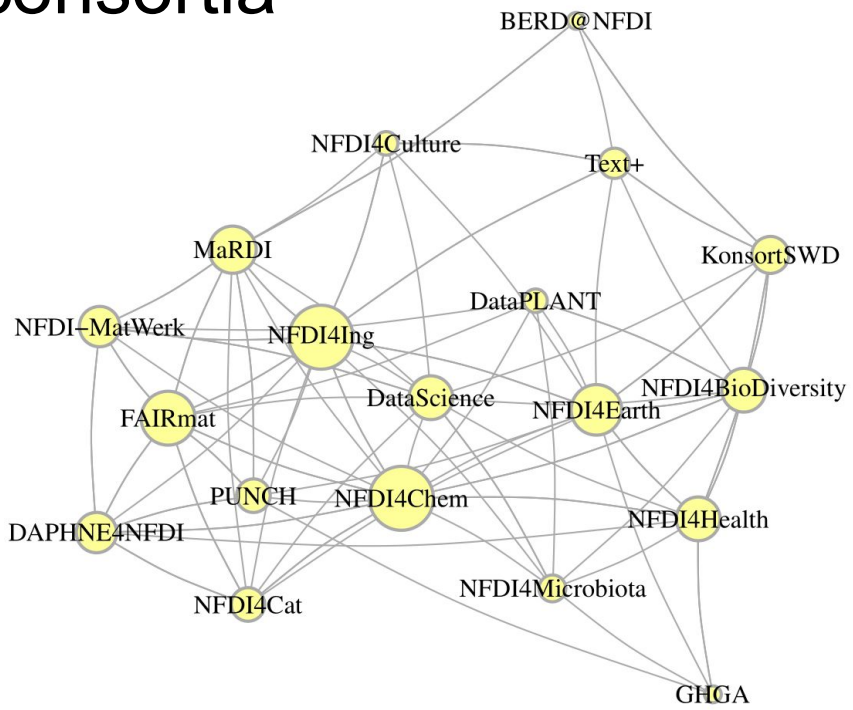
# NFDI’s promise to partner

## Hands on: Analyzing and visualizing connections between (planned) NFDI consortia


| | | |
|---|---|---|
| *Sama Majidian* | *RWTH Aachen University* | https://orcid.org/0000-0002-7428-6713 |
| *Lukas C. Bossert* | *RWTH Aachen University* | https://orcid.org/0000-0003-3076-3968 |
| *Évariste Demandt* | *RWTH Aachen University* | https://orcid.org/0000-0002-5705-0071 |

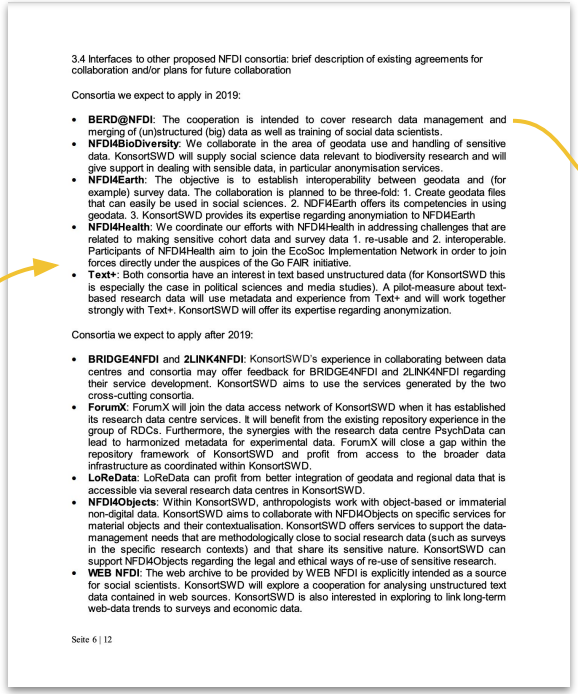
# Connections between NFDI consortia




Funded:
9+10 nodes

# Data collection from published binding letters of intent

check out https://github.com/dorothearrr/NFDI_Netzwerk

3.4 Interfaces to other proposed NFDI consortia: brief description of existing agreements for collaboration and/or plans for future collaboration

Consortia we expect to apply in 2019:

- **BERD@NFDI**: The cooperation is intended to cover research data management and merging of (un)structured (big) data as well as training of social data scientists.
- **NFDI4BioDiversity**: We collaborate in the area of geodata use and handling of sensitive data. KonsortSWD will supply social science data relevant to biodiversity research and will give support in dealing with sensible data, in particular anonymisation services.
- **NFDI4Earth**: The objective is to establish interoperability between geodata and (for example) survey data. The collaboration is planned to be three-fold: 1. Create geodata files that can easily be used in social sciences. 2. NDFI4Earth offers its competencies in using geodata. 3. KonsortSWD provides its expertise regarding anonymiation to NFDI4Earth
- **NFDI4Health**: We coordinate our efforts with NFDI4Health in addressing challenges that are related to making sensitive cohort data and survey data 1. re-usable and 2. interoperable. Participants of NFDI4Health aim to join the EcoSoc Implementation Network in order to join forces directly under the auspices of the Go FAIR initiative.
- **Text+**: Both consortia have an interest in text based unstructured data (for KonsortSWD this is especially the case in political sciences and media studies). A pilot-measure about text-based research data will use metadata and experience from Text+ and will work together strongly with Text+. KonsortSWD will offer its expertise regarding anonymization.

Consortia we expect to apply after 2019:

- **BRIDGE4NFDI** and **2LINK4NFDI**: KonsortSWD's experience in collaborating between data centres and consortia may offer feedback for BRIDGE4NFDI and 2LINK4NFDI regarding their service development. KonsortSWD aims to use the services generated by the two cross-cutting consortia.
- **ForumX**: ForumX will join the data access network of KonsortSWD when it has established its research data centre services. It will benefit from the existing repository experience in the group of RDCs. Furthermore, the synergies with the research data centre PsychData can lead to harmonized metadata for experimental data. ForumX will close a gap within the repository framework of KonsortSWD and profit from access to the broader data infrastructure as coordinated within KonsortSWD.
- **LoReData**: LoReData can profit from better integration of geodata and regional data that is accessible via several research data centres in KonsortSWD.
- **NFDI4Objects**: Within KonsortSWD, anthropologists work with object-based or immaterial non-digital data. KonsortSWD aims to collaborate with NFDI4Objects on specific services for material objects and their contextualisation. KonsortSWD offers services to support the data-management needs that are methodologically close to social research data (such as surveys in the specific research contexts) and that share its sensitive nature. KonsortSWD can support NFDI4Objects regarding the legal and ethical ways of re-use of sensitive research.
- **WEB NFDI**: The web archive to be provided by WEB NFDI is explicitly intended as a source for social scientists. KonsortSWD will explore a cooperation for analysing unstructured text data contained in web sources. KonsortSWD is also interested in exploring to link long-term web-data trends to surveys and economic data.

Seite 6 | 12

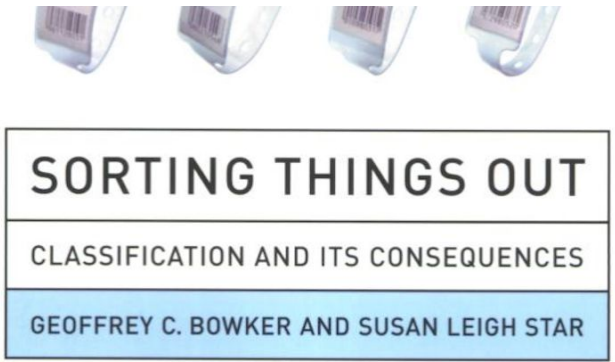
| from | to |
|---|---|
| KonsortSWD | BERD@NFDI |
| KonsortSWD | NFDI4BioDiversity |
| KonsortSWD | NFDI4Earth |
| … | … |

# A few quotes on infrastructure:

Bowker, G. C., Baker, K., Millerand, F., & Ribes, D., 2010:
„Beyond bricks, mortar, pipes or wires, infrastructure also encompasses more abstract entities, such as protocols (human and computer), standards, and memory.“ (p. 97)

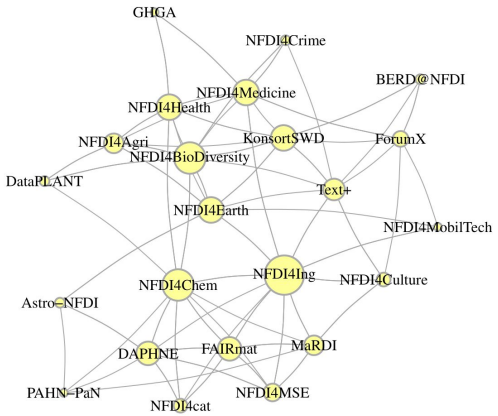
Ibid.: „... we take infrastructure as a broad category referring to *pervasive enabling resources in network form*, and we argue that a theoretical understanding of infrastructure is crucial to its design, use, and maintenance.“ (Emphasis added by the authors, p. 98)

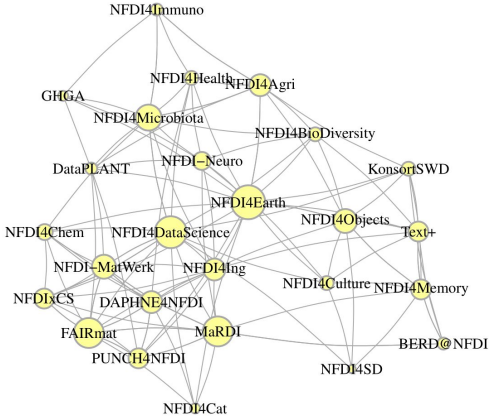
Geoffrey C. Bowker et Susan Leigh Star, 1999:
„Infrastructure does more than make work easier, faster or, more efficient; it changes the very nature of what is understood by work.“ (p. 108)

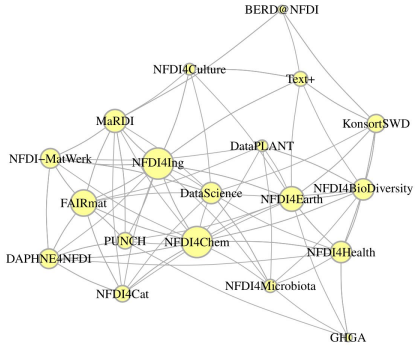



https://voxeu.org/content/what-counts

# Visualization of consortia: applied vs. funded



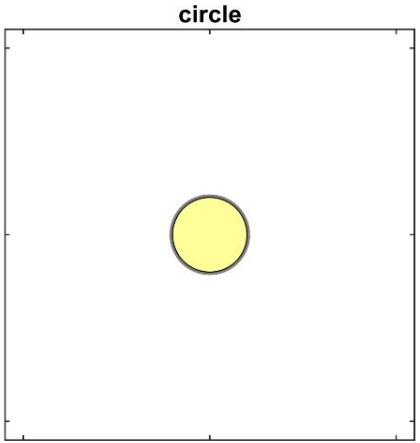



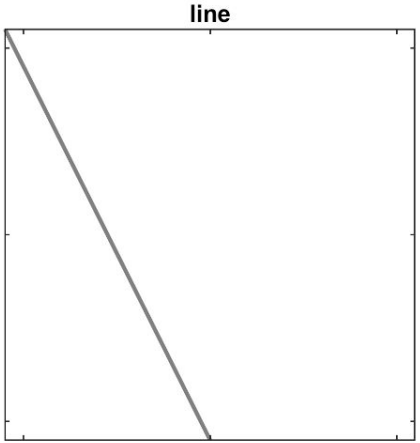

2019:
23 nodes

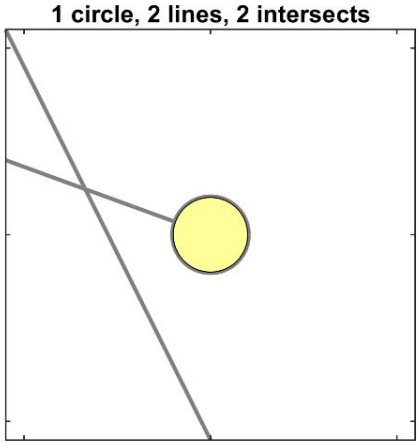
2020:
9+17 nodes

Funded:
9+10 nodes

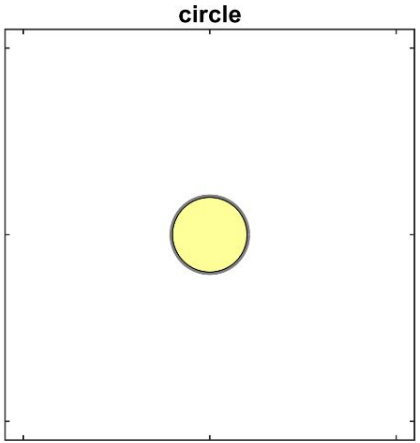
# Visual components (building blocks) of networks

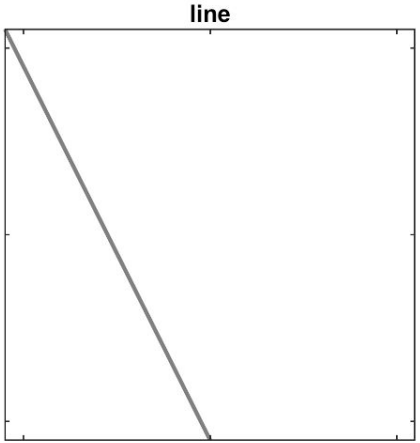


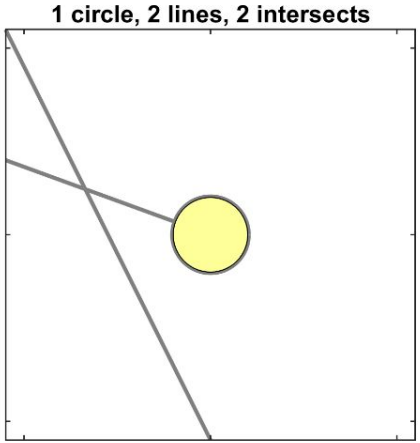



# Visual components (retinal variables) of networks




circle = simple cell (“bug detector”), Frog retina

line = complex cell, Cat visual cortex (V1),

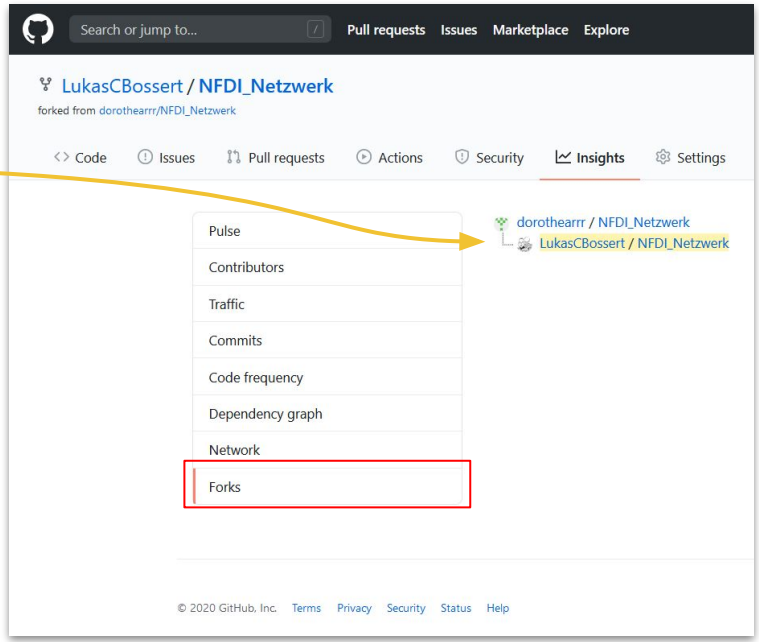
intersect = corner cell Cat visual cortex (V2)

# Network analysis – sharing is important :)

- A GitHub repository, where it all began: https://github.com/dorothearrr/NFDI_Netzwerk
- The Publication of software and data
- Description of the data in a README file
- Indication of sources and software used
- Version control (the repository was updated with the release of the binding LoIs for the second round of applications)
- Licensing with a "permissive software license" as a signal to potential downstream users

# Network analysis: Re-use

- Data can not only be downloaded, but also "forked" (e.g. in a new public repository)
- The data origin remains traceable
- Further development is possible
- (New) data is fed back into the community



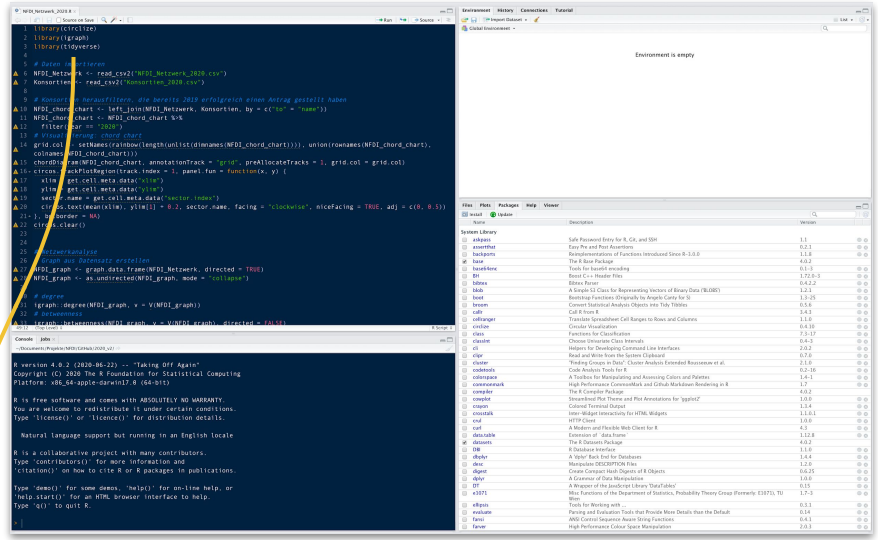
# Good Scientific Practice and Software: Disclaimers

- We should always ponder that “in many ways software is frozen organisational and policy discourse“. (Geoffrey C. Bowker et Susan Leigh Star, 1999, p. 135)
- We should bear in mind that GitHub is not sufficient when it comes to depositing or deep-freezing software in a "persistent, citable and documented" manner. (Deutsche Forschungsgemeinschaft, 2019, p. 14)
- We shouldn’t be oblivious to the observation that dealing with research software comes with many challenges that also need to be discussed in relation to the German science system. (cf. Anzt et al., 2020)

# Network analysis – using R with the library igraph

In preparation of this workshop, the analyses and visualizations were performed using the statistical software R, the programming environment RStudio and the library igraph (all freely available).

Check out the carpentry R for Reproducible Scientific Analysis !



igraph – The network analysis package

igraph is a collection of network analysis tools with the emphasis on **efficiency**, **portability** and ease of use. igraph is **open source** and free. igraph can be programmed in **R**, **Python**, **Mathematica** and **C/C++**.

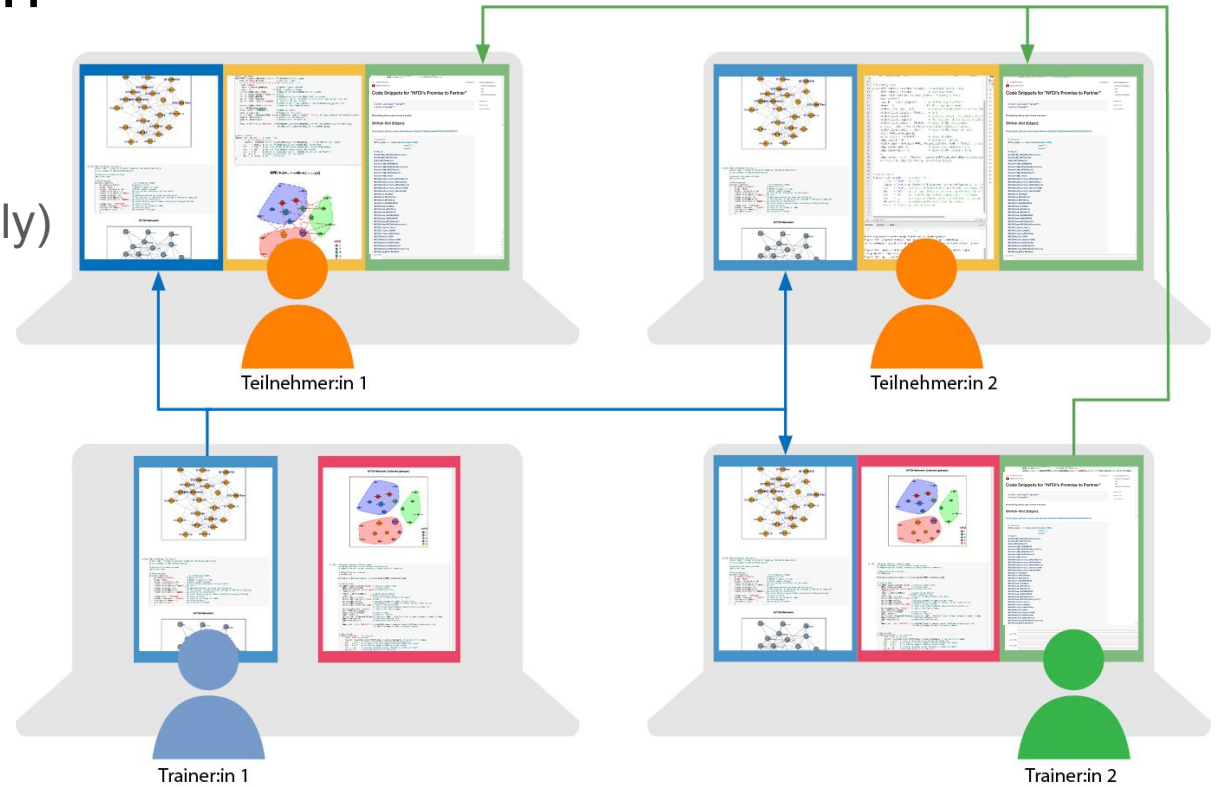
# Hands-On-Session

Red (master)

Blue (video stream/ read only)

Green (shared/ copy paste)

Orange (your own code)

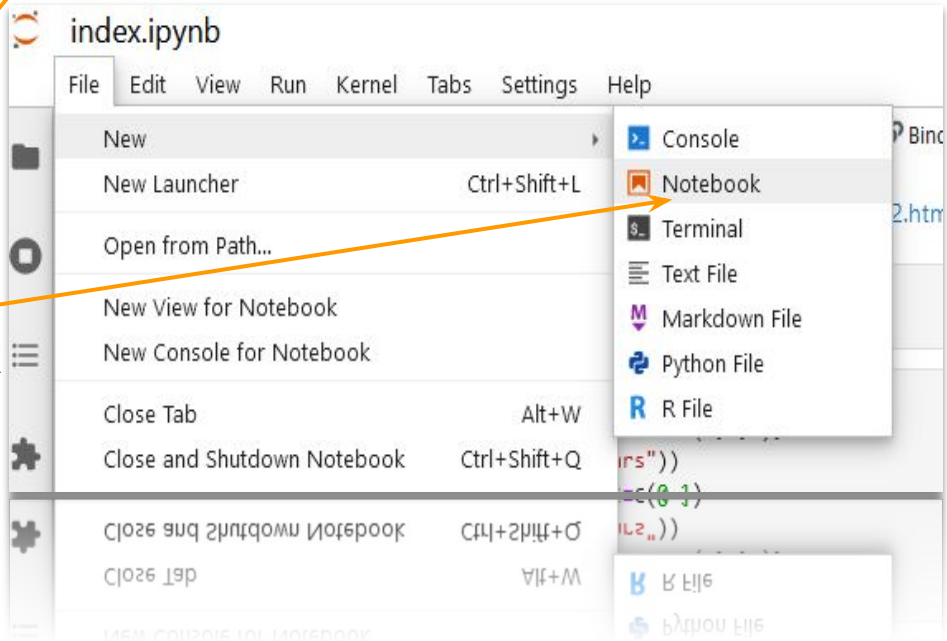


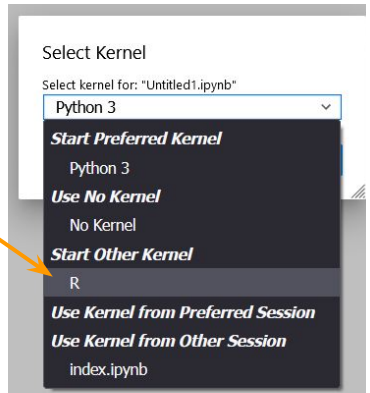

Adapted from “Das Versprechen der Vernetzung der NFDI”, Bausteine FDM (upcoming)

# Hands-On-Session (option 1)

We work with R Notebooks (aka Jupyter Notebooks). Get yours in just 4 steps:

1. Visit https://jupyter.org/try
2. Click “Try JupyterLab”
3. Click File and create empty notebook
4. Select the kernel “R”

# Hands-On-Session (option 2)

We work with R Notebooks (aka Jupyter Notebooks).
Get yours in just 3 steps:

1. Visit https://jupyter.org/try
2. Click “Try JupyterLab”
3. Click R under the section for Notebook