4th NFDI Neuroscience Community Workshop –
Shaping Research Data Management for the Neurosciences Together
July 1, 2020

Notes from the Discussion Sessions

General session

- How can we find ways to motivate the community to use RDM tools and share data and reward those in community approaches? Which incentives do we need?
    - Best incentive: scientists are more effective with proper RDM.
    - Realize that sharing data makes research easier
    - Automatized ways of data handling are required
    - Existing repositories are too complicated to use
    - Incentives for data publication needed
    - Rewards necessary for PhD students who improve RDM
    - with good training students actually like data management and apply it readily
    - the first round of students will have a hard time establishing proper RDM, but the second round of students can already benefit from the work.
    - Instead of creating incentives RDM should be made a natural component of reearch and so easy that it does not cost much extra effort

TA Common Infrastructure

- How can we make sure people can use datalad?
    - Datalad too complicated for use by not-nerds? How make it more accessible?
    - simple, single and specific user interfaces for specific user groups are needed
    - individual solutions can be built on top of existing structure
    - Same problems, switch to datalad, but needs a front-end for users
- Need for adapters to file systems for data that cannot be moved away
    - Standardized annotations are lacking
    - a web interface for good metadata is needed
    - Start with little, but be ready to become better
    - Datalad should be interoperable with as many infrastructures used in Germany as possible, LRZ will be a participant of NFDI Neuroscience in doing a datalad integration
- Market place idea is very much liked
    - NFDI is expected to coordinate and provide the platforms, a kind of gateway to all individual archives within the specific research domains
- Other tools than datalad?
    - Of course, Datalad is a connector between other tools: Independently extensible, as decentral as possible
- There is a gap between developers, analysts and people in a lab

- The collaboration aspect in development is the proper approach; a common layer for collaborative development needs to be established via NFDI Neuro and made known to the experts within individual labs so they can participate and benefit.
- Infrastructure is a wide field and should not just be Datalad. The core purpose is not to consolidate everything into a single tool but to connect existing tools. The underlying infrastructure should be and stay independently extensible and as de-central as possible.

TA Neuroimaging

- There are too many island solutions at lab level
- Not only think at individual lab level but also on a broader, institutional level
- The OMERO platform can be used individually at smaller labs as well as in larger institutions, however, the setup will be different: this should be kept in mind when setting up also other tools: different requirements on different levels

- Think big: in terms of both study size and lab size, however we should not limit to "powerful" institutions, but also include smaller, single labs and such. This does not mean, there should not be lighthouse projects initiated by institutions. The entire size spectrum should be covered

- International outreach is necessary as German researchers work with international tools and technologies

- Must people at lower level be convinced (going into labs and distribute the work here) or more on institutional level?
    - Having a large institute adopting the principles of NFDI will be better than going at lab level. However, this should not be an exclusive solution. Both should be addressed. The principles of NFDI need to be advertised as much as possible. Big institutions using our tools will be very convincing multipliers
    - For example: Omero can do some of these things, but it needs support in big institutions
    - There is a need to prioritise; surveys should first be done and the inquire is out there to be an active participant. It would be very good to already have an institute wide solution that could be introduced to NFDI Neuro.

TA Systems and Behavioural Neuroscience

- Very exploratory research with very heterogeneous approaches are being done in very many labs, every month new analysis tools are being developed
    - How can we deal with the issue of what the standard should be? Who decides? How to deal with different opinions? Come to a consensus in the community, process should be bottom-up

- Hope to reduce the number of solutions and converge (come to just one or a few)
- NFDI-Neuro should not tell what to do, but how to document, how to compare
- Most important is to pick up the people where/at the level they are and provide them with something they benefit from
- Documentation is most important, the tool is secondary (Excel is ok, as long as people understand the structure of the content!)
- NFDI Neuro should not enforce standards of how to do things, but standards that work should be documented and guidelines should be available, and it should provide and be a forum to discuss one workflow via another.

TA Computational Neuroscience

- Authority directives on accounting need to be considered and can differ substantially between different districts, often not compliant with modern lab practices. Can we lobby here? Useful link provided by J Colomb on this topic: www.labfolder.com/features/#compliance

- For computational neuroscience many models were developed and extensive data was harvested, but the problem was how these two could be connected. The technical level is a constant barrier between the computational scientists and the experimental scientists. NFDI Neuro should also work on reconciling these two disciplines so one can comment on the other from their own expertise.

- main goal is to co-develop the models and the data: guarantee interoperability and find translations.
- Interaction between students to learn from each other (i.e. marketplace for data, for tools) should be supported, not only from central 'teaching' instances. The NFDI Neuro platform should help people learn from each other and the idea of a marketplace is very helpful in this respect. A platform where things can be changed and also exchange and discussion of e.g. models can take place.

TA Molecular and Cellular Neuroscience

- Should the teaching concept be domain specific or more general, or both?
    - The goal of NFDI-Neuro is to pick-up the community where it is, both starting courses and more advanced courses should be offered. All/as much as possible levels of specificity should be offered; develop tailored solutions on top of a generic framework
- Collaborative setting are mostly fine within an intra-institutional setting, however, inter-institutional and certainly transregional settings are often very problematic ('Babylonean conditions')

- Software providers are often not very willing to put data export and exchange possibilities or solutions for their software in place, approaching such companies as a NFDI consortium could have more power to get them engaged in such solutions. Furthermore, some good strategies already exist (e.g. Omero) and one could learn from each other experiences
- RDM Plans are becoming more and more important for a successful SFB proposal (and probably in the near future also for individual proposals), we should start an initiative on collecting information on what should be enclosed in a good proposal for CRCs/SFBs with regards to RDMP and also write this in our NFDI-Neuro application

- The neuroimaging community already has some strategies on how to use and open up formats but these strategies should be shared and distributed to a broader audience.

TA Clinical neuroscience

- Engage in and further benefit from ongoing discussions with other NFDIs (e.g. NFDI4Health, NFDI4Medicine) to collaborate and work together (complement each other and not compete)
- how much overlap is there with TA2 (Systems and Behavioral Neuroscience) and how can this be addressed?

→ **General discussion**
- How much would imaging overlap with the Task area 2 (Systems and Behavioral Neuroscience) – anonymisation is not needed animal research but inevitable in clinical neuroscience.
    - The acquired data are much more standardised since they are routinely and constantly collected. Maybe the recorded data and metadata can overlap with Task area 2 and save them some work.
    - The discussion is very similar to their grant proposal which contains intertwining human and animal data; the answers are complex; maybe these plans can also be shared with the community. The data repositories that are needed should remain in parallel for human and animals but should be aligned in terms of standards. The medical side is ahead in some respect since they have standards and rules that are currently still lacking in animal research; maybe these can be applied to animal research as well. In the long run a consolidation between these standards and also with the computational neuroscience Task area should take place.

Further general comments:
- Find a way to facilitate the way of interaction between TAs
- Next workshop could cover different aspects of DM: tools, measures instead of communities

- Get working groups started on these tools and measures…
- We should aim at defining minimal standards for the whole neuroscience field
- NFDI must make some concrete suggestions for standards, not just a discussion board
- RDM is also science and not just administrative stuff
- Interchange between centers is a real problem, no experts, but 16 people with different opinions
- my vision of a promising and very useful contribution of the consortium is a repository, which allows me to quickly overview the type of data used in a published or running study (in best case with unique identification), so I can oversee who has used what type of data for what purpose and how can I have access to the data (even if I only get an email address from the study author) … this can be done retrospectively on existing publications (at the best automatically) … this can be seen as an structured addon to all existing and future publications in neurosciences (as long as they used data in the study), which can be used to find studies and of course data