

Achieving reproducible data workflows: Lightweight tools for safe and efficient data management

Achilleas Koutsou¹, Michael Sonntag¹, Christian Garbers¹, Christian Johannes Kellner¹, Jan Grewe², Thomas Wachtler¹

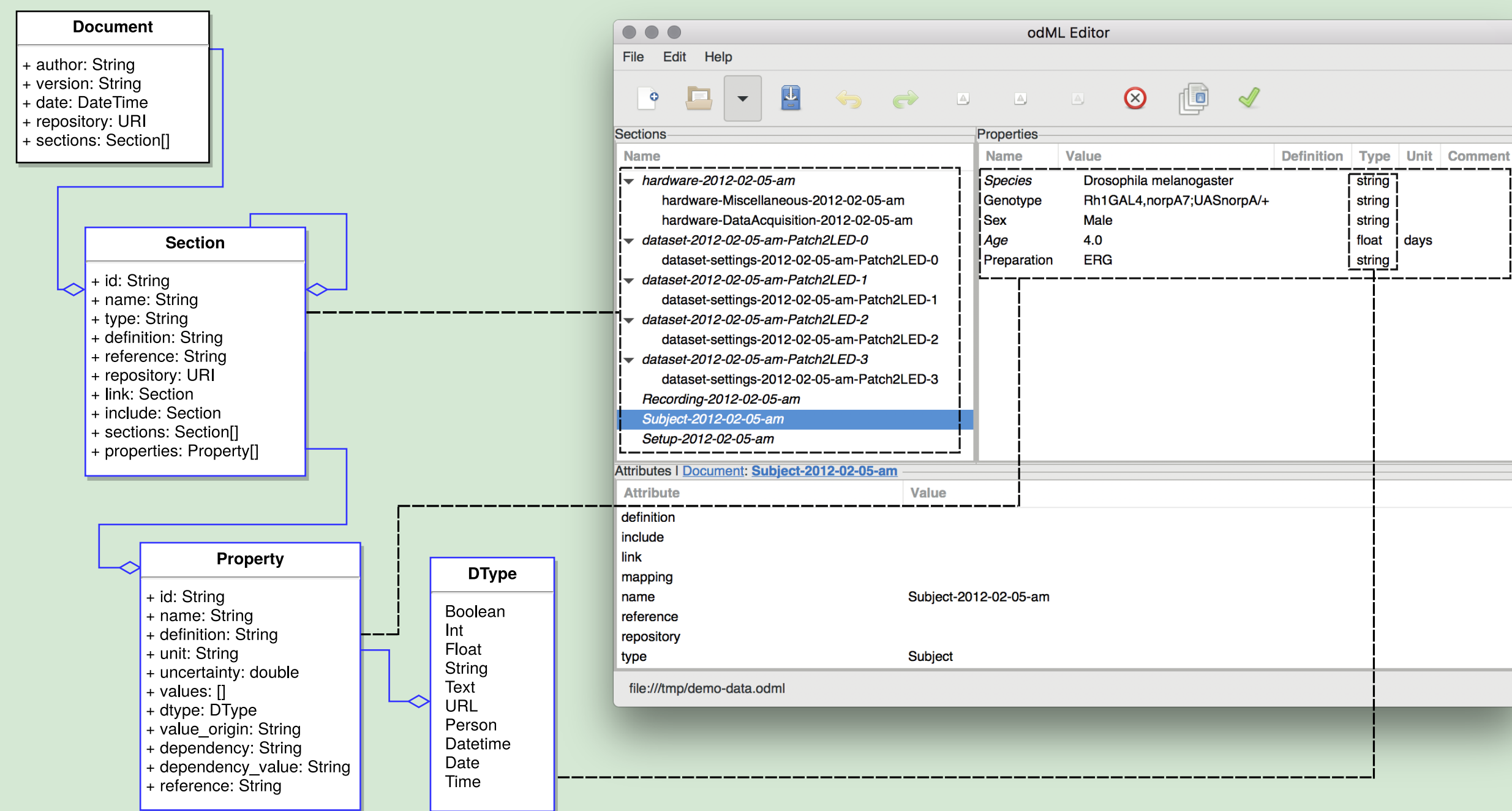
¹German Neuroinformatics Node, Department Biologie II, Ludwig-Maximilians-Universität München, Germany;

²Institut für Neurobiologie, Universität Tübingen, 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 minimize time and effort scientists have to spend on these tasks, we provide a suite of tools designed for comprehensive, reproducible and versioned management of scientific data.

Organize and Store Data and Metadata

Manage all information about your experiment

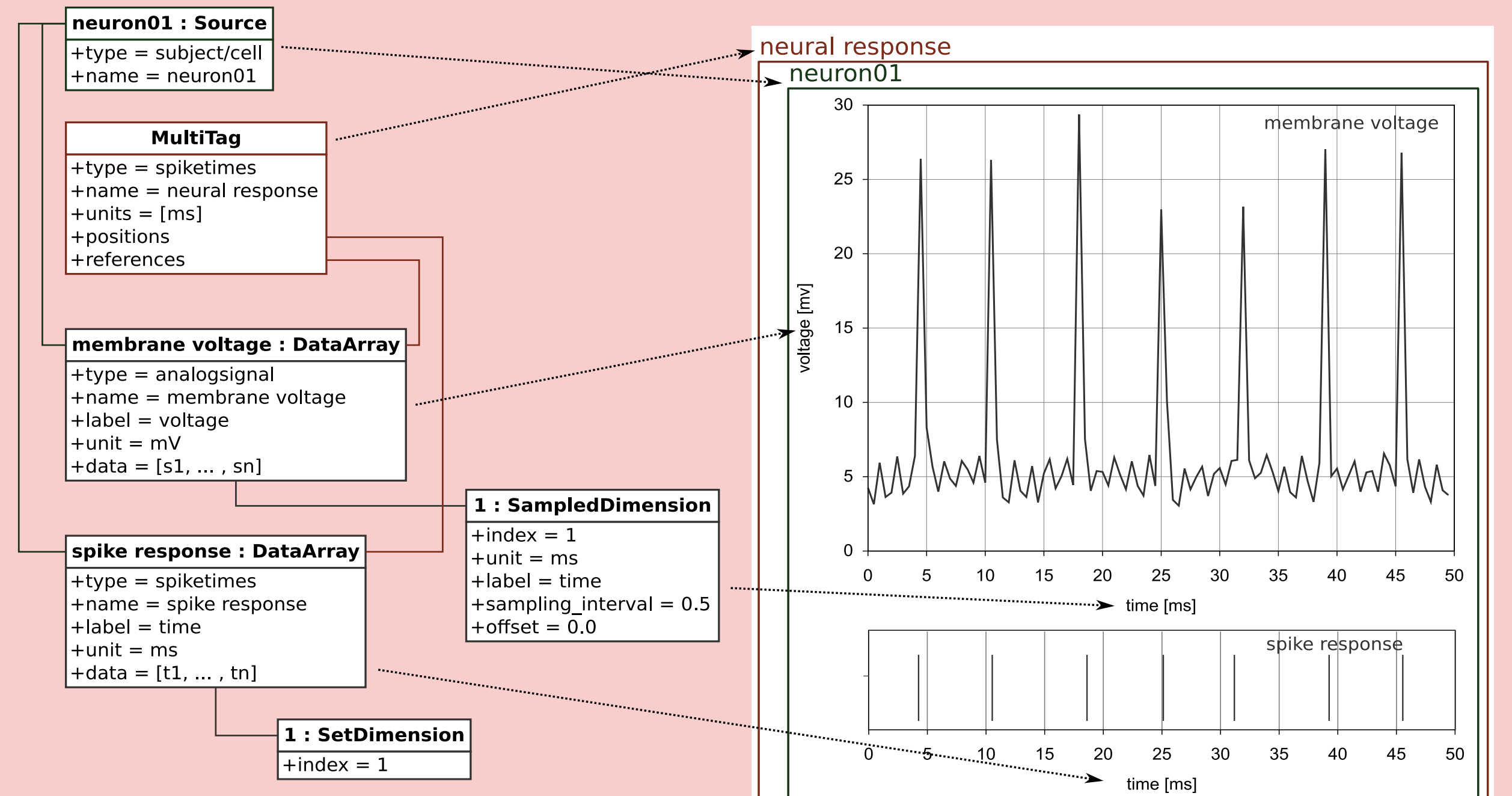


The odML format

- Metadata format used in NIX
- Read and write metadata using library or editor
- Export to RDF: Query using Semantic Web



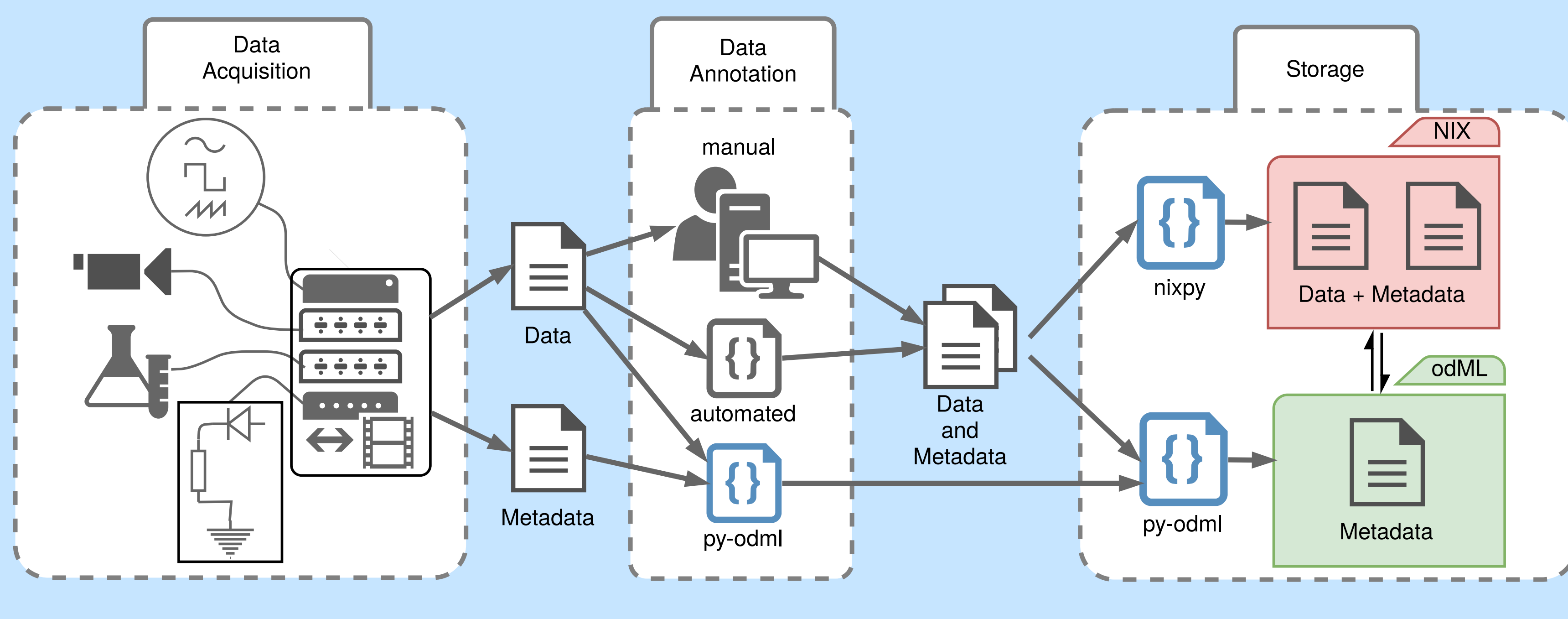
Manage data and metadata in one versatile format



The NIX format

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

Data / Metadata acquisition workflow using odML and NIX



Changes to files can be tracked in GIN (see below)

Libraries

Free open source libraries for:



Store data securely; publish and collaborate with ease



- #### Workflow
- Manual access via web and command line
 - Script automation

GIN Client

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

$ gin get G-Node/odmlFiles
Fetching repository 'G-Node/odmlFiles'... done.

$ gin ls
Synced:
  metadata.xml
  metadata080707.xml
  metadata080708.xml
  metadata080709.xml
$
```



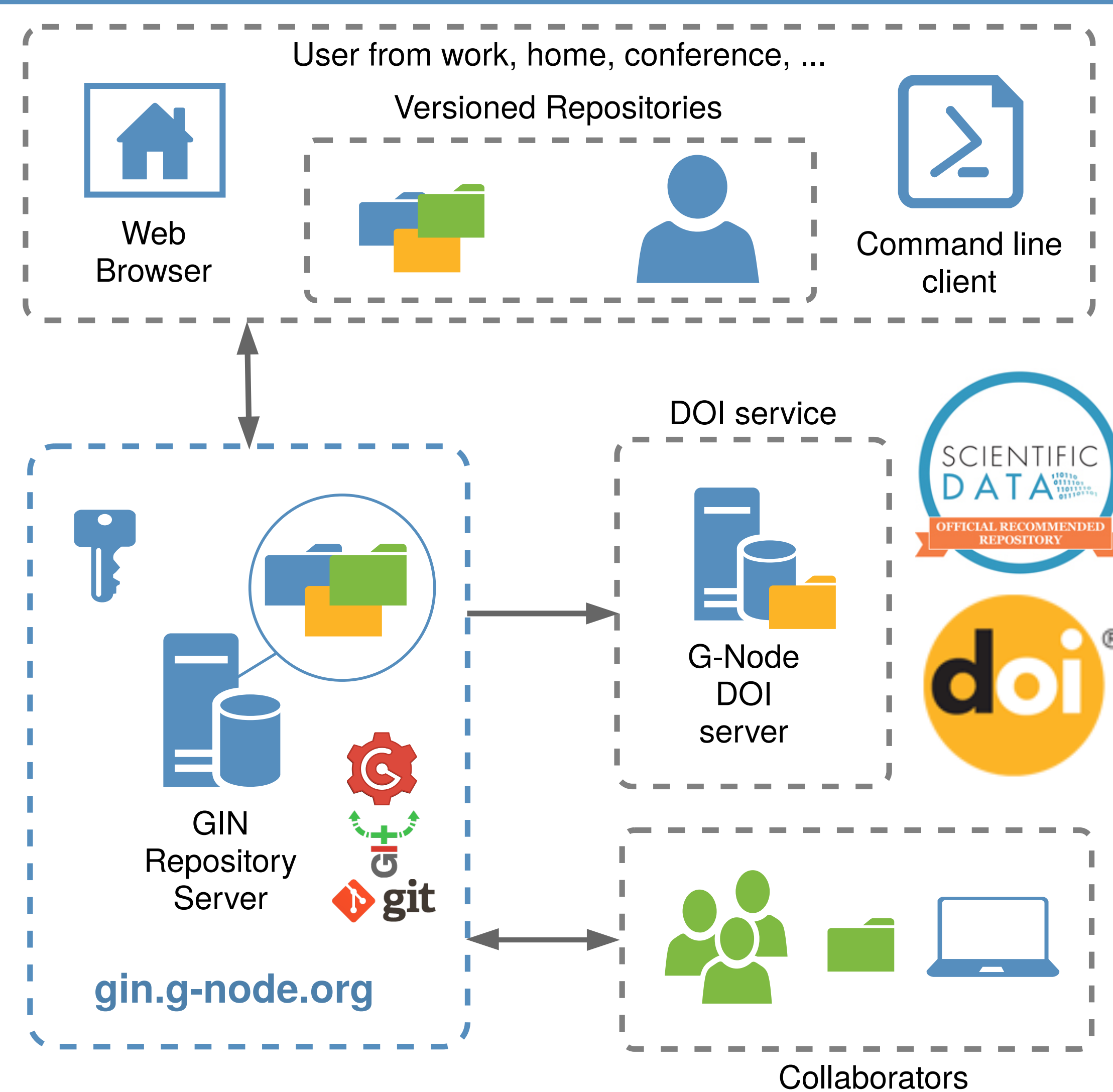
Main features

- Access data from any location
- Backup
- Built-in versioning

Collaboration

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

GIN Services



Resources



Contact:
dev@g-node.org

Poster presented at the NWG Conference 2019 (Goettingen)

<https://gin.g-node.org>
<https://github.com/G-Node/nix>
<https://github.com/G-Node/python-odml>
<https://github.com/G-Node/gin-cli>
<https://github.com/G-Node/gogs>
<http://neuralensemble.org/neo>
<http://neuralensemble.org/elephant>

Supported by BMBF grants
01GQ1302, 01GQ1509

