Open tools and services for efficient data management, collaboration and data publication

Michael Sonntag[1], Achilleas Koutsou[1], Jan Grewe[2], Thomas Wachtler[1]

[1]German Neuroinformatics Node, Department of Biology II, Ludwig-Maximilians-Universität München, Martinsried, Germany; [2]Institute for Neurobiology, Eberhard-Karls-Universität Tübingen, Tübingen, Germany

Maintaining reproducible analysis workflows while keeping data in sync, backed up, and accessible from within and outside the lab is a growing challenge in research. To facilitate these tasks, the G-Node provides tools and services for research data management that are designed to seamlessly integrate with and support the scientist's lab data workflows. Specifically, these tools address key components of efficient data management: Metadata collection, data organization and storage, and reproducible data workflows, sharing and publication.

Collect all information about your experiment

Main features

- Open metadata format [1]
- Flexible hierarchical key-value storage
- Save to common structured formats: XML, JSON, YAML

Validation engine

- automatically validate documents
- define your own data validations

File structure example

```
Document
+-Section
| +-Property-[Values]
+-Section
  +-Section
  | +-Property-[Values]
  | +-Property-[Values]
  +-Section
```

Manage data and metadata in one open, versatile format

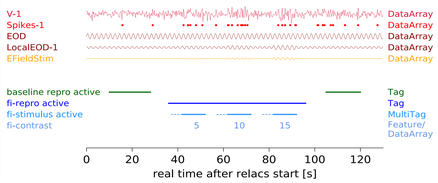
Main features

- Open data format
- Store data, analysis results, and metadata conveniently in the same file
- Descriptive associations between data, analysis results, and metadata
- Built for recording script automation

Free open source libraries [5]

Integrates with [6,7]

Recording to NIX file: Relacs [8]



Electrophysiology data raw and analysed in a single file

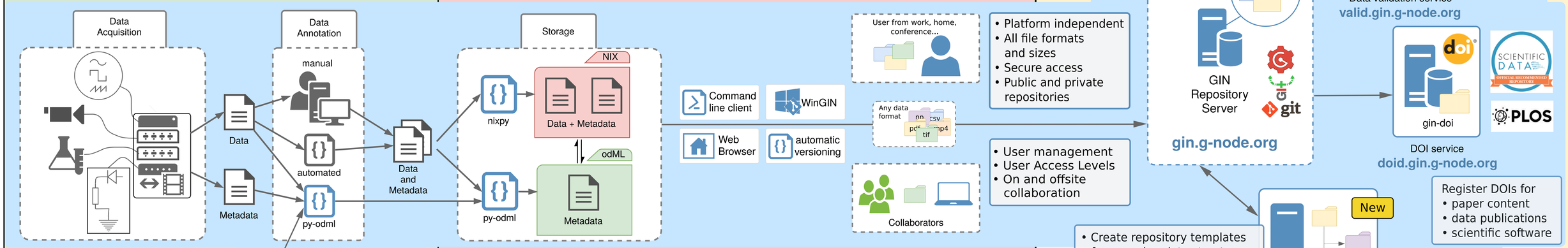
Store data securely; publish and collaborate with ease

Main features

- Access data from any location
- Free storage for scientific data [10]
- 'Built-in' versioning (based on git [11])
- Microservice architecture for easy addition of new services.

GIN microservices

- Automatic data validation
- Add custom validations

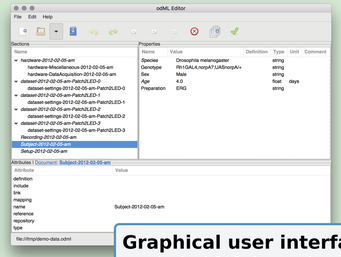


- Platform independent
- All file formats and sizes
- Secure access
- Public and private repositories

- User management
- User Access Levels
- On and offsite collaboration

Register DOIs for
- paper content
- data publications
- scientific software

- Create repository templates for recurring data structures
- Automate repository setup
- Manage lab-wide permissions



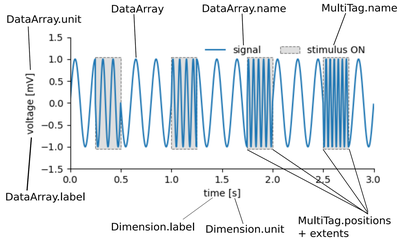
templates.g-node.org
terminologies.g-node.org

Re-usable concepts [2]
- metadata building blocks
- importable definitions

- file creation automation
- file content standardization

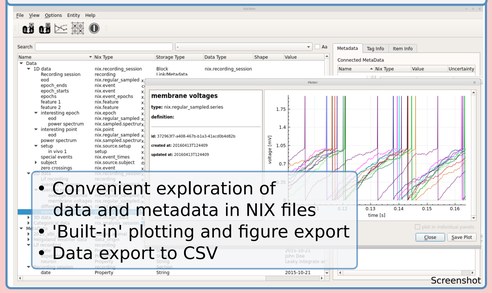
Graphical user interfaces

- odml-ui [3]: browsing and template editing
- odmltables [4]: filtering and tabular export



Fully annotated file contains all information required to understand and interpret the data and their properties.

NIX View GUI [9]



- Convenient exploration of data and metadata in NIX files
- 'Built-in' plotting and figure export
- Data export to CSV

Screenshot

Resources

[1] Grewe et al (2011), doi:10.3389/fninf.2011.00016
[2] http://www.g-node.org/projects/odml/terminologies
[3] https://github.com/G-Node/odml-ui
[4] https://github.com/INM-6/python-odmltables
[5] http://g-node.github.io/nix
[6] http://neuralensemble.org/neo
[7] http://neuralensemble.org/elephant
[8] https://github.com/relacs/relacs
[9] http://bendalab.github.io/NixView
[10] https://gin.g-node.org
[11] https://git-scm.com
[12] https://github.com/G-Node/tonic

Contact: info@g-node.org

Visit us at the G-Node virtual booth for demonstrations.

Supported by BMBF grants 01GQ1302, 01GQ1509

Federal Ministry of Education and Research