

SCIENTIFIC DATA

OPEN Data Descriptor: Massively parallel recordings in macaque motor cortex during an instructed delayed reach-to-grasp task

Received: 23 October 2017

Accepted: 30 January 2018

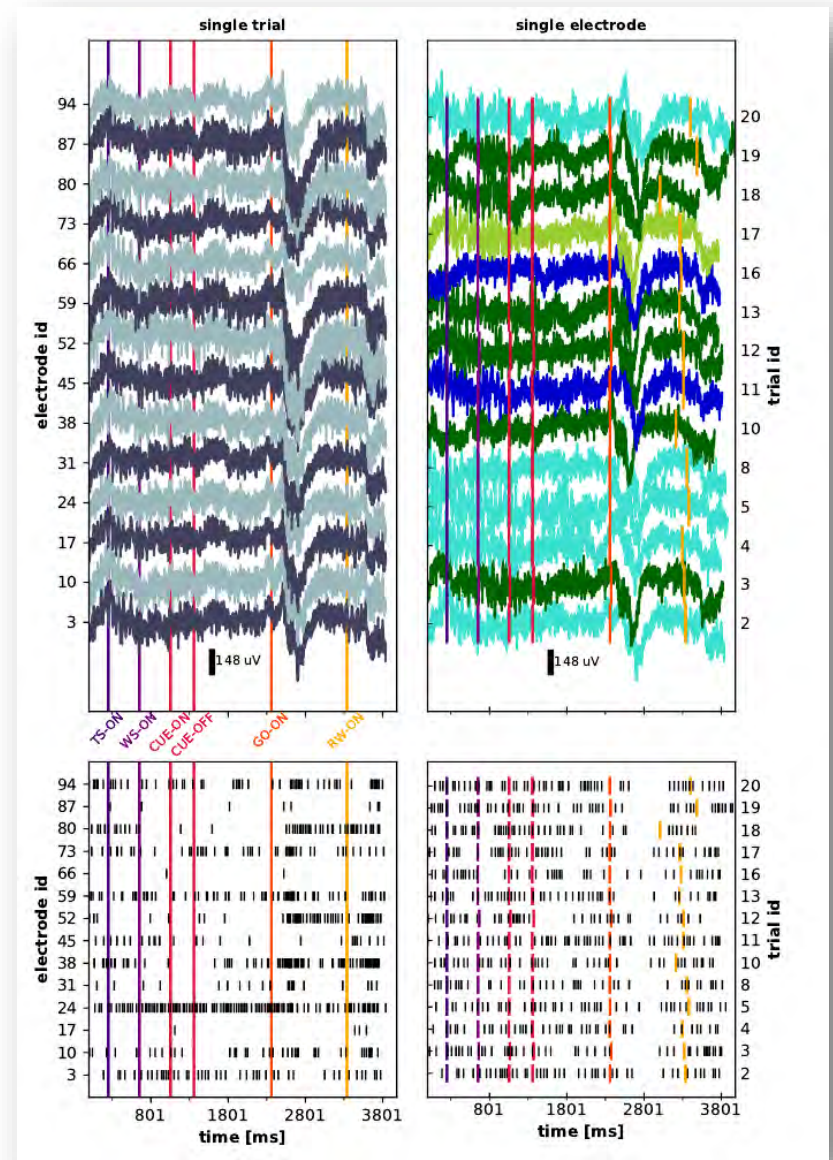
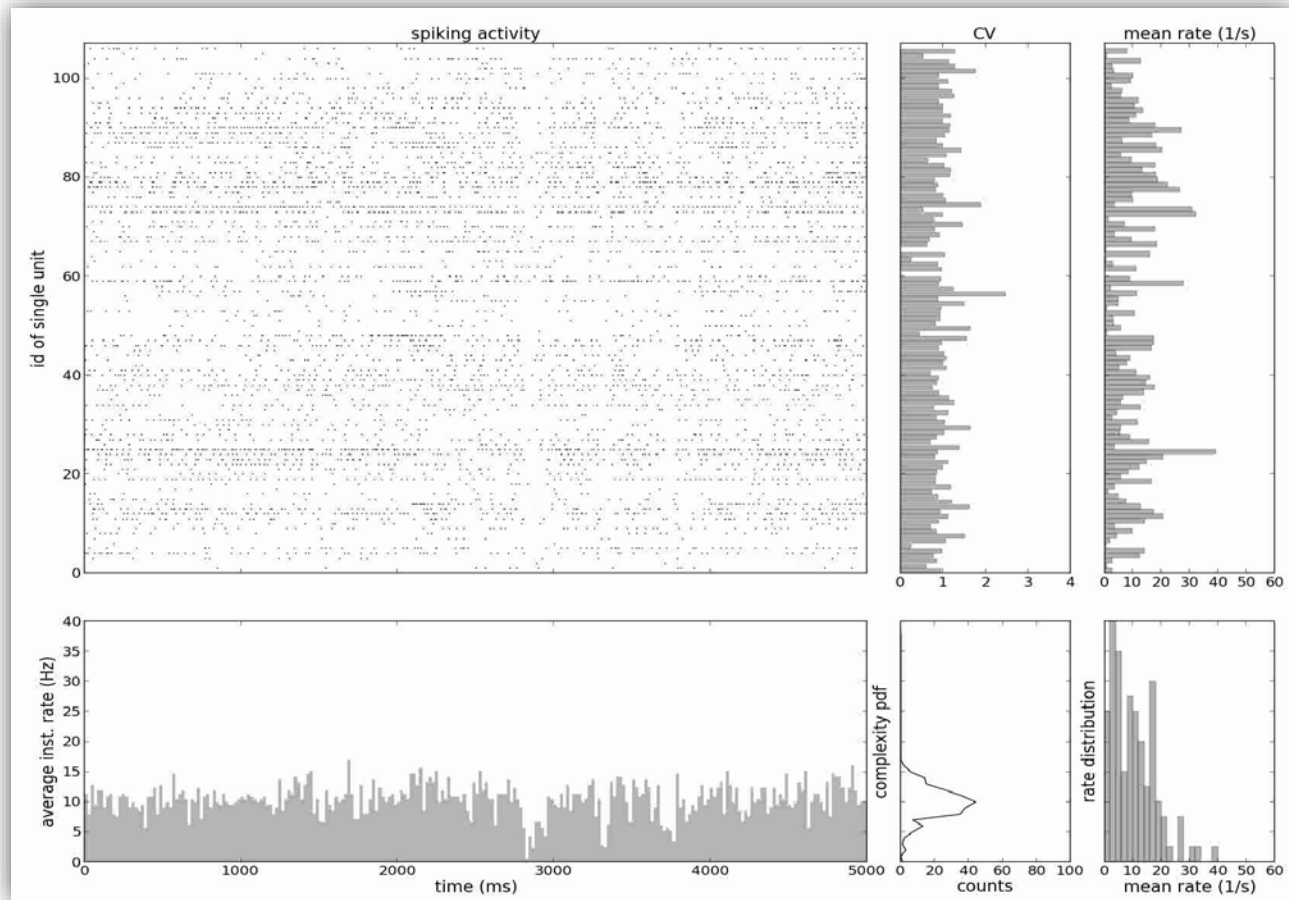
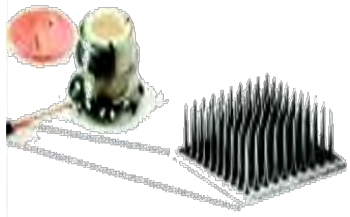
Published: 10 April 2018

Thomas Brochier¹, Lyuba Zehl², Yaoyao Hao¹, Margaux Duret¹, Julia Sprenger², Michael Denker², Sonja Grün^{2,3,4} & Alexa Riehle^{1,2,4}

BUILDING THE RIGHT WORKFLOW FROM THE START FACILITATES PUBLISHING OF NEUROSCIENCE DATA

NOVEMBER 8, 2018 | MICHAEL DENKER | CHALLENGE DATASETS WORKSHOP, FZJ, GERMANY

https://web.gin.g-node.org/INT/multielectrode_grasp



RECORDING DATA IN AN EXPERIMENTAL SETUP

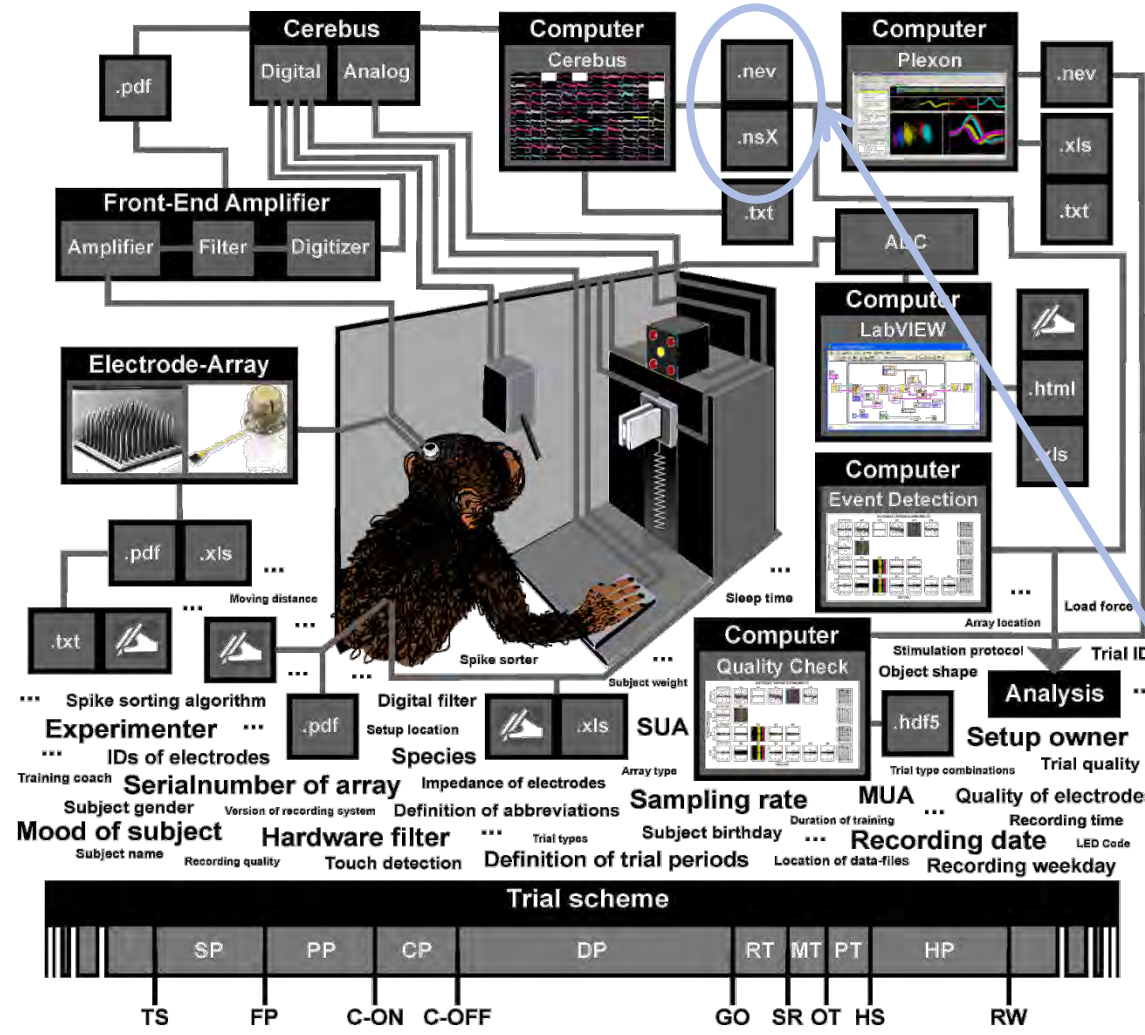
Metadata is often stored in heterogeneous formats, distributed over files, software code, file names, spreadsheets, handwritten ...

Separate organization and storage of data and metadata prevents rigorous identification and selection of data.

Publishing a FAIR dataset implies:

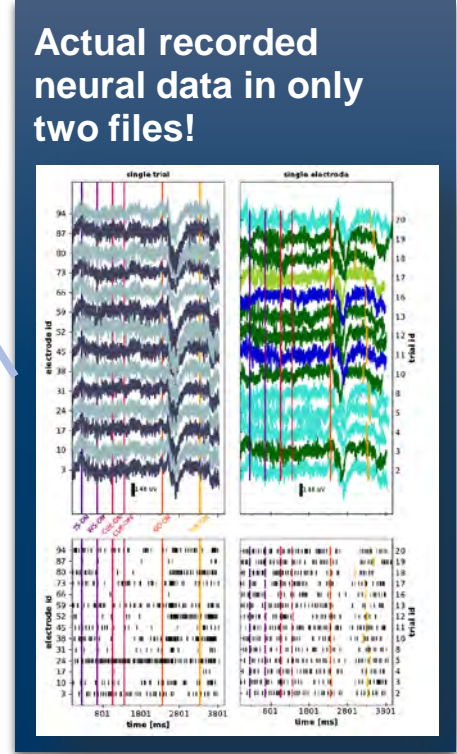
- Creating a single, self-explanatory, inter-linked, rigorous, standardized and complete representation of **heterogeneous data**
- **Tracked control** over the data generation and modification
- **Collaboratively** managing acquisition and preprocessing

Member of the Helmholtz Association



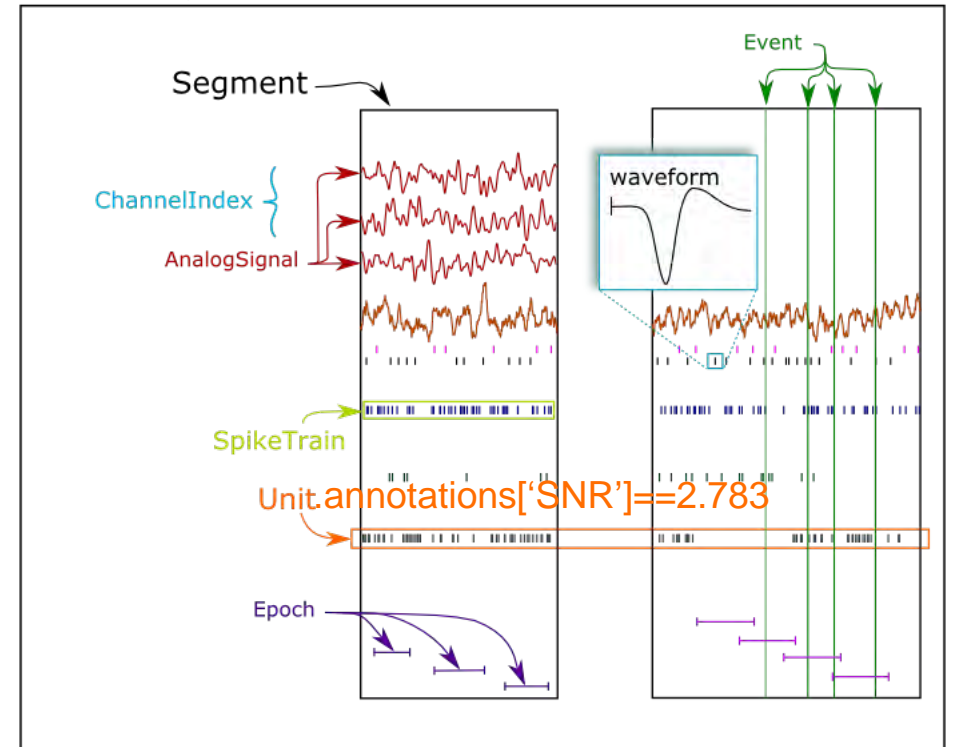
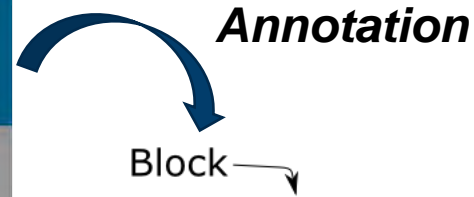
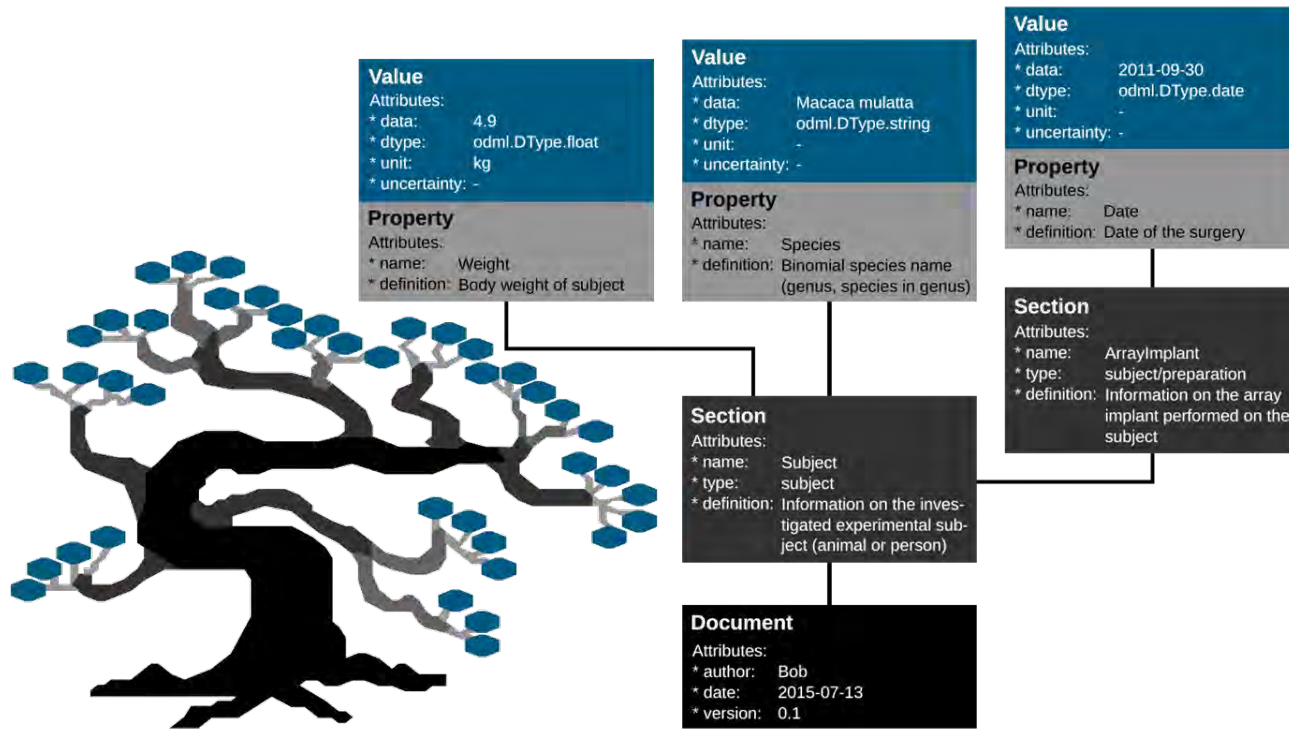
Reach-to-grasp study:

- 120 trials / recording
- ~ 5 recordings / day
- ~ 70 days / monkey
- 3 monkeys



Zehl, Jailet, Stoewer, Grewe, Sobolev Wachtler, Brochier, Riehle, Denker, Grün (2016) Front Neuroinf.

NIX: COMBINING NEO AND ODML



How combine metadata and data?

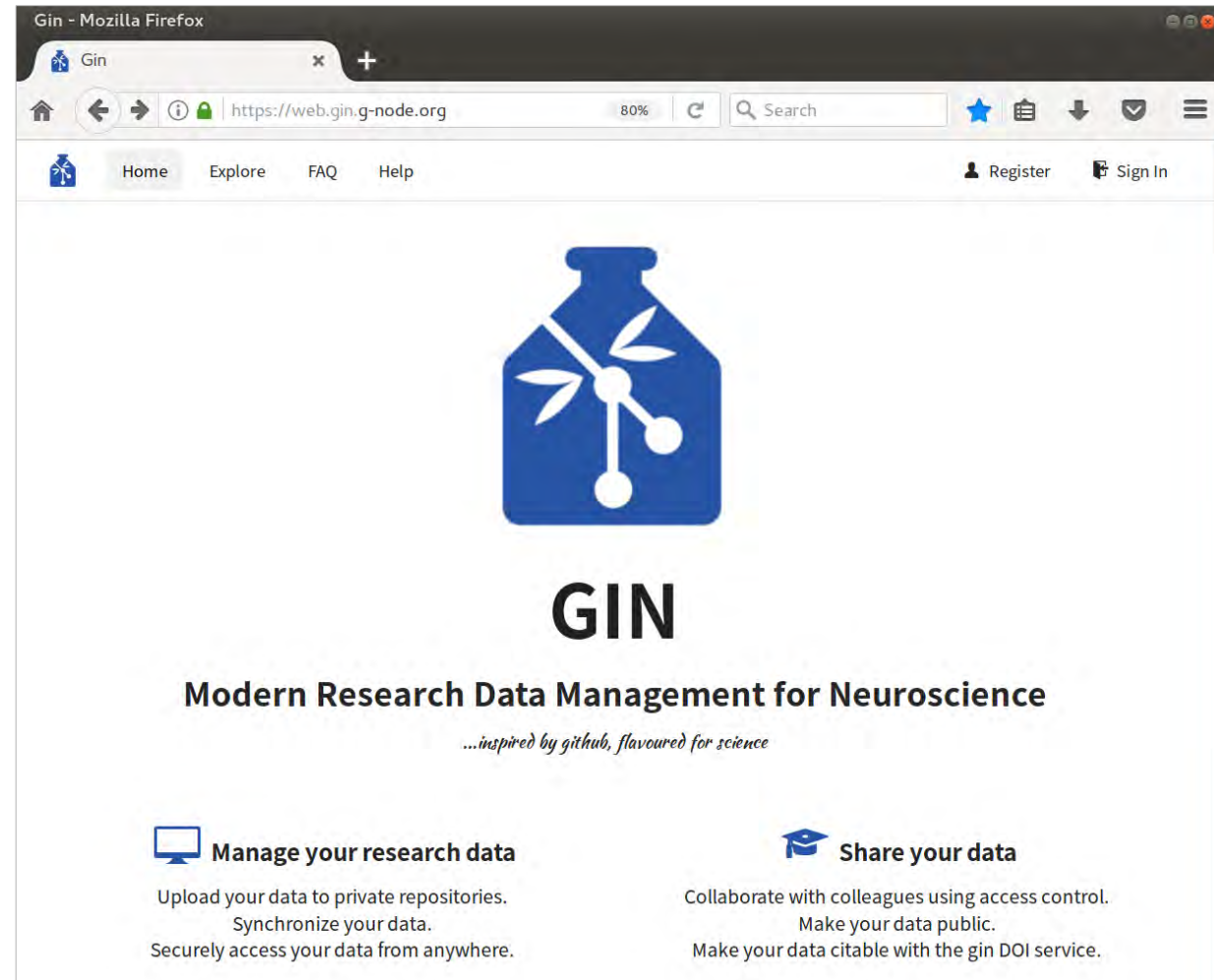
- People have different *mental structures* for metadata and data
- Some metadata not trivial (e.g., SNR for spike sorted unit)

NIX as a common format to store data, metadata and associations

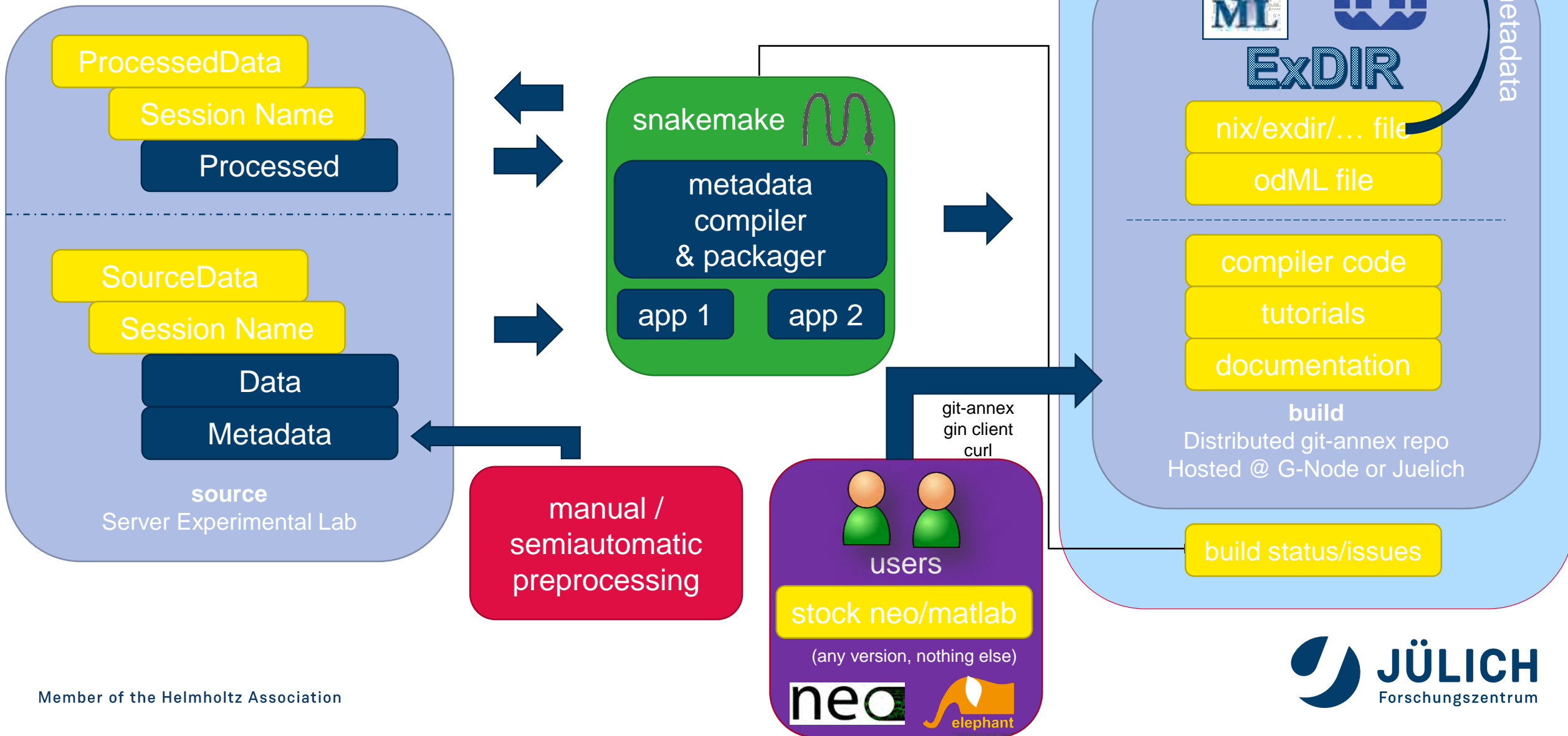
GIN: SERVICES FOR DATA ORGANIZATION AND SHARING

- Versioned management of data repositories (git)
- Distributed data management via file or web browser, or command line client (analysis script integration)
- Secure access and sharing
- Services for search, indexing, and publication (DOI)
- Single place of contact for docs, code, and data

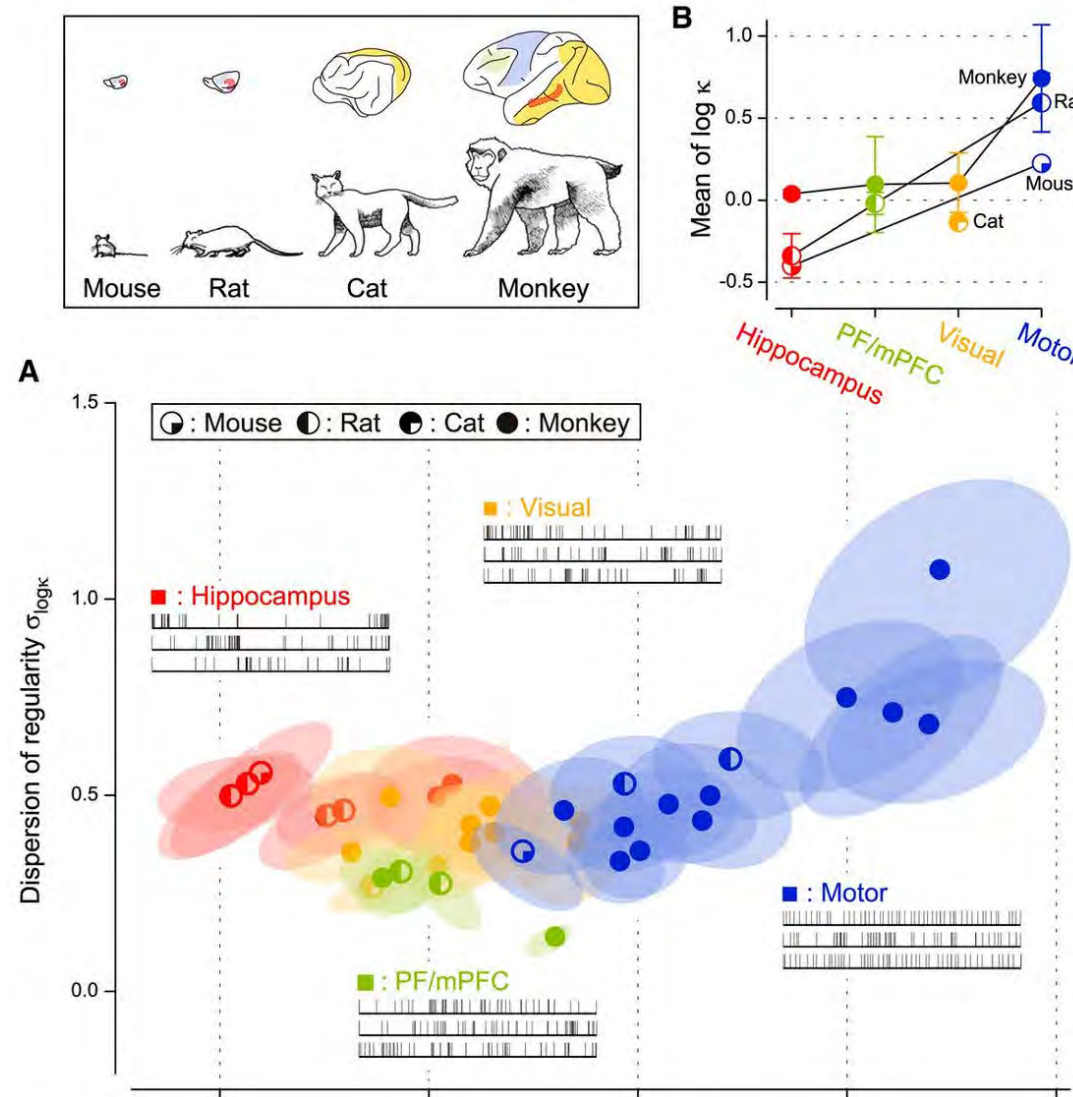
<http://gin.g-node.org>



DATA ANNOTATION AND CROSS-LAB DATA EXCHANGE



FULL ANNOTATION OF DATASET ALLOWS TO CREATE CHALLENGES BY REDUCING DATA FROM MULTIPLE STUDIES



Mochizuki et al (2016)
J Neuroscience, 36 5736-5747

FULL ANNOTATION OF DATASET ALLOWS TO CREATE CHALLENGES BY REDUCING DATA FROM MULTIPLE STUDIES

- Preprocessing challenges (e.g., “spike sorting”)
- Detection of features in the neuronal dynamics
- The ANDA challenge: Which dataset is real? → Surrogate data

