Services Roadshow by

Jupyter®4.nfdi

base4 nfdi

May 27, 2025

A central Jupyter Hub

<u>Arnim Bleier</u> and Björn Hagemeier

on behalf of the project team:
Bernd Flemisch, Matthias Lieber, Klaus Reuter, George Dogaru,
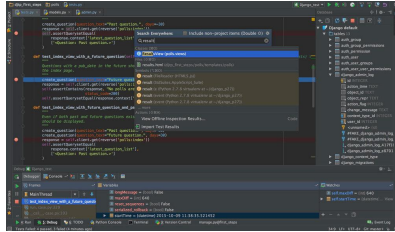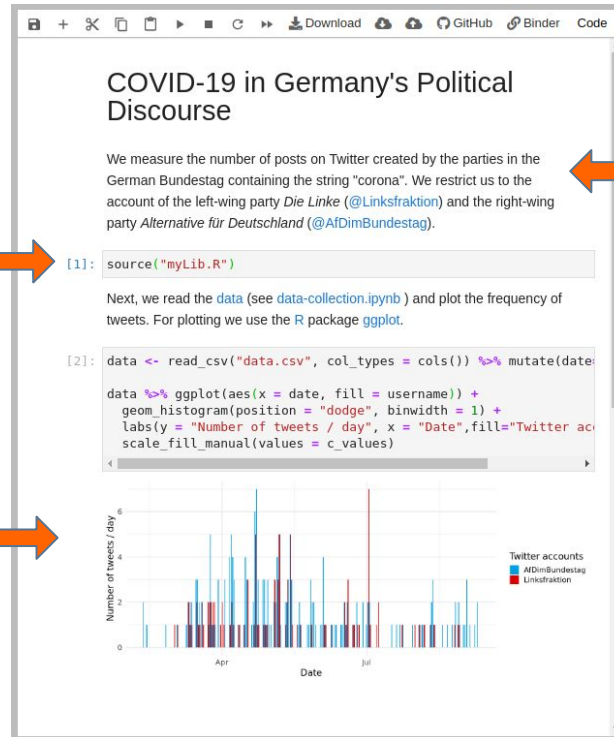Tim Kreuzer, Diego Jahn, Anett Seeland, Ines Ouled Elhaj, Julian Kohne

What are (Jupyter) Notebooks?
Technology overview

Source code

Visualization
(potentially interactive)

Natural language

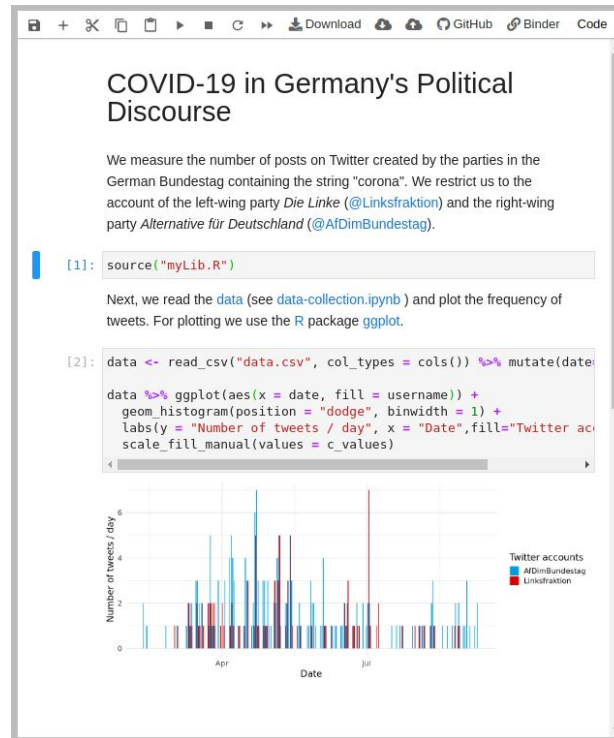**Examples:**
- ❏ **Jupyter**
- ❏ Quarto
- ❏ Pluto.jl
- ❏ ...

What are (Jupyter) Notebooks?
Computation

Cloud:
- ❏ potentially large Data
- ❏ standardized environment
- ❏ 1-Click reproducibility

COVID-19 in Germany's Political Discourse

We measure the number of posts on Twitter created by the parties in the German Bundestag containing the string "corona". We restrict us to the account of the left-wing party *Die Linke* (@Linksfraktion) and the right-wing party *Alternative für Deutschland* (@AfDimBundestag).

```
[1]: source("myLib.R")
```

Next, we read the data (see data-collection.ipynb) and plot the frequency of tweets. For plotting we use the R package ggplot.

```
[2]: data <- read_csv("data.csv", col_types = cols()) %>% mutate(date
data %>% ggplot(aes(x = date, fill = username)) +
  geom_histogram(position = "dodge", binwidth = 1) +
  labs(y = "Number of tweets / day", x = "Date",fill="Twitter ac
  scale_fill_manual(values = c_values)
```

Personal Computer:
- ❏ only small data
- ❏ every environment different
- ❏ time-consuming to set up

base4 nfdi

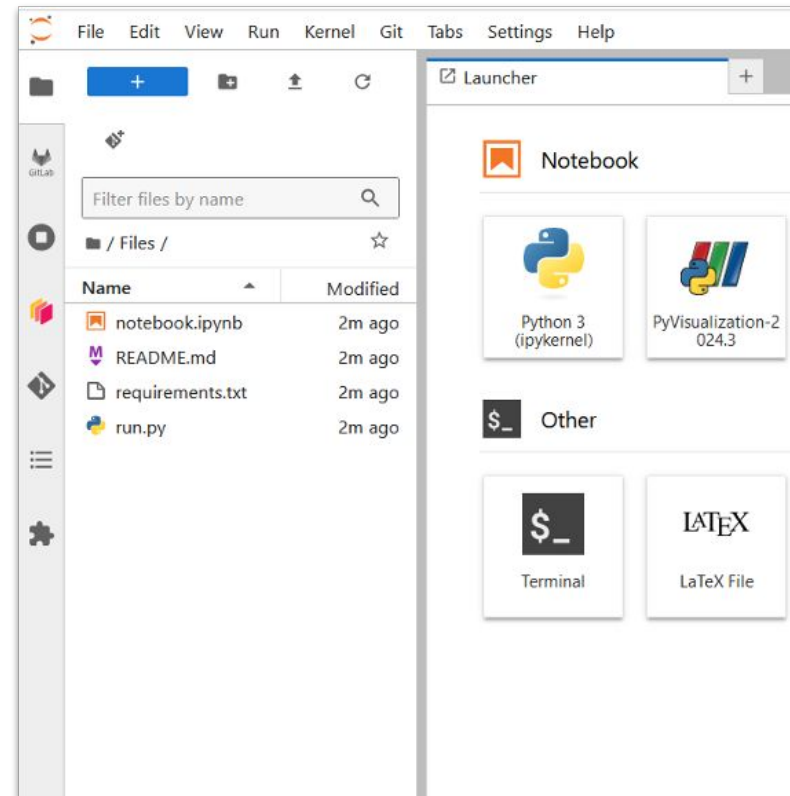# JupyterLab (IDE)
Components

Backend for notebooks
- executes code and serves the notebook interface in a browser

Web interface
- provides a browser-based web interface

JupyterLab
- Next-gen interface
- Extendible (terminal, text editor, file browser)
- Endless customization

Jupyter®4 nfdi

… is a **cloud-based computational analysis service** that enables researchers to focus on the methodological aspects of their work while increasing their reach and following best open-science practices at the same time.

The service does not try to reinvent the wheel, but reuses and integrate existing components whenever possible.

Why Jupyter4NFDI?
Heterogeneous Availability

Dozens of JupyterHubs in Germany

Challenges

- Many are not accessible to external users
- Which JupyterHubs are available to me?
- What are the differences?

Proposed solution

- Centralized "default" JupyterHub
- Combining resources of providers
- Not a replacement, but **complementary service** to enhance access and usability for all users
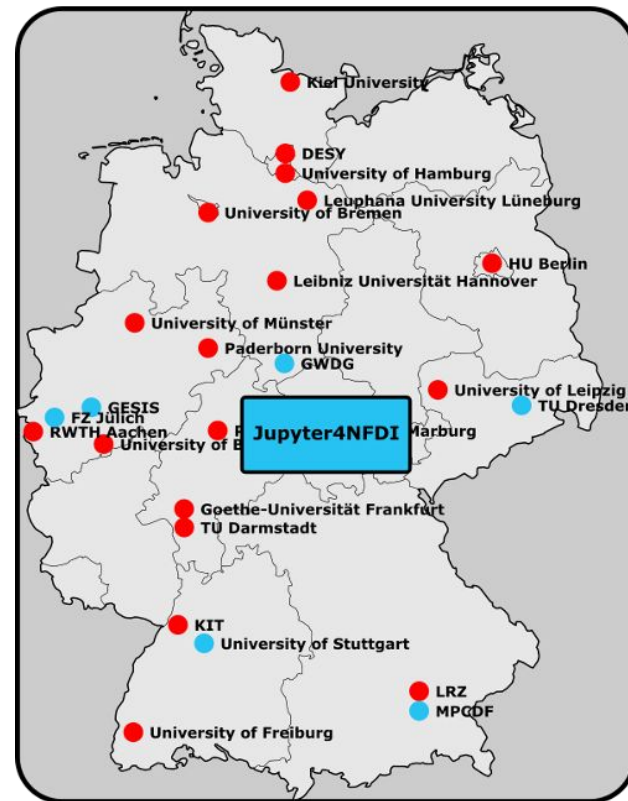
Why Jupyter4NFDI?
Heterogeneous Availability

Dozens of JupyterHubs in Germany

Challenges

- Many are not accessible to external users
- Which JupyterHubs are available to me?
- What are the differences?

Proposed solution

- Centralized "default" JupyterHub
- Combining resources of providers
- Not a replacement, but **complementary service** to enhance access and usability for all users

Why Jupyter4NFDI?
Benefits of the service

Easy access to resources
- Login with your Home IdP or social IdPs

Exclusive features
- Repo2Docker (Binder) integration into JupyterHub
- Shareable links

Highly customizable
- Run your own Docker images
- Create your environment

Persistent storage
- Varies by provider
- Centralized storage is planned for the future
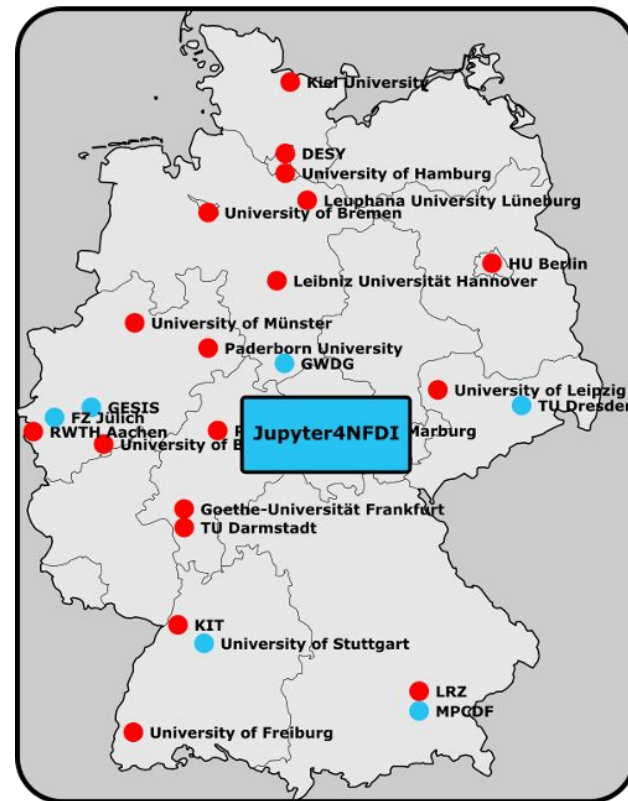
Jupyter4NFDI
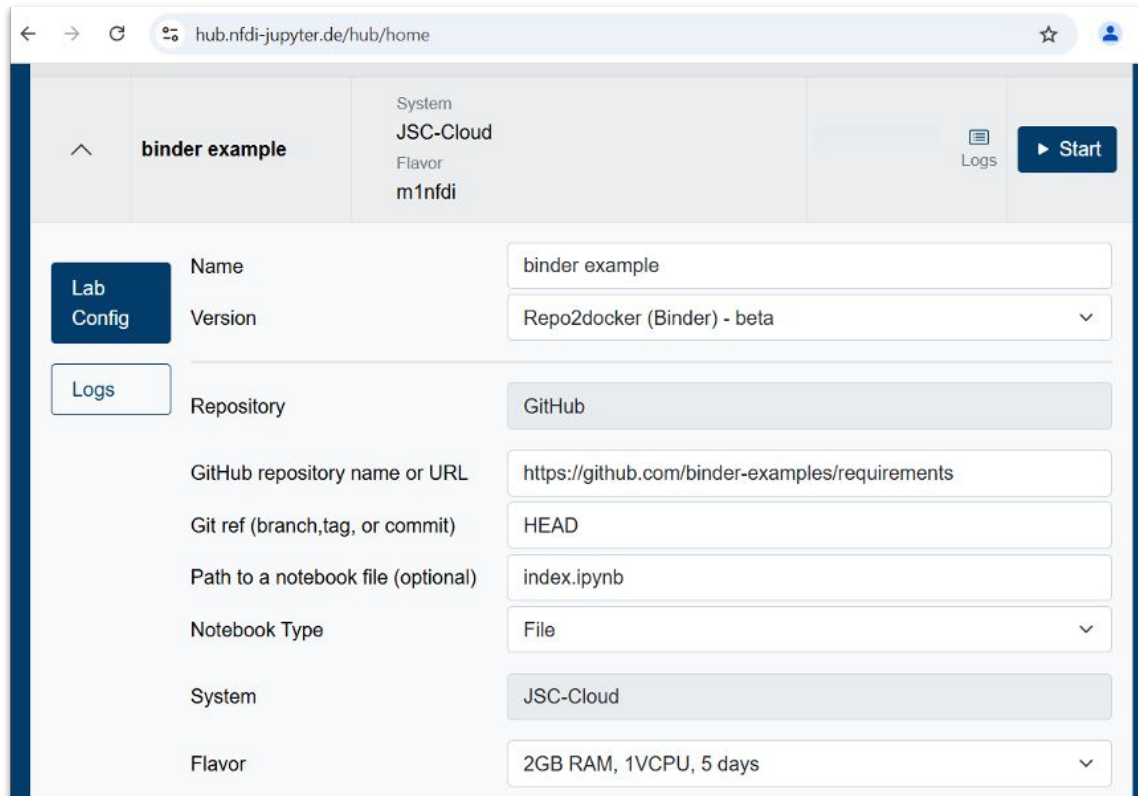*NFDI JupyterHub & Service Documentation*

Jupyter®4∩fdi

JupyterHub

🌐 https://hub.nfdi-jupyter.de

1-Click Import

launch ∩fdi

🌐 https://hub.nfdi-jupyter.de/v2/gh/gesisc
ss/Jupyter4NFDI_survey_results/HEAD

base4
∩fdi

Jupyter4NFDI
Technical concept

- The **central JupyterHub** acts as an **access point** to multiple resources in the background

- Resource providers can install a "JupyterHub Outpost" to **integrate their resources with the central JupyterHub** following general standards

- Throughout the project, additional resource providers will be integrated into the central service, expanding the available **resources, environments, features, and capabilities**

Partners and Resources:

- Text+ (FZ Jülich, GWDG)
- NFDI4DS (GESIS, TU Dresden, Univ. of Stuttgart)
- FAIRmat (MPCDF)
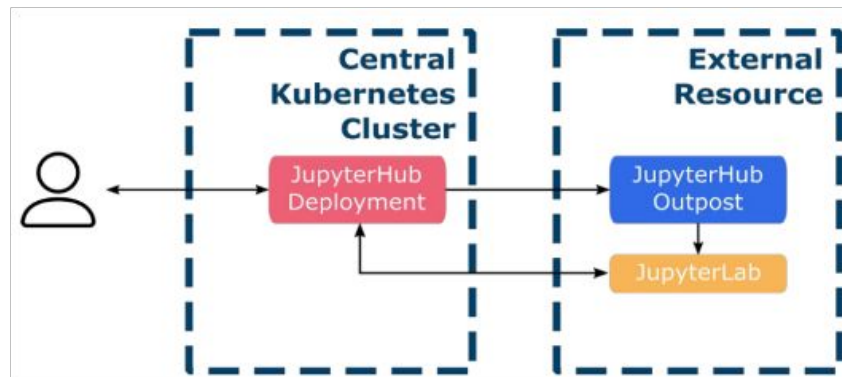
Jupyter4NFDI
Outlook on Integration Phase

Potential start:
October 2025

- **Onboarding of Content**: Content from Base services of other consortia, as well as from individual content producers and researchers.

- **Verifiable Reproducibility**: Based on FAIR Jupyter, we will share a knowledge graph for computational notebook reproducibility.

- **Training**: We will provide reusable training materials for both regular users and researchers, aiming to make their computational workflows Jupyter4NFDI-ready.

Initialization   Integration   Ramping Up
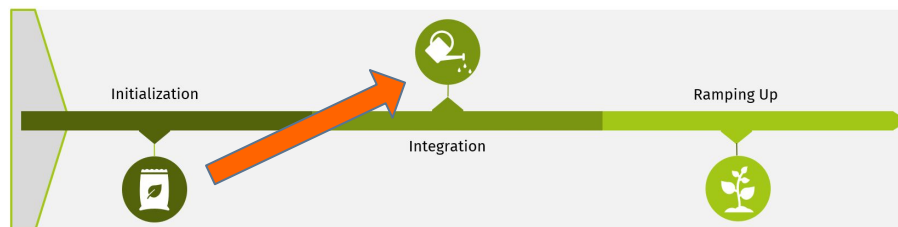
Quiz: Which benefits does Jupyter4NFDI offer their users?

A) Run your own Docker images
B) Import binder-ready analysis pipelines
C) Get central access to multiple resources via a browser-based interface

Join us during our **Open Hours**
https://events.hifis.net/event/2625/

Thank you!
Questions?

Jupyter®4.nfdi

✉ **jupyter4nfdi@lists.nfdi.de**

✉ base4nfdi-servicestewards@lists.nfdi.de for general inquiries

🌐 **base4nfdi.de/projects/jupyter4nfdi**