## German Neuroinformatics Node

# Research Data Management, Hosting and Sharing

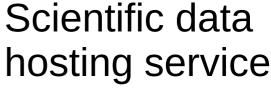
The GIN research data platform and associated tools





## GIN - G-Node Infrastructure







#### GIN

#### **Modern Research Data Management for Neuroscience**

...distributed version control, flavoured for science

#### Manage your research data

Upload your data to private repositories. Synchronise across devices. Securely access your data from anywhere.

#### Version your data

Uploaded files are automatically versioned. Retrieve any previously uploaded version of a file. Never lose file history.



#### Share your data

Collaborate with colleagues. Make your data public. Make your data citable with the GIN DOI service.

#### Open Source

Based on open source projects such as Git, git-annex, and Gogs. You can even set it up in your lab!



mouse

doi/Resolving and rescuing developmental miswiring in a mouse model of cognitive impairment \u00a7 ★ 0 P 0

Search

**★** 0 🔑 0

**★** 0 🔑 0

Updated 1 week ago

#### hiobeen / Mouse hdEEG ASSR Hwang et al

A set of high-density EEG (electroencephalogram) recording obtained from awake, freely-moving mice (mus musculus).

Updated 1 week ago

#### mdp10vv / DCA

Supplementary Figure 1. Human bile increased afferent firing frequency in mouse proximal colon

Updated 5 months ago

#### pgoltstein / Mini1p2pcomparison Glas Goltstein 2018

Dataset of "Benchmarking miniaturized microscopy against two-photon calcium imaging using single-cell orientation tuning in mouse visual cortex". By Annet Glas, Mark Hübener, Tobias Bonhoeffer & Pieter M Goltstein; Max Planck Institute of Neurobiology







## GIN - G-Node Infrastructure

 Neuroscience data repository recommended by:







- Statistics
  - 3000+ repositories (~1200 public)
  - ~2900 users
  - 62 TB of research data





#### General features

- Github like service
  - Public and private repositories
  - Access control via user permissions, organizations and teams
  - Issues, pull requests, wiki
- Platform independent data access
- Enables easy data publishing (DOI)





#### General features

- Indexing and search of text file content
- Built-in file versioning
- Manual and automated workflow integration
- Specific handling of large files via git annex
  - Handles large files better than git
  - Enables to keep only required file content locally





#### Handle collaboration

- Repositories and permissions
  - add registered users to a repository
  - Invite non-registered users by email address
  - assign Read only; Write access; Admin access

#### Organisations

- Create repositories using the organisation
- Create teams using the organisation
  - assign read; write; admin; owners
  - add registered users to teams





### Handling data using the web interface

- Upload files
- Create new folders and files
- Edit files online
- Delete files
- Download single files
- Download repository zip
  - Download large file content separately
- Extensive user documentation

https://gin.g-node.org/G-Node/Info/wiki/Web+Interface





# Handling data using the gin command line client

- Simplified git commands
  - Download, upload
  - Local user login
  - Version control
  - Local large file content handling
  - Work with different remotes
  - Work with different servers
  - Extensive command help
- Full git and git annex commands supported





# Handling data using the gin command line client

- Automatic ssh key session handling
- Automate workflows; integrate into scripts
- Handle large file content
  - download only content required
  - specify file size as large file content
- Available for all operating systems
   https://gin.g-node.org/G-Node/Info/wiki/GIN+CLI+Setup
- Beginner and advanced usage documentation

https://gin.g-node.org/G-Node/Info/wiki/GIN+CLI+Usage+Tutorial





## Handling data using the datalad client



- Use the datalad client with GIN
- ssh upload and download supported
- Requires ssh key upload
- GIN specific how-to documentation

https://handbook.datalad.org/en/latest/basics/101-139-gin.html





## Features - Data Publication

### Data publication



- Free data publication service
- Citable datasets (DOI)
- 290+ published data sets
- Total published data size: 14TiB
- Free for all public repositories on gin
- Provide metadata about the dataset
- How-to documentation

https://gin.g-node.org/G-Node/Info/wiki/DOI



https://doi.gin.g-node.org



## Features - self-hosted GIN

### Deploy own GIN instance

- Open source
- Docker deployment
  - Pre-built Container available
  - Build your own container
- Local or public
- Customizable look
  - stylesheets
  - templates
- Use gin command line client with all GIN instances
- Full installation documentation

https://gin.g-node.org/G-Node/in-house-gin





# Features – Repository validator

#### Repository validation service

Companion service to gin

https://valid.gin.g-node.org

- Validate repository content
  - BIDS repository format
  - NIX neuroscience data file format
  - odML metadata file format
- Login using gin credentials
- Validate any public GIN repository
- Set up automatic validation on upload to owned repository
- Link results badge on GIN repository





# Features – Repository validator

## Repository validation service

- Submit custom validators
  - Core code easy to extend
  - Bash, python, npm scripts, executables
  - Validator extension documentation
- Run your own validation server
  - Open source
  - Build on all platforms and deploy via docker
  - Connect to GIN or to self-hosted GIN service

https://github.com/G-Node/gin-valid



#### Resources

#### Linklist

https://gin.g-node.org

https://gin.g-node.org/G-Node/Info/wiki

https://doi.gin.g-node.org

https://gin.g-node.org/G-Node/Info/wiki/In+House

https://github.com/G-Node/gogs

https://gin.g-node.org/G-Node/in-house-gin

https://valid.gin.g-node.org

Get in touch

info@g-node.org

