The G-Node Infrastructure Services: Efficient Research Data Mangagement for Neuroscience and Beyond

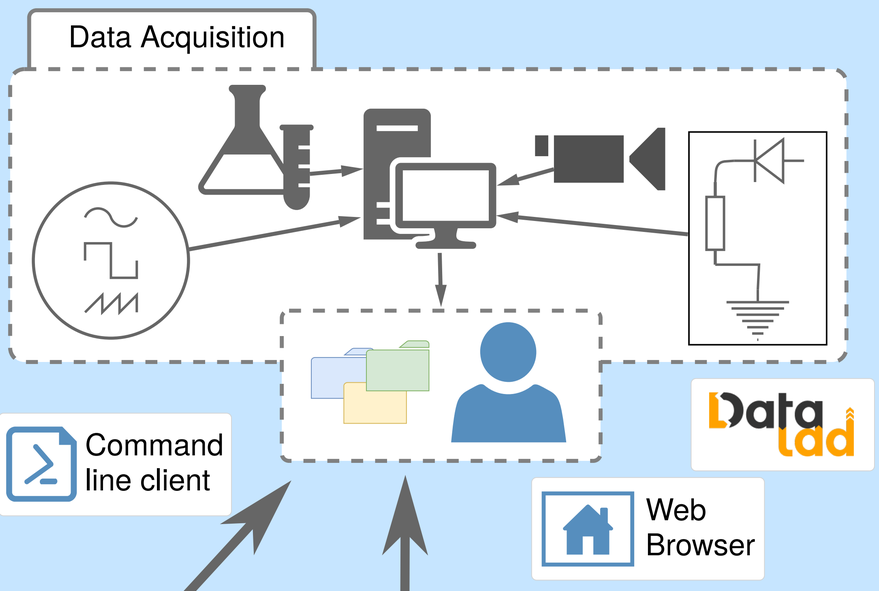
Achilleas Koutsou, Michael Sonntag, Christian Garbers, Thomas Wachtler

German Neuroinformatics Node, Department Biologie II, Ludwig-Maximilians-Universität München, Germany

Maintaining reproducible data workflows while keeping data in sync, backed up, and easily accessible from within and outside the lab is a key challenge in research. To help minimize the time and effort required for these tasks, the GIN services provide support for comprehensive, reproducible and versioned management of scientific data throughout the data lifecycle.

GIN Services for Data Storage, Collaboration and Data Publication

GIN core features
- Data access from any location
- Built-in versioning
- Platform and format independent
- Secure access
- Public and private repositories



Supporting research data management throughout the data lifecycle
- Version control for code and data
- Efficient collaboration through access control and tracking of changes
- Validation services to ensure data quality
- Data publication (DOI) with a button click

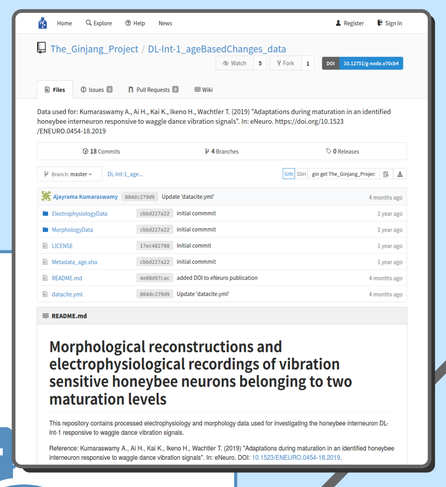
In-lab Local Instance

Local GIN Server

Intranet

Command line client

Local Hosting
- GIN is open source
- prebuilt docker containers for easy installation
- use your own data storage



GIN Platform Server

gin.g-node.org

Collaboration

- User management
- User access levels
- On and offsite collaboration
- Issue tracker
- Supports repository integrity by versioning and Pull Requests

Automation and Validation

gin-valid

Data validation service

valid.gin.g-node.org

gin-proc

Data processing service

proc.gin.g-node.org

Data Publication and Findability

gin-doi

DOI service

doi.gin.g-node.org

Recommended by:

gin-dex

Data search service

gin.g-node.org/explore/data

Automated Data Validation
- Automatically runs validation on selected repositories
- Supported validation formats:
  - BIDS
  - odML
  - NIX
- Easily extensible to more formats
- Format validation contributions are welcome

Automated Data Processing
- Automatically runs pre-defined processing pipelines
- Triggered on repository changes
- Automatically returns specified results
- Based on SnakeMake and DroneCI

Findable Data via GIN
- Automatic indexing of text based files
- Online search for repository content
- Interactive rendering of markdown formats

Data Publication
- Seamless publication
- Persistent identifiers (DOI)
- Data - Article links

Resources and References

Contact:
info@g-node.org

Poster presented at RDA-DE 2020

www.bernstein-network.de

www.nfdi-neuro.de

GIN (RRID:SCR_015864):	*https://gin.g-node.org*
BIDS (RRID:SCR_016124):	*http://bids.neuroimaging.io*
NIX (RRID:SCR_016196):	*http://www.g-node.org/nix*
odML (RRID:SCR_001376):	*http://www.g-node.org/odml*
SnakeMake (RRID:SCR_003475):	*https://snakemake.readthedocs.io*
DroneCI:	*https://drone.io/*

Supported by BMBF grants 01GQ1302, 01GQ1509