NIX
An HDF5-based data file format

Achilleas Koutsou

December 18, 2017

G-Node

- Development and free distribution of tools for handling and neurophysiological data.
- All tools developed within the G-Node are open source and freely available.
- Main projects:
  - NIX: Manage data and metadata together in an open, versatile format.
  - odML: Collect and manage all information about your experiment.
  - GIN: Secure data storage, easy collaboration and publication.

Data, results, and metadata

Data:

- Voltage traces, EEG recordings
- Subject location across time
- Subject preference to selection task

Analysis results:

- Spike times
- Total time the subject spent in an area
- Frequency of correct response over time

Data, results, and metadata

Data:

- Voltage traces, EEG recordings
- Subject location across time
- Subject preference to selection task

Analysis results:

- Spike times
- Total time the subject spent in an area
- Frequency of correct response over time

Metadata:

- Recording equipment
- Subject age, gender, physical attributes
- Date, time, experimenter

Problems

The problem with (not necessarily) old datasets:

- Where did this dataset come from?
- When was it created?
- What were the simulation parameters when I generated these numbers?
- Who ran the experiment?
- What's the meaning of the second row of numbers?

NIX

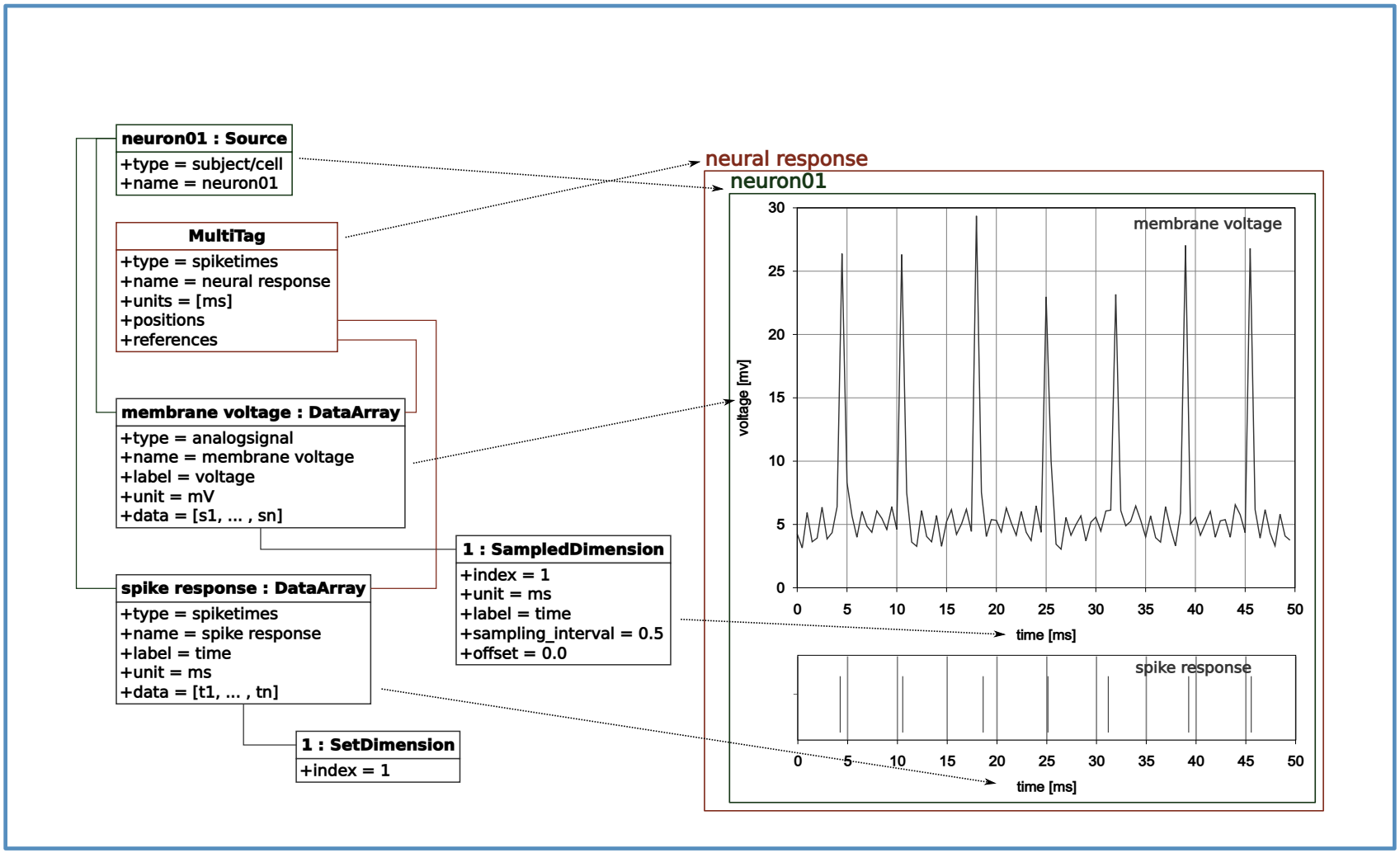
Features

Main features

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

NIX

Object hierarchy schema



NIX

Libraries

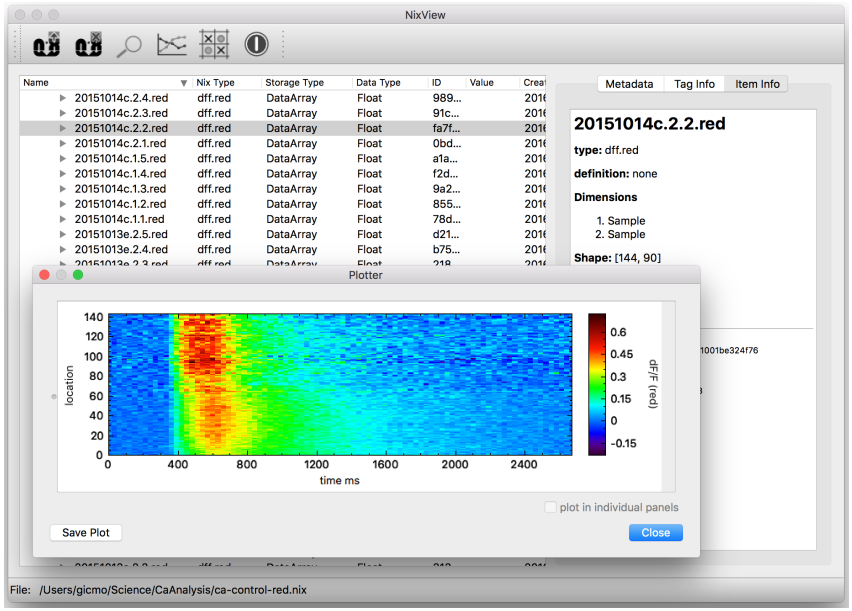
Libraries available for multiple languages

C++	core library and reference implementation.
Python	bindings for core lib as well as complete reimplementation.
Matlab	bindings for core lib.
Java	bindings for core lib.
Neo IO	allows interoperability with Neo, an API for organising electrophysiological data.

NIX

Tooling

NixView — Cross-platform GUI viewer



- Convenient exploration of data and metadata.
- Exports data to CSV.
- Plotting of data.

NIX

Python code

```
import nixio as nix
import numpy as np


data = np.load("bunch-o-numbers.npz")["data"]

# create new nix file (overwrite truncates existing file)
nixfile = nix.File.open("mydata.nix", nix.FileMode.Overwrite)
# block: top level data grouping
block = nixfile.create_block("DataBlock", "experiment")
# data array: data storage object
darray = block.create_data_array("raw-01", "raw data", data=data)
time = darray.append_sampled_dimension(0.1)
time.unit = "ms"
darray.append_set_dimension()

nixfile.close()
```

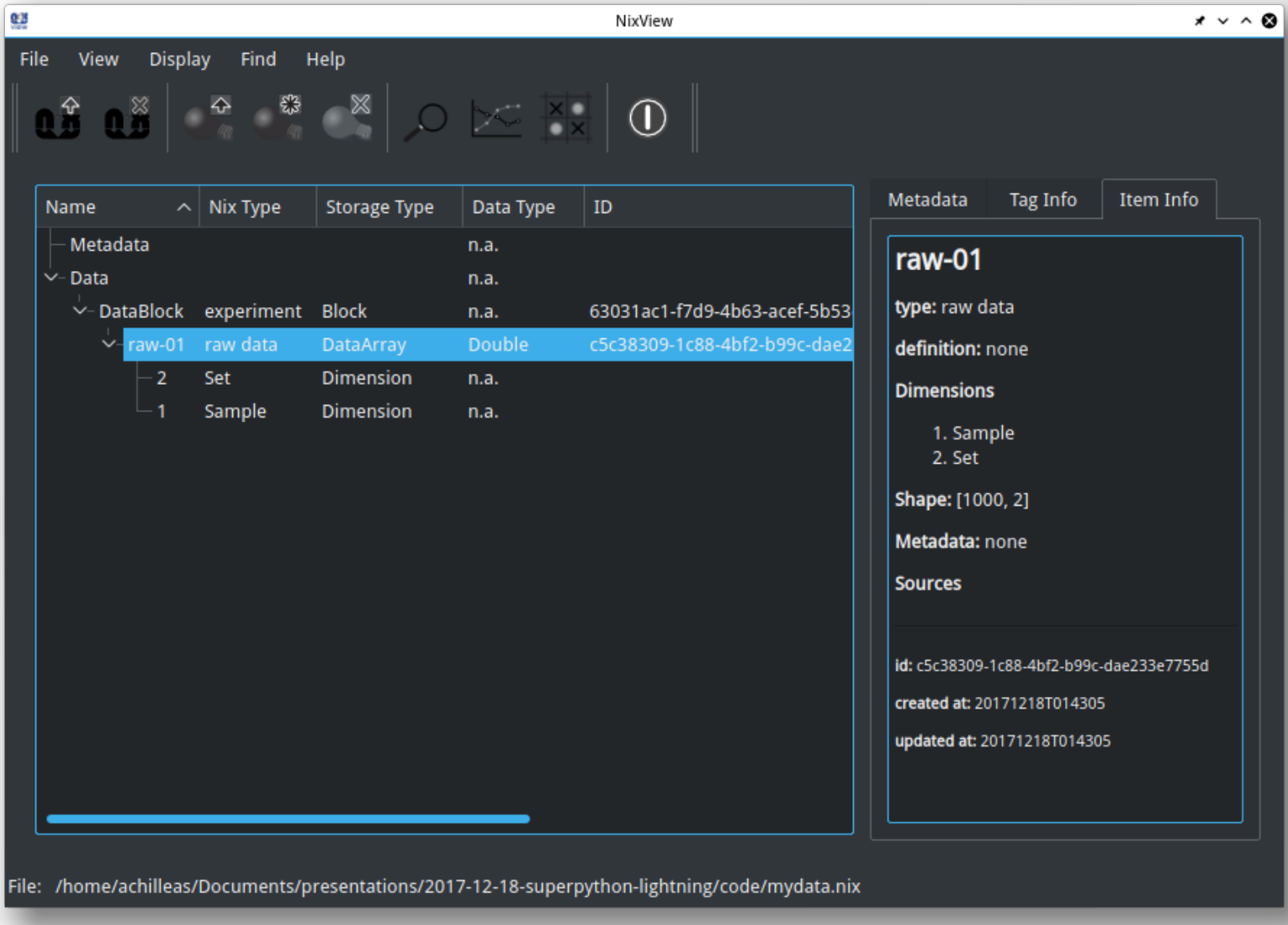
NIX

HDF5 structure

```
/                                Group
/data                            Group
/data/DataBlock                  Group
/data/DataBlock/data_arrays Group
/data/DataBlock/data_arrays/raw-01 Group
/data/DataBlock/data_arrays/raw-01/data Dataset {1000/Inf, 2/Inf}
/data/DataBlock/data_arrays/raw-01/dimensions Group
/data/DataBlock/data_arrays/raw-01/dimensions/1 Group
/data/DataBlock/data_arrays/raw-01/dimensions/2 Group
/metadata                        Group
```

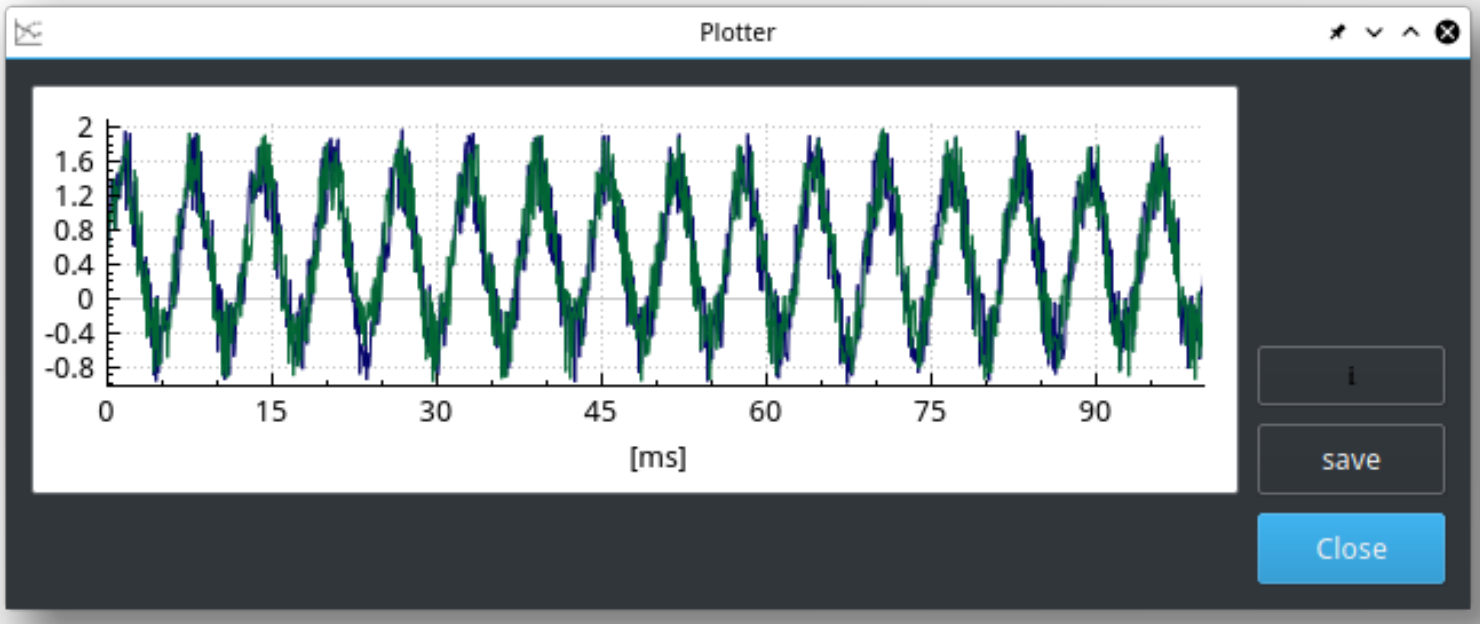
NIX

Data in NixView



NIX

Data in NixView



Resources

- NIX info and documentation: `https://github.com/G-Node/nix/wiki`
- NIX source: `https://github.com/G-Node/nix/`
- NIX Python source: `https://github.com/G-Node/nixpy/`