# *How to spend less time on data management and more on the science:*
# Open tools for efficient data organization, reproducible workflows, and collaboration

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

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

Maintaining reproducible data workflows while keeping data organized, annotated, backed up, and easily accessible from within and outside the lab is a key challenge in research.
To minimize the time and effort scientists spend on these tasks, we provide a suite of tools designed for comprehensive and versioned management of scientific data.
These tools facilitate the storage, publication, annotation, and analysis of data and metadata in the interest of reproducibility, data sharing and re-usability.

## Collect and manage all information about your experiment

**Main features**
- Open metadata format [1]
- Flexible hierarchical key-value storage
- Save to common structured formats: XML, JSON, YAML
- Terminology repository [2] for reusable definitions
- Template system [3] for reusable metadata structures

**Graphical user interfaces**
- odml-ui [4]: browsing and template editing
- odmltables [5]: filtering and tabular export
- Available for macOS and Linux

**Utilising semantic web technologies**
- Export to PDF
- Query metadata
- Search cross datasets via SPARQL

## Manage data and metadata together in one open, versatile format

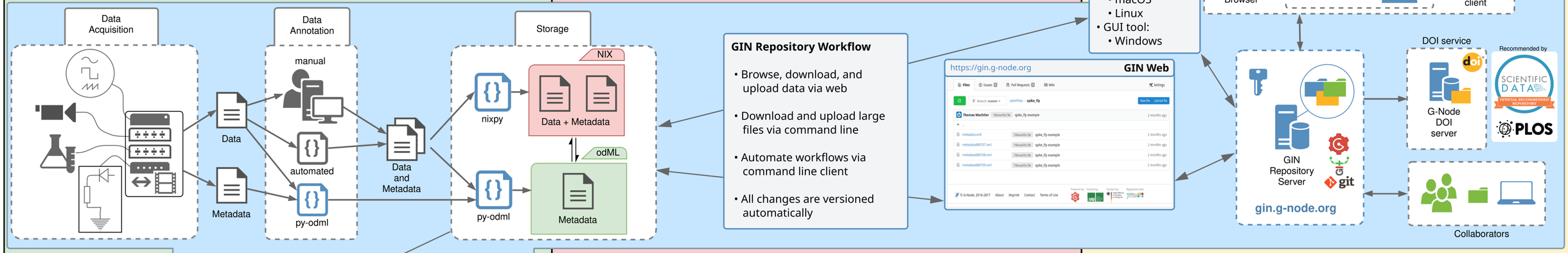
**Main features**
- Open data format
- Store data, analysis results, and metadata conveniently in the same file
- Descriptive associations between data, analysis results, and metadata

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

## Store data securely for easy access, collaboration, and publication

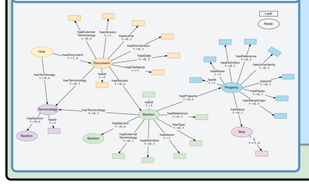
**Main features**
- Access data from any location
- Free storage for scientific data [10]
- Built-in versioning (built on git [11])

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

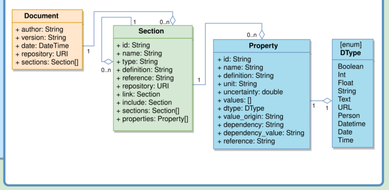



**meta.g-node.org**
- access diverse metadata datasets
- all datasets are publicly available
- searchable by SPARQL via API and web


## odML RDF schema

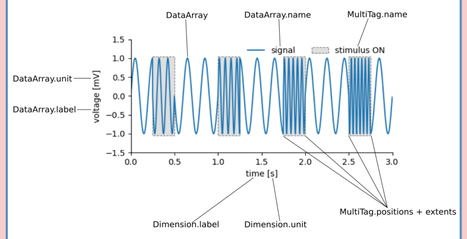


## The odML Metadata format

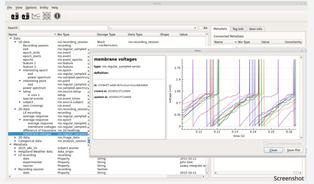



## NIX usage example

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




## NixView GUI [9]




- Linux, macOS and Windows
- Convenient exploration of data and metadata of NIX files
- Browse raw data via tabular display and export to CSV
- Built-in plotting and figure export

**Collaboration**
- User management
- User Access Levels
- Local and remote collaboration

**NEW Data Validation Service**
- Automatic validation of data files on change
- Facilitates adherence to data format standards
- Currently supported formats:
  - odML, NIX, BIDS

Resources

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

Contact: dev@g-node.org

Supported by BMBF grants 01GQ1302, 01GQ1509

Federal Ministry of Education and Research