PanNeuro: leveraging a community-based approach for big data neuroscience

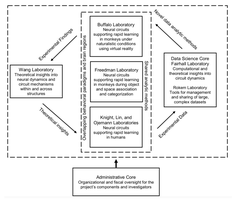
Ariel Rokem[1,2,*], Joe Hamman [3], Ryan Abernathy [4] & Adrienne Fairhall[2,5]

1. The eScience Institute, 2. Computational Neuroscience Center 5. Dept. of Physiology and Biophsyics, Univ. of Washington
3. National Center for Atmospheric Research, 4. Earth and Environmental Sciences, Columbia University
*Contact: arokem@uw.edu | Download: http://arokem.org/presentations/brainpi-panneuro-2019

Introduction

Computational and circuit mechanisms underlying rapid learning

NINDS/NIH U19 funded through the BRAIN Initiative



"Learning to learn"[2]

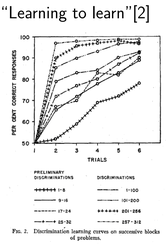
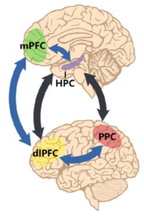


Fig. 2. Discrimination learning curves on successive blocks of problems.

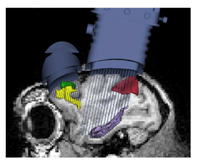


Aims

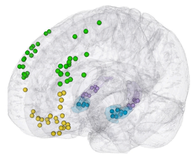
- Identify the neural mechanisms that support schema development and rapid learning in association and categorization paradigms in monkeys and humans.
- Develop and validate novel techniques for large-scale single unit recordings from multiple distributed regions of the nonhuman and human primate brain, during learning, through reversible inactivation, and during sleep.
- Generate and test a multi-region computational understanding of circuit mechanisms that underlie schema development and rapid learning.

High-throughput/resolution recordings in human and non-human primate

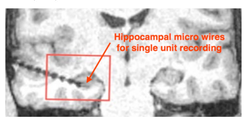
Multi-channel recordings in monkey



ECoG recordings in human

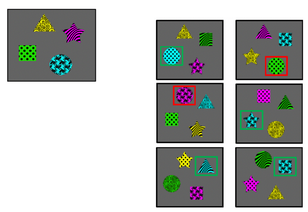


Single/multi-channel recordings in human

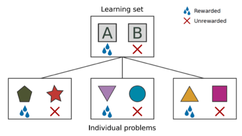


Same tasks used across species

Variation on WCST



Variation on context-dependent learning task [1]



Data types

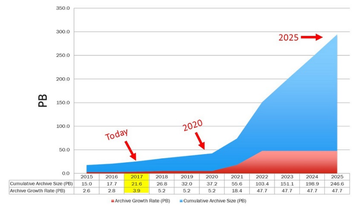
- Behavioral results
- Non-human primate multichannel recordings
- Human grid and electrode recordings
- Model and simulation results

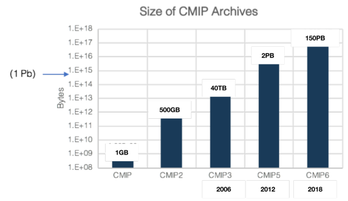
Data volumes

- ∼5-20 TB/week at steady-state
- ∼1-2 PB to store at steady-state
- ∼10%-20% of that needs to be routinely accessed

The *PanNeuro* framework

Inspiration in geoscience big data





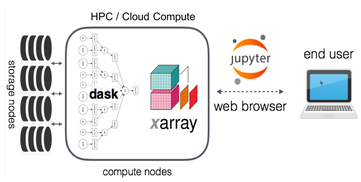
PANGEO

A community platform for Big Data Geoscience

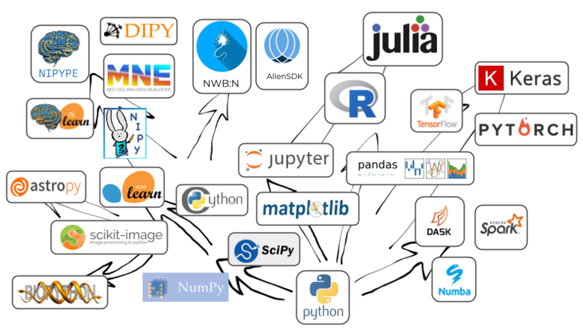
Cloud computing

- Bring the compute to the data.
- Scalable computing.
- Minimize data duplication.
- Reproducibility.

The Pangeo/PanNeuro architecture

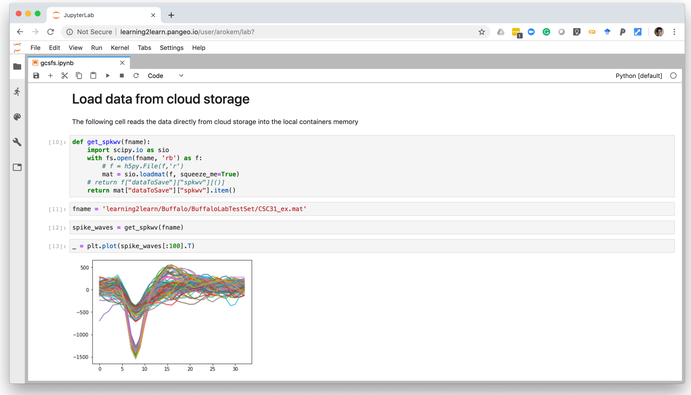


Scientific computing in Python



Jupyter notebook interface

- Shareable, reproducible computational narratives
- Access controlled
- Scalable cluster accessed through the notebook



Try it yourself!

https://learning-2-learn.github.io/panneuro_binder_demo

Barriers to adoption

- Concerns about cost
- Reluctance to share data
- New skills required
- The tools are rapidly evolving
- Data formats and data standardization

References

[1] Timothy EJ Behrens, Timothy H Muller, James CR Whittington, Shirley Mark, Alon B Baram, Kimberly L Stachenfeld, and Zeb Kurth-Nelson. What is a cognitive map? organizing knowledge for flexible behavior. *Neuron*, 100(2):490–509, 2018.

[2] H F Harlow. The formation of learning sets. *Psychol. Rev.*, 56(1):51–65, January 1949.

Acknowledgements

BRAIN INITIATIVE — BRAIN RESEARCH THROUGH ADVANCING INNOVATIVE NEUROTECHNOLOGIES

National Institute of Neurological Disorders and Stroke

Alfred P. Sloan Foundation

Gordon and Betty Moore Foundation

eScience Institute — Advancing Data-Intensive Discovery in All Fields