

G-Node Microservice infrastructure for continuous validation, processing and indexing of research data on an open platform



2 months ago

2 months ago

2 months ago

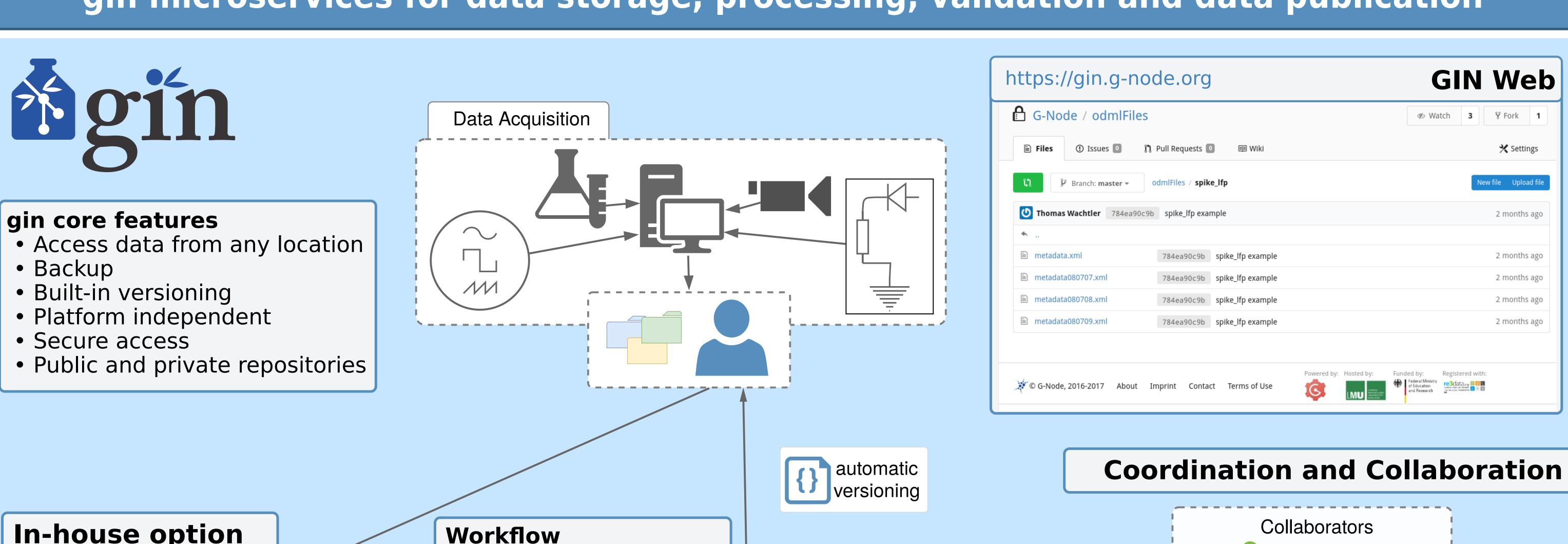
2 months ago

Michael Sonntag¹, Achilleas Koutsou¹, Christian Garbers¹, Jiří Vaněk¹, Mrinal Wahal¹, Thomas Wachtler¹

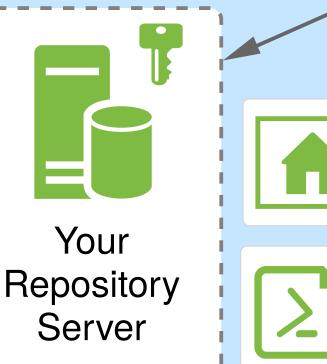
¹German Neuroinformatics Node, Department Biologie II, Ludwig-Maximilians-Universität München, Germany;

Research requires data to continuously be processed, analyzed and visualized; data needs to be quality checked, verified and backed up. Data and metadata need to be made publicly available in an easy to find and use manner. Many of these tasks can be automated, which usually leads to fewer errors and a higher results quality. To facilitate these tasks, we introduce a suite of microservices for the G-Node data infrastructure (GIN), an open platform for collaboration and sharing of research data and code.

gin microservices for data storage, processing, validation and data publication



In-house option





Command

line client

- Manual access via web and command line
- Script automation





- User management
- User access levels
- On and offsite collaboration
- Online issues help coordination
- Ensure repository integrity with versioning and "Pull requests".

gin or local hosting

- gin is open source
- use the free online gin service
- use your own in-house instance

Automation and validation tools

GIN Repository Server

gin.g-node.org

Data publication and searchability

Persistent publications

 Every public gin repository can be published

DOI service

- DOIs for paper publication data
- Research software publication
- Data set publication



Data validation service valid.gin.g-node.org

gin-proc

Data processing service proc.gin.g-node.org

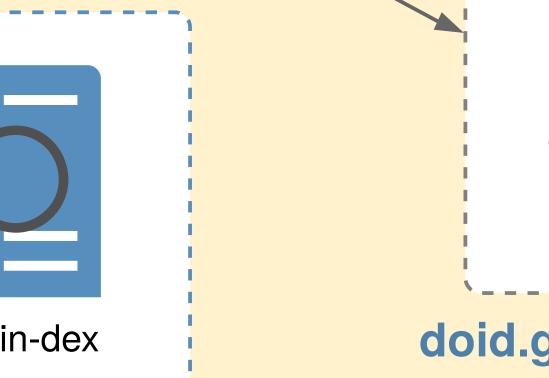
Automated Data validation

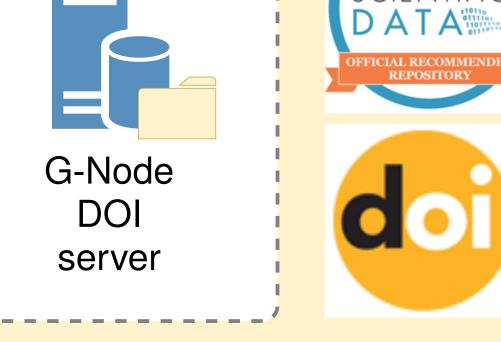
- Opt in gin service
- Automatically runs validation on selected repositories
- Supported validation formats:
- BIDS
- odML
- NIX
- Easily extensible with custom formats

Automated Data processing

- Opt in gin service
- Based on snakemake and DroneCl
- Automatically runs pre-defined processing pipelines
- Triggered on repository changes
- Automatically returns result files









SCIENTIFIC

doid.gin.g-node.org

Findable data via gin

- gin provides automatic indexing of all text based files
- online search for repository content
- Online browsing of
- Markdownyaml
- json

XML

Resources



Poster presented at **INCF Neuroinformatics 2019** Warsaw, Poland

Contact: dev@g-node.org GIN (RRID:SCR_015864): BIDS (RRID:SCR 016124): *NIX (RRID:SCR 0*16196): odML (RRID:SCR 001376): SnakeMake (RRID:SCR 003475) DroneCI: https://drone.io/

https://gin.g-node.org http://bids.neuroimaging.io http://www.g-node.org/nix http://www.g-node.org/odml https://doi.org/10.1093/bioinformatics/bts480 https://drone.io/

Supported by BMBF grants 01GQ1302, 01GQ1509



■ Federal Ministry of Education and Research