

## mHealth Datenmanagement Machine Learning

## Step by step to FAIR data and reproducibility management.

FAIR data management is the foundation of scientific discovery and efficiency, but difficult to achieve in practice. DataLad aids data management with version control, streamlined data publication and -retrieval routines, and reproducible data analysis.

Data discovery: File-level access to > 500TB of openly available datasets

Simplified version control: Easy tracking, built upon industry standards

**Decentral:** No central server/ service required - your computer suffices

**Interoperable:** Integrates with third party services for data/code hosting

Data provenance for science and security, optional encryption & credentials

Open and established: Used by portals and consortia (OpenNeuro, CRCs, ...)



Free and open source Python software, available on all operating systems.

DataLad is an ecosystem of software, with just the right extension for your usecase.

website:



chat:



handbook:





## Manage, reproduce, reuse:



Create a dataset



Track data regardless of size or format



Capture processing reproducibly Willearn



Publish data to infrastructure you like

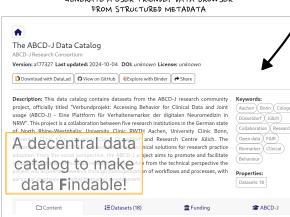


Clone and get others' datasets



Extract and expose metadata





Data sharing with no breach of privacy