National Research Data Infrastructure for the Neurosciences

Ad Aertsen\textsuperscript{1}, Torsten Bronger\textsuperscript{2}, Michael Denker\textsuperscript{2}, Sonja Grün\textsuperscript{2}, Michael Hanke\textsuperscript{2}, Jan Klein\textsuperscript{3}, Steffen Oeltze-Jafra\textsuperscript{4}, Frank Ohl\textsuperscript{4}, Guido Prause\textsuperscript{3}, Stefan Rotter\textsuperscript{1}, Hansjörg Scherberger\textsuperscript{5}, Jan Schreiber\textsuperscript{2}, Constanze Seidenbecher\textsuperscript{4}, Oliver Speck\textsuperscript{4}, Alexandra Stein\textsuperscript{6}, Thomas Wachtler\textsuperscript{7}

\textsuperscript{1}Bernstein Center Freiburg; \textsuperscript{2}Research Center Jülich; \textsuperscript{3}Fraunhofer Institute for Digital Medicine; \textsuperscript{4}Center for Behavioral and Brain Sciences, Magdeburg; \textsuperscript{5}German Primate Center, Göttingen; \textsuperscript{6}Bernstein Coordination Site, Freiburg; \textsuperscript{7}German Neuroinformatics Node, LMU Munich
National Research Data Infrastructure - Background

- 2016-2018: RfII publishes three documents on establishing a national research data infrastructure (Nationale Forschungsdateninfrastruktur, NFDI)
- November 2018: The Joint Science Conference (Gemeinsame Wissenschaftskonferenz, GWK) decides on funding program, mandates DFG to organize the process
  3 rounds of calls (2019, 2020, 2021)
  Overall volume up to 85 M€/year for 10 years
- May 2019: First NFDI Conference
- June 2019: DFG issues first call
- July 2019: Letters of Intent

http://www.dfg.de/nfdi
The NFDI is intended

- to be a **structure** / a process, not a funding program
- to be connected and **cooperative**, not competitive
- to be a coordinated **network of consortia** tasked with providing **science-driven data services** to research communities
- to establish close **collaboration** between users and information infrastructures
- to address the **needs of researchers**
- to **facilitate** access to data
- to foster sustainability by enabling reliable RDM services tailored to the needs of researchers

after Allgöwer, Gehring
NFDI Conference 13.05.2019
http://www.dfg.de/nfdi/
Research Data Management in Neuroscience

- High diversity of experimental approaches, model systems, methods, signal types, acquisition systems, data formats, analysis methods, ...

- Rapid increase of volume and complexity of experimental data: High-dimensional, heterogeneous data streams and complex analysis workflows require efficient data management and new concepts for analysis

- Lack of standardization

- Need for innovative concepts for research data management that address established lab data workflows

Complex data and metadata flow in an electrophysiology setup modified from Zehl et al 2016
Research Data Management in Neuroscience

- Community efforts to improve research data management have started out of neuroinformatics initiatives.
- Solutions are emerging in various areas including data formats, metadata, analysis and visualization, data sharing (e.g., NWB, NIX, odML, Neo, Elephant, DataLad, GIN), particularly in the field of neuroimaging (e.g. BIDS, NIDM, XNAT, MRIQC, MeVisLab, Slicer, OpenNeuro).
- Need for further development, integration and support of community adoption.
NFDI Neuroscience - Concept and Approach

- Coordinated **bottom-up process** to empower scientists with skills and tools for competent research data management
- Development of **practical solutions** to facilitate data management in the labs
- Community-driven process supported by
  - coordinating **expert teams**
  - **working groups** of users and providers
  - flexible **funding** to develop solutions for identified needs
- Connected with **international** neuroinformatics initiatives and neuroscience infrastructures
- Teaching and **training**
NFDI Neuroscience Community

- Community network of scientific labs, institutions and infrastructure providers
- Building on established communities (Bernstein Network, NWG)
- Strong links to international neuroinformatics and neuroscience initiatives (INCF, HBP, BRAIN Initiative, ...)

www.nfdi-neuro.de
Domain-specific Infrastructure: SimLab Neuroscience

**Mission:** To enable use of supercomputers for neuroscience

- Making software and applications ready for HPC
- Managing data in HPC-based simulation/analysis workflows
- Assist scientists in writing HPC applications
Domain-specific Infrastructure: G-Node

Development of tools and services for research data management in neuroscience, facilitating data access, data analysis and data sharing

- Methods and tools for data and metadata management
- Services supporting collaboration and reproducibility
- Data publication services
- Hosting services
- Teaching and training

www.g-node.org
International Initiatives

- **EOSC**: European Open Science Cloud, [www.eosc-portal.eu](http://www.eosc-portal.eu)
  Integrating European infrastructures (EUDAT, OpenAIRE, EGI, PRACE, ...), funded by EC through H2020

- The **FAIR** Principles, [www.force11.org](http://www.force11.org)
  Guidelines to make research data *Findable, Accessible, Interoperable, Reusable*
  adopted by many stakeholders incl. H2020

Neuroscience:

- **INCF**, [www.incf.org](http://www.incf.org)
  International organization to foster neuroinformatics developments, standardization and training

- **HBP**, [www.humanbrainproject.eu](http://www.humanbrainproject.eu)
  European Flagship project building research infrastructure for neuroscience