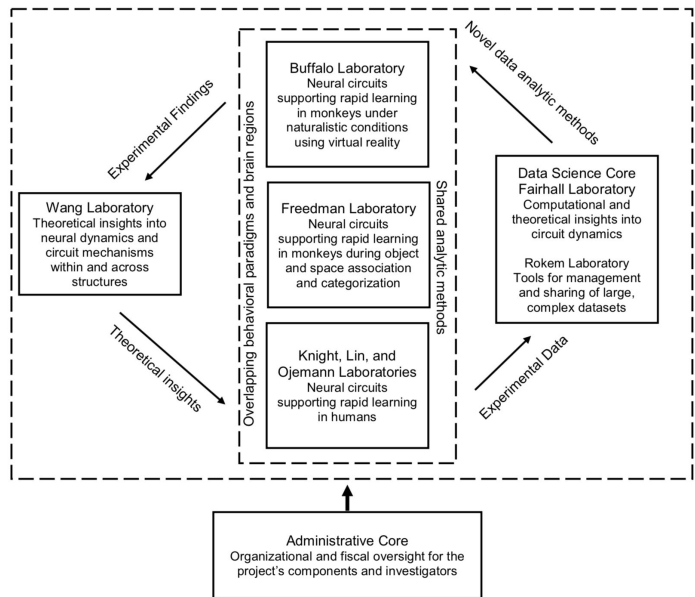


The learning2learn Data Science Core

February 25th, 2020



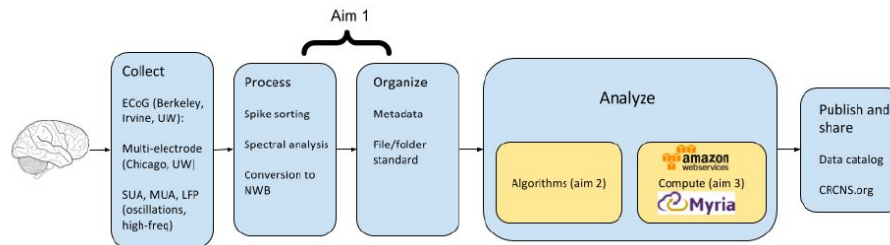
Adrienne Fairhall



Ariel Rokem



John Ferre

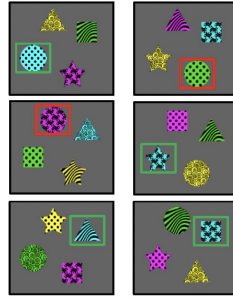
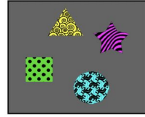
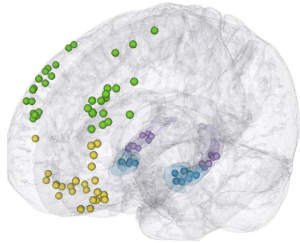
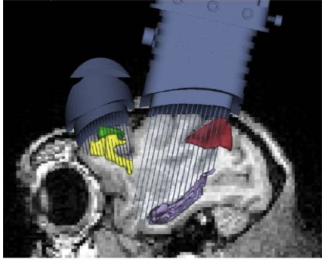


Identify neural mechanisms supporting rapid learning in monkeys and humans.

Develop and validate novel techniques for large-scale single unit recordings from multiple brain regions during learning

Generate and test a multi-region computational understanding of circuit mechanisms that underlie rapid learning.

The data



~5-20 TB/week at steady-state

=> ~1-2 PB to store at steady-state

~10%-20% of that needs to be routinely accessed

Challenges

Computational resources

Statistical methods and algorithms

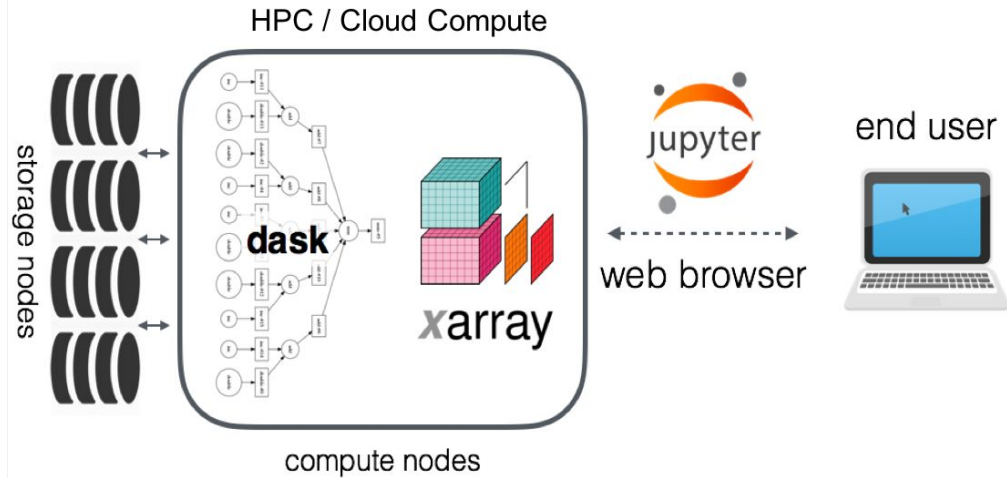
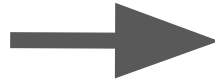
Integration of different data types

Integration of theory and experiments

Reproducibility and extensibility

PANGEO

A community platform for Big Data Geoscience



PanNeuro

The image shows a dual-pane view of a data processing workflow. On the left, a Jupyter Notebook contains Python code for data loading and processing. On the right, the DASK web interface provides a visual overview of the task execution, including progress bars and task stream graphs.

```
project = "learning-2-learn-201802"
fs = get_fs(CSF1) if platform.system() == "linux" else None

folder = "learning/learn/Buttata/buttata/learn/learn/"
files = fs.glob(fs.join(folder, "*.parquet"))

[init_files]

218

data = get_paths(files, project)

data.write()

dask.array.from_delayed(delayed=1, shape=(100000, 30), dtype=float64, chunksize=(10000, 30))

from dask_ml.dask import DaskML

sp_data = DaskML.from_delayed(delayed=1)

sp_data.write()

pct = 0.10

sp_data.compute()
```

The DASK interface displays the following information:

- Bytes stored:** 17.18 GB
- Tasks Processing:** A horizontal bar chart showing the distribution of task sizes.
- Task Stream:** A waterfall chart showing the sequence of tasks over time.
- Worker Group:** A horizontal bar chart showing the distribution of worker sizes.
- Progress:** Total 433, In-memory 175, processing 228, wrote 0. 716/218
- Conclusion:** 00/218