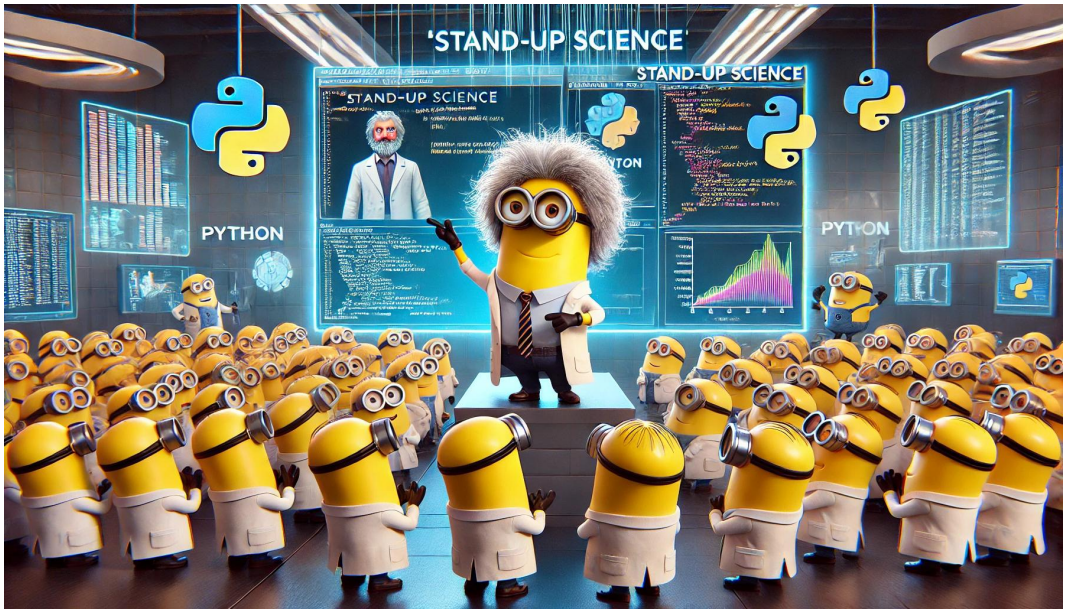
RSE Survey - Favorite git commands

bisect
commit
merge
pull
push
status

Let us do some *Stand-Up-Science*

Questions

1. Name / last academic degree / Affiliation / current task
2. What are your connections to the NFDI?
3. What was the last thing I coded/worked on?
4. Which research data did I last reuse?
5. Which research software did I last reuse?
6. Favourite git command?
7. Give a reason why you were told that the software/data could not be published …
8. What advice (best/good practice) would you give to a new RSE?



It is your turn! Do some *Stand-Up-Science*

Finis!
Thx :-)

Florian Thiery (CAA-DE e.V. & LEIZA)
Bernd Flemisch (University of Stuttgart)
Jan Bernoth (University of Potsdam)
Corinna Seiwerth (FAU Erlangen-Nürnberg)
Jan Linxweiler (TU Braunschweig)
Linnaea Söhn (Academy of Sciences and Literature | Mainz)