

Establishing Sustainable Research Data Management in the Neurosciences: The RETAIN Fellowship Program

René Bernard, PhD
Cluster of Excellence NeuroCure
BIH QUEST Center for Responsible Research
rene.bernard@charite.de

15.10.2024 FDM Netzwerktreffen Berlin Brandenburg

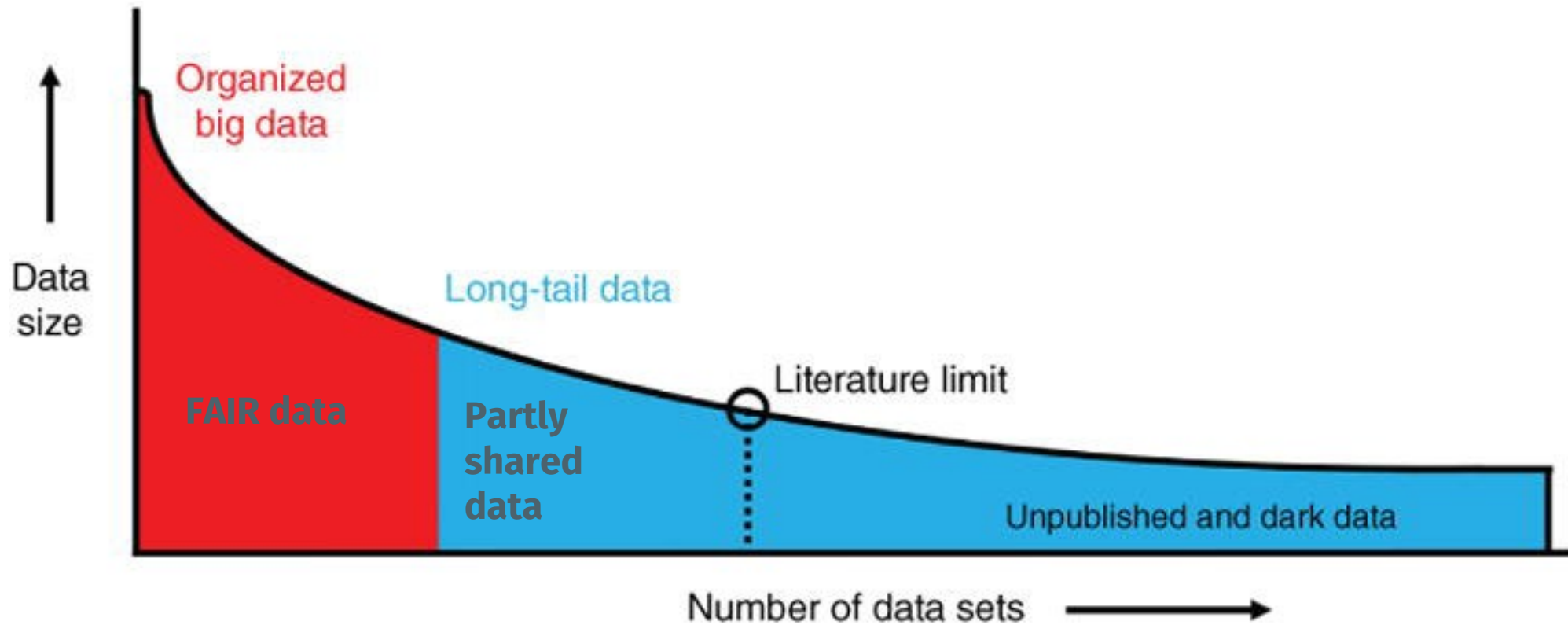
 **BIH QUEST**
Center for Responsible Research



NEUROCURE
Cluster of Excellence

 **BIH** Berlin Institute
of Health
@Charité

Datafication in Biomedicine (applicable to Neuroscience)



Complexity in the Neurosciences – Subfields & Subdisciplines

- Computational
- Behavioral
- Molecular
- Developmental
- Clinical
- Cellular
- Sensory
- Systems
- Social
- Evolutionary

Many interdisciplinary approaches and techniques in combination

Many data have a spacial and temporal component

- Neurogenetics
- Neuropharmacology
- Neurophysiology
- Neuroanatomy
- Neuroendocrinology
- Neuroinformatics
- Neuropsychology

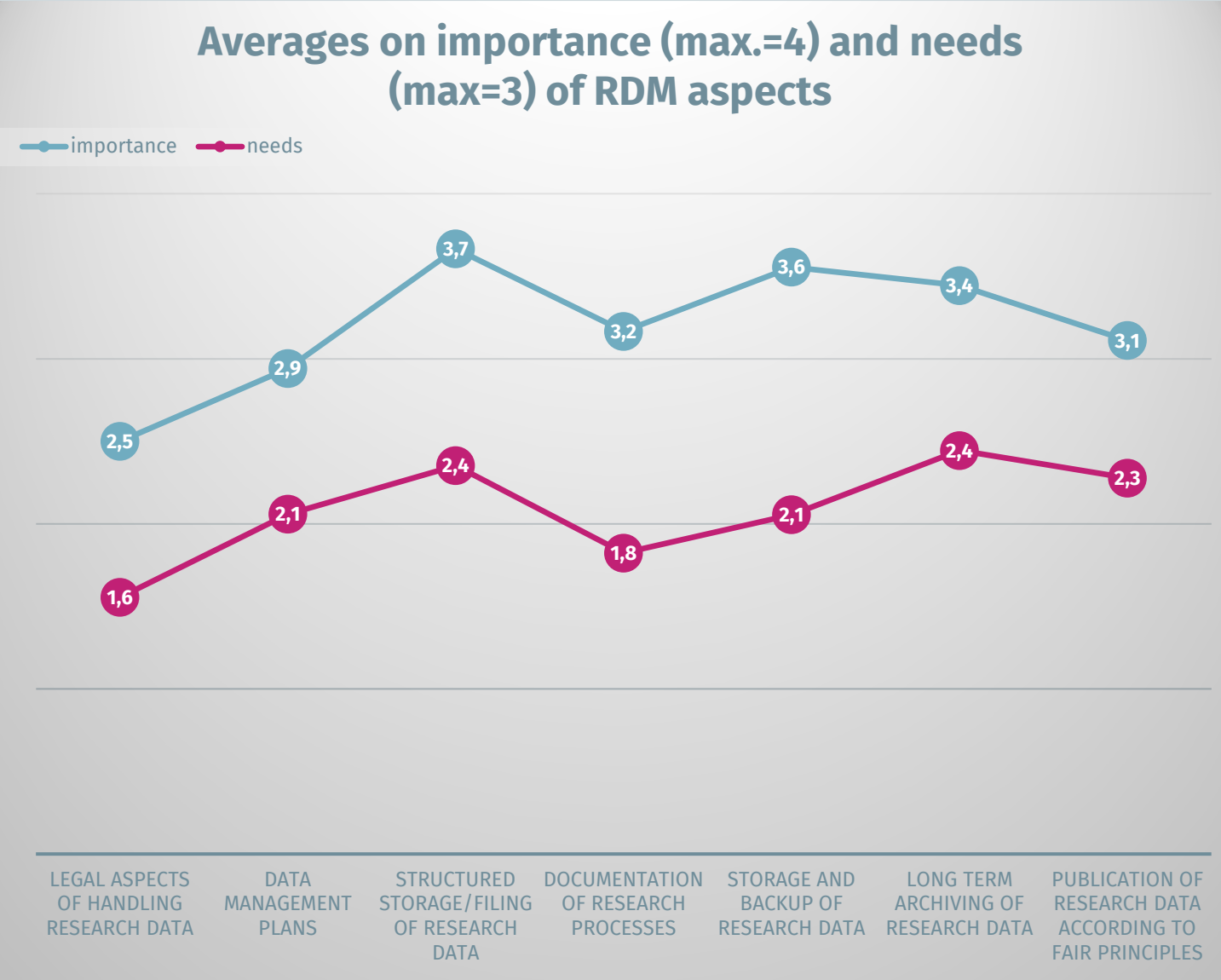
This complexity is a big hurdle in effective RDM

Research Data Management Questionnaire among NeuroCure PIs and Group Leaders

Researcher's perspective:

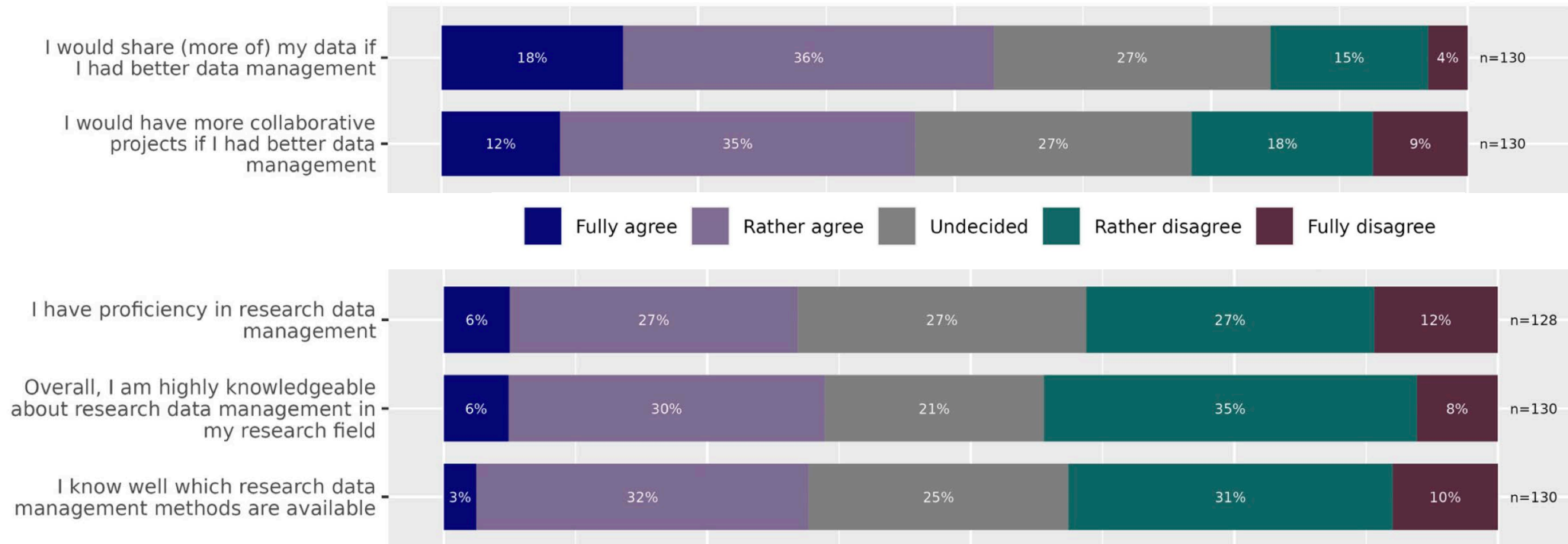
- **Structured storage and filing**
- **Long term archiving and access**

are most important –
reflect the use and
access to data



Research Data Management and Data Sharing for Reproducible Research—Results of a Community Survey of the German National Research Data Infrastructure Initiative Neuroscience

What is your opinion on the following statements?



Awareness of potential benefits, yet there is a lack of knowledge and best practices

Neuroscientists need support to build a lasting RD infrastructure

From the abstract:

and opinions about RDM. The German neuroscience community perceives barriers with respect to RDM and data sharing mainly linked to (1) lack of data and metadata standards, (2) lack of community adopted provenance tracking methods, (3) lack of secure and privacy preserving research infrastructure for sensitive data, (4) lack of RDM literacy, and (5) lack of resources (time, personnel, money) for proper RDM. However, an overwhelming majority of community members (91%) indicated that they would be willing to share their data with other researchers and are interested to increase their RDM skills. Taking advantage of this willingness

- **NFDI did not fund the NFDI-Neuro consortium**
- **NFDI4health, NFDI4bioimaging, GHGA**
- **Omics researcher have established international standards and databases**
- **What can we do as Cluster of Excellence ?**

VOS coordinator actions can help but are limited

- Counseling for DMPs and funding applications
- Course (ReproducibiliTeach) that teaches basic RDM best practices, data dictionaries, and data repository uploads
- But: Review and counseling from within almost impossible + specialized knowledge is required
- There is a clear limit what an outside person is able to do → true changes in RDM infrastructure must be initiated from within a research environment

Overview of RETAIN Program (Research Data Management Implementation in the Neurosciences)

- Goal: Develop a sustainable RDM infrastructure in a neuroscientific research environment
- Who: last PhD student / postdoc student receives a fellowship for 2 years that covers 50% of salary
- Applicant shows interest in RDM, data literacy, and willingness to pursue further specialized education
- Written proposals require statement of support from the PI for RDM implementation
- Fellowship program is overseen by VOS Coordinator (incl. progress tracking)
- Proposals and Progress Reports to the NeuroCure Board of Directors

Programm Benefits

- VOS coordinator guides and consults fellows
- Provides sufficient time to develop and implement best-fitting solutions that work for all
- Fellows expertise and skills qualify them for data manager and data steward positions in the future
- Show-case labs serve as blueprints for other NeuroCure groups and beyond
- RETAIN is a proof-of-concept program but if successful – it is adaptable and scalable

RETAIN application package

1. Description of RD handling in the research environment (lab standards, conventions, organizational structures for storage, archiving; workflow, if they exist)
2. Table with : data types, metadata, data processing, data archiving, use of ELN, data access for collaborators, data sharing, legal obligations
3. If you were to re-use data created by others, which properties and metadata should a dataset have to qualify as reusable?
4. Description of the ideal RD solution achievable in the research environment
5. For applications information: general project outline, workplan, and resources
6. Commitment of support by applicants PI
7. CV of applicant

Timeline of RETAIN fellowships

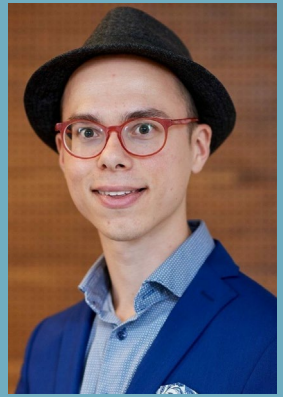
- **Call for applications: 1.6.2023 – 31.7.2023**
- **Review & Evaluations: 15.8.23**
- **Proposal Presentations to BOD: 30.9.23**
- **Start of fellowships: 1.1.24**
- **VOS Progress Meetings: 25.4 and 25.9.24**
- **BOD progress report: 30.11.24**
- **End of fellowships: 31.12.25**
- **Evaluations and follow up monitoring in 2026**

RETAIN fellows



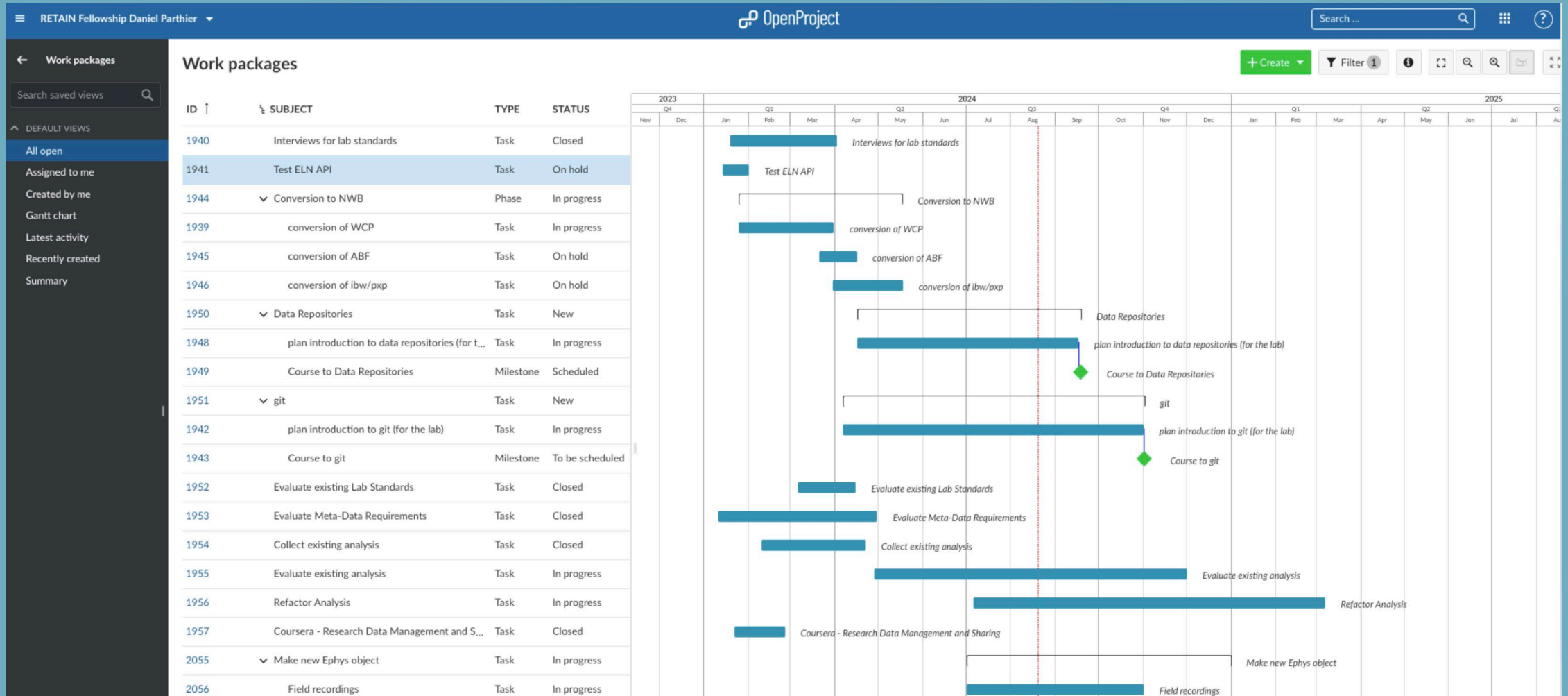
- Daniel Parthier, PhD in the lab of Dietmar Schmitz (preclinical)- Charité
- Electrophysiology, imaging and behavior
- 3 recording setups (different makers, different software, different output formats) – get them converted to open NWB format (neurodata without borders)
- Synchronized behavioral data and electrophysiological recordings
- Setup of “Data onboarding for new lab members”
- RETAIN project poster: <https://doi.org/10.5281/zenodo.13772103>

RETAIN fellows



- Jojo Vanhöcke, PhD student in the lab of Andrea Kühn (clinical)- Charité
- Deep brain stimulation device in patients with Parkinson
- human clinical data (e.g. behavioral, clinical, and routine clinical care data), sometimes "event data" or continuous (home) recordings
- Construct a data container analogous to the Brain Imaging Data Structure (BIDS) standard
- Share metadata with publications and speed up data transfer agreements by prepared templates
- DMP for most common recording setups

OpenProject for low-threshold RETAIN progress tracking



OpenProject for low-threshold progress tracking

- Fellow break up their projects into phases, tasks, and milestones
- Fellows' first exposure to Project Management Software
- Transparency increased; potential problems detected
- Read access: VOS coordinator, other fellow
- Encourage fellows to update OpenProject → email notification
- Questions/clarifications: MS Teams
- Joint progress meetings every 5 months

SUMMARY

- **Fellowships for RDM implementation have the potential to generate true progress towards shareable FAIR data**
- **Sharing of the developed solutions can be to adopted by other Neuroscience community members**
- **Generate domain-data-experts that can help others with similar setups, while empowering ECRs**
- **RETAIN concepts needs refinement but overall could be a funding mechanism that can have an impact in the Neuroscience community and make sharing reusable research data the norm**
- **If NeuroCure is continued to be funded: more RETAIN fellowships**