Neuroimaging and Neuroengineering Lab

How to establish and maintain a multimodal animal MRI data repository

Aref Kalantari Sarcheshmeh[1], Michał Szczepanik[2], Stephan Heunis[2], Christian Mönch[2], Michael Hanke[2,3] ,Thomas Wachtler[4], Markus Aswendt[1,5]

[1] University of Cologne, Faculty of Medicine and University Hospital Cologne, Department of Neurology, Cologne, Germany
[2] Research Centre Jülich, Psychoinformatics Lab, Institute of Neuroscience and Medicine, Brain & Behavior (INM-7), Jülich, Germany
[3] Heinrich Heine University, Institute of Systems Neuroscience, Medical Faculty, Düsseldorf, Germany
[4] Ludwig-Maximilians-Universität München, Computational Neuroscience, Faculty of Biology, Planegg-Martinsried, Germany
[5] Research Centre Jülich, Institute of Neuroscience and Medicine (INM-3), Jülich, Germany

Objectives

What are the goals?

- **Version control** on
- **Remote collaboration** for
- **Standardized structure** of
- **Provenance** tracking of
- **Workflow** management of

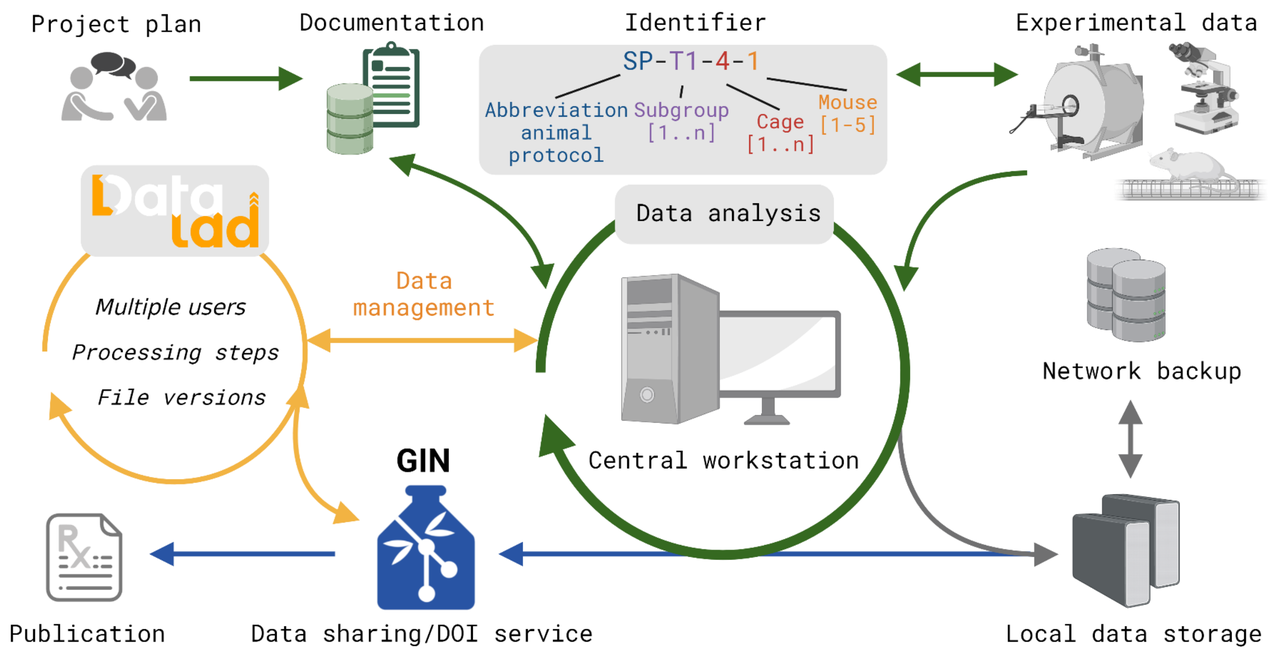
Data

Findable Accessible Interoperable Reusable

Replacement Reduction Refinement

Structure

How does the workflow look like?



Methods

What are the main tools needed?

Datalad

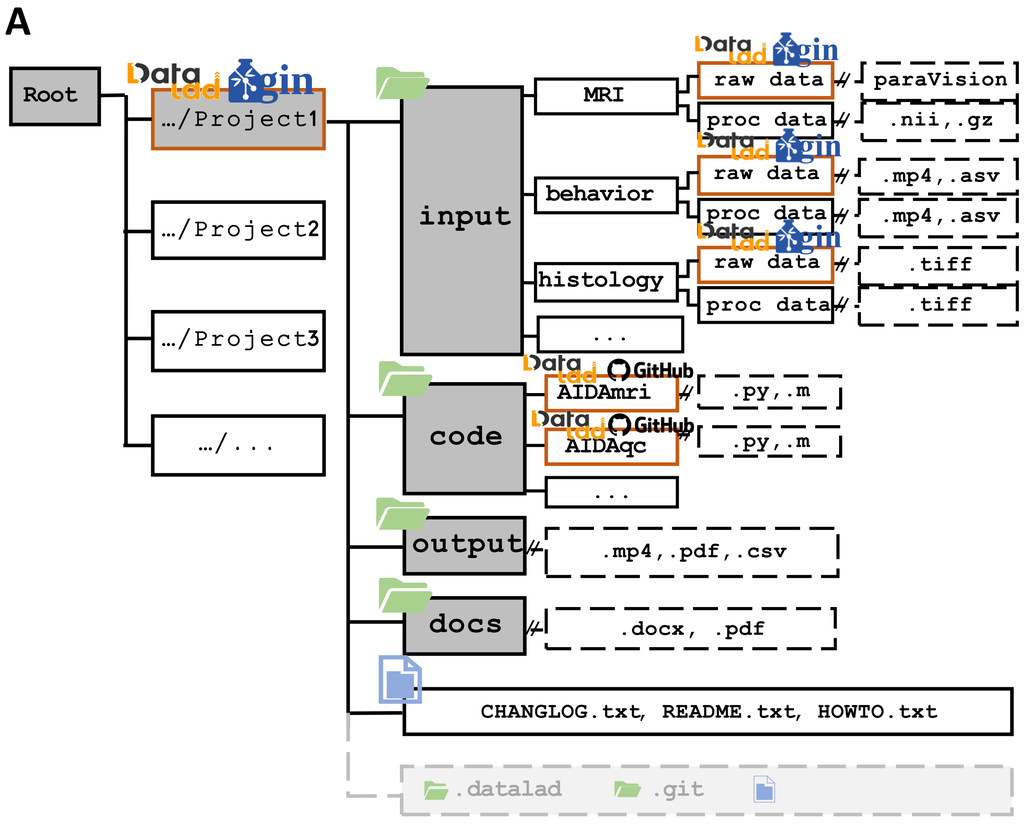
1. Datalad is a data management tool for scientific research.
2. It tracks changes, simplifies collaboration, and facilitates data sharing.
3. Datalad provides a unified interface for managing and analyzing large datasets.

gin GitHub

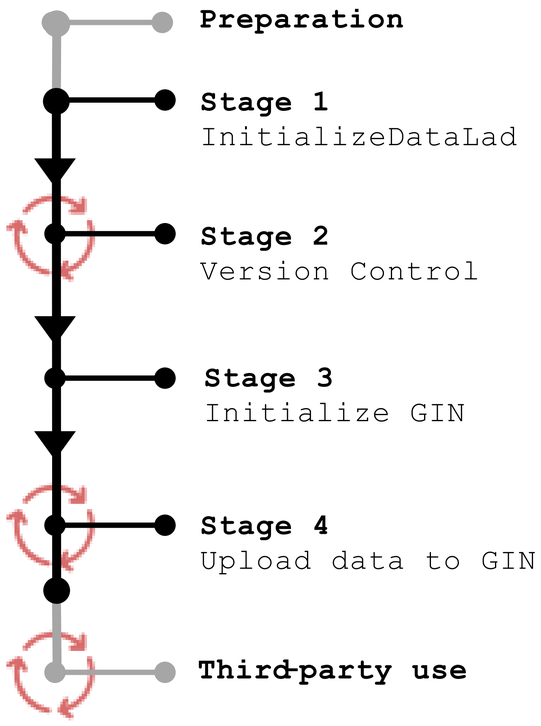
1. Gin-gnode is a web-based platform for version control of scientific data.
2. It allows researchers to store and share data, code, and documentation in a unified platform.
3. Gin-gnode facilitates collaboration and reproducibility of scientific experiments.
4. GitHub also provides version control for code, although it is more commonly used for software development

Results

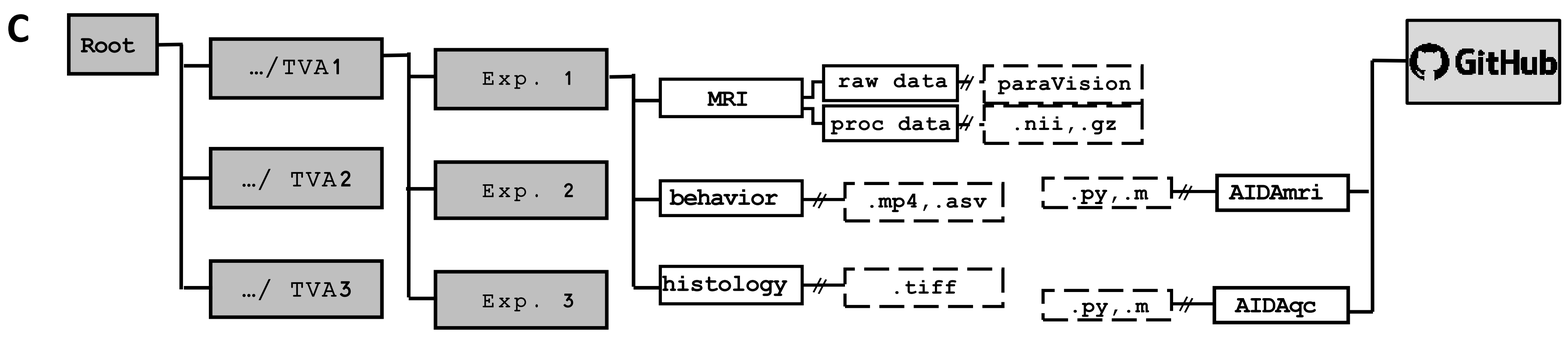
A



B



C



A

- **YODA-directory structure**
- **Integration of DataLad, GIN and GitHub**

B

- **Step-by-step guide for creating the DataLad dataset and Third-party use**

C

- **Folder Structure based on the permit of performing animal experiments without DataLad (TVA)**

Outlook

How does the future look like?

DataLad also offers provenance tracking (re-executable annotation of changes), and **metadata management**. Together with DataLad's **run-command** and **rerun-command**, which allow **"tracked execution"** of operations on a dataset, DataLad enables truly reproducible research

Limitation

… for users without programming experience

The use of this workflow in research poses challenges for researchers who lack familiarity with programming and related technical aspects. The installation and setup of tools such as git and git-annex (required for DataLad), as well as the use of command line interfaces, may dissuade interested users. Simplifying these pre-installation steps could improve the usability of the workflow and promote collaboration among researchers.

- Example datasets are available at: *https://gin.g-node.org/Aswendt_Lab*

Necessary information on every project can be found in the datasets

esmi european society for molecular imaging
EMIM 2023 - Salzburg

DFG Deutsche Forschungsgemeinschaft German Research Foundation
FRIEBE-STIFTUNG

References:

1. Begley CG, Ioannidis JPA. Reproducibility in science: improving the standard forbasic and preclinical research. Circ Res. 2015;116: 116–126.
2. Poldrack RA, Baker CI, Durnez J, Gorgolewski KJ, Matthews PM, Munafò MR, et al. Scanning the horizon: towards transparent and reproducible neuroimaging research. Nat Rev Neurosci. 2017;18: 115–126.

CONTACT
Aref Kalantari
Aref.kalantari-sarcheshmeh@uk-koeln.de