Formats, tools and services for efficient data management, reproducibility and collaboration in neuroscience

A. Koutsou[1], C Garbers[1], M Sonntag[1], C J Kellner[1], J Grewe[2], T Wachtler[1]

[1]Ludwig-Maximilians-Univität München, Germany; [2]Eberhard Karls Universität Tübingen, Germany

Supported by BMBF grant 01GQ1302

Federal Ministry of Education and Research

Download this poster!

Management of scientific data, including consistent organization and storage of data, is a challenging task. Data needs to be annotated with metadata to provide information about the underlying experiment to ensure reproducibility. Accessing and managing data from multiple workplaces while keeping it in sync, backed up, and easily accessible from within or outside the lab, is even more demanding. To minimize the time and effort scientists have to spend on these tasks, we here present formats and tools designed for comprehensive and reproducible management of scientific data.

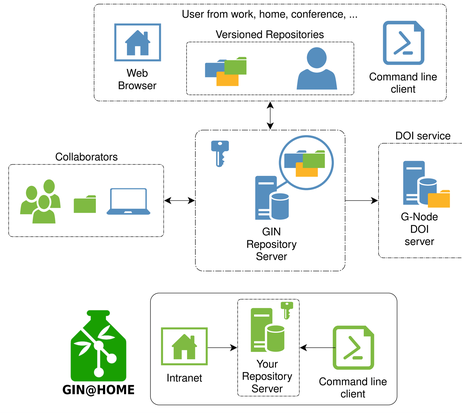
Secure data storage, easy collaboration and publication

Main features
- Access data from any location
- Free storage for scientific data [1]
- Built in versioning (built on git [2])

- Platform independent
- Secure access
- Public and private repositories
- Citable data by DOIs

Workflow
- Browse, download, and upload data via web [1]
- Download and upload large files via command line [3]
- Automate workflows using command line client

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

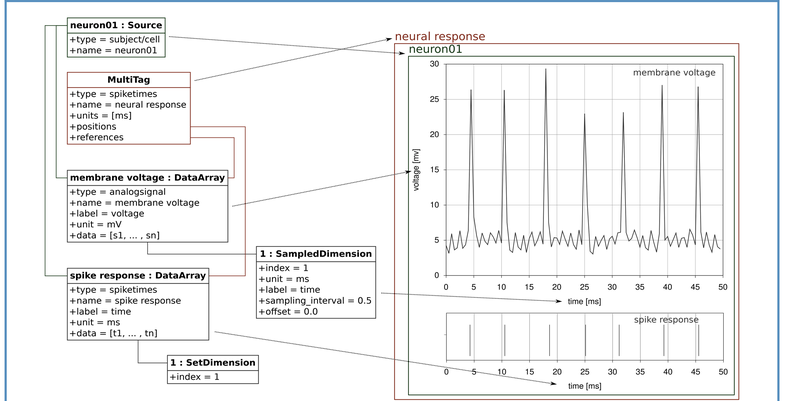


Manage data and metadata together in an open, versatile format

Main features
- Open data format
- Store data, analysis results, and metadata conveniently in the same file
- Descriptive associations between data, analysis results, and metadata

- Free open source libraries for multiple programming languages: C++ [4], Python [5], Matlab [6], Java [7]
- NIX IO for Neo [8]
- Enables interoperability with Neo compatible tools, e.g., the Elephant toolkit [9]

The NIX data model



- **NixView** [10] Cross-platform GUI viewer
- Available for Windows, macOS, and Linux
- Convenient exploration of both data and metadata of NIX files
- Raw data can be browsed via tabular display and easily exported to CSV
- Facilitates plotting support for a large variety of raw data as well as the export of plots

Collect and manage all information about your experiment

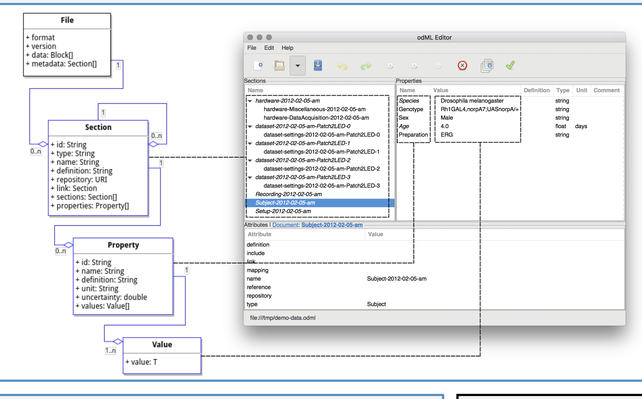
Main features
- Open metadata format [11]
- Flexible hierarchical key-value storage
- Save to common structured formats: XML, JSON, YAML

- Terminology repository [12] for reusable definitions
- Template system for reusable metadata structures

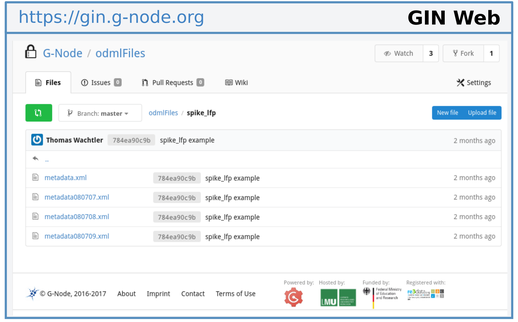
- Export to RDF
- Query metadata using semantic web technologies

- GUI editor [13]
- Available for Windows, macOS, and Linux
- Cross-document drag-and-drop for metadata subtrees

odML Metadata format: Schema and GUI

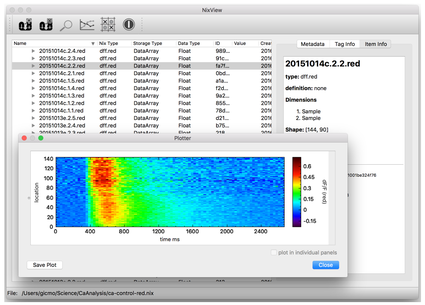


GIN Web

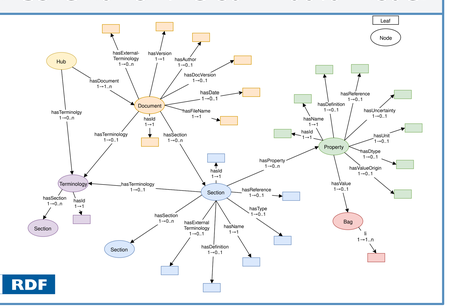


GIN Client

```
$ gin login
Login: achilleas
Password: **********
Hello achilleas. You are now logged in.

$ gin get G-Node/odmlFiles
Fetching repository 'G-Node/odmlFiles'... OK
Initialising local storage... OK

$ gin ls
Synced:

        metadata.xml
        metadata080707.xml
        metadata080708.xml
        metadata080709.xml

$
```



RDF schema for the odML data model



Resources

Contact: dev@g-node.org

[1] https://gin.g-node.org
[2] https://git-scm.com
[3] https://web.gin.g-node.org/G-Node/Info/wiki/gin-cli
[4] https://github.com/G-Node/nix
[5] https://github.com/G-Node/nixpy
[6] https://github.com/G-Node/nix-mx
[7] https://github.com/G-Node/nix-java
[8] http://neuralensemble.org/neo
[9] http://neuralensemble.org/elephant
[10] http://bendalab.github.io/NixView
[11] Grewe et al (2011), Frontiers in Neuroinformatics 5:16
[12] http://www.g-node.org/projects/odml/terminologies
[13] https://github.com/G-Node/odml-ui