



Mining open access data for teaching electrophysiology

Ashley Juavinett, PhD
Associate Teaching Professor

UC San Diego Neurobiology

Slides &
materials





By the end of this talk you'll be able to:



- Identify multiple ways in which open access data can be used to teach basic principles of electrophysiology
- Access **Allen Institute** cell types data using Google Colab
- Expand your teaching (and research?) with **Neurodata Without Borders** datasets

Slides &
materials



Why use open access data?

- Limited time and resources in undergraduate teaching laboratories
 - With open data we can conduct experiments on data we did not collect!
 - Developing skills to work with open access data are broadly useful in both research and the workforce
(Juavinett, 2022; Shah & Juavinett, 2022, Grisham et al., 2016, and many others)



Why Allen Institute data?

- Well curated with very detailed metadata
- High quality media for students
- Dedicated team to educational efforts & responsive to technical questions
(<https://community.brain-map.org>)



The collage includes:

- A photograph of a scientist in a lab coat and gloves working at a bench.
- A photograph of two scientists, one male and one female, looking at a screen or document together.
- A photograph of a scientist from behind, working at a desk.
- A screenshot of a YouTube video player for 'ALLEN CELL TYPES DATABASE'. The video shows a collection of colorful, branching neural cell models against a black background. The title reads 'ALLEN CELL TYPES DATABASE' and 'UNDERSTANDING THE FUNDAMENTAL BUILDING BLOCKS OF THE BRAIN'. Below the video are standard YouTube controls: play, volume, progress bar (0:00 / 4:34), and a 'Intro' link. To the right of the video are icons for CC (Creative Commons), a gear (settings), and a square (share).
- A circular logo for 'NEUROSCIENCE 2023' featuring a stylized brain cell with yellow and red branches.

Allen Cell Types Database: Understanding the fundamental building blocks of the brain



Subscribe



95



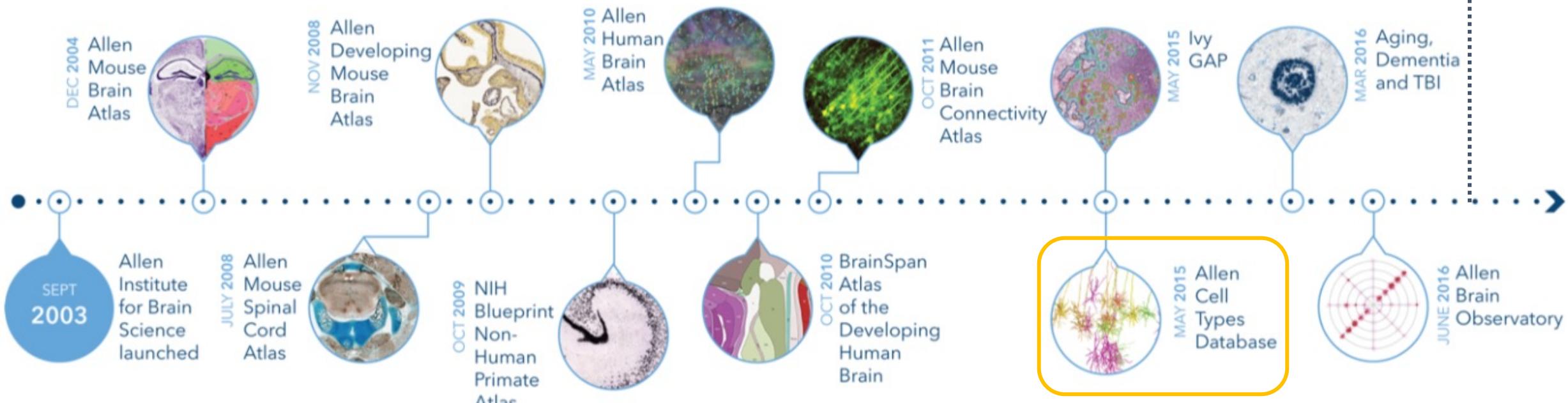
Share



...

Many different data are available (brain-map.org):

+ Neuropixels
& behavior
data

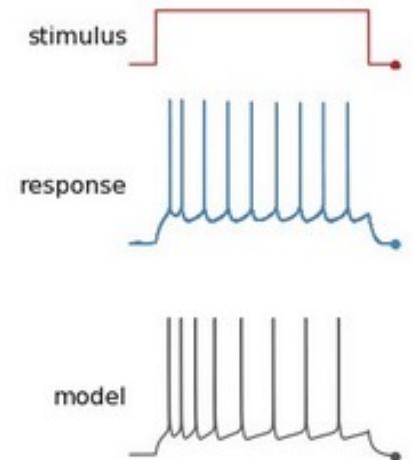
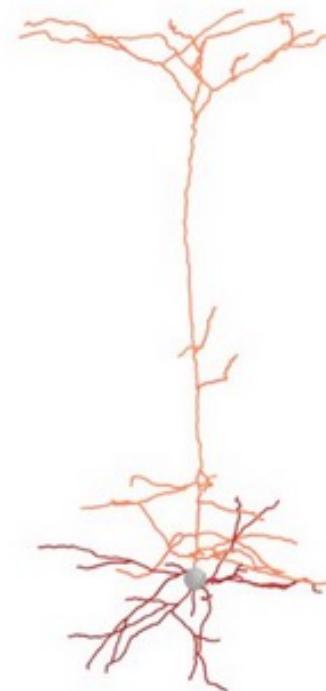


Focus for today



About the cell types database

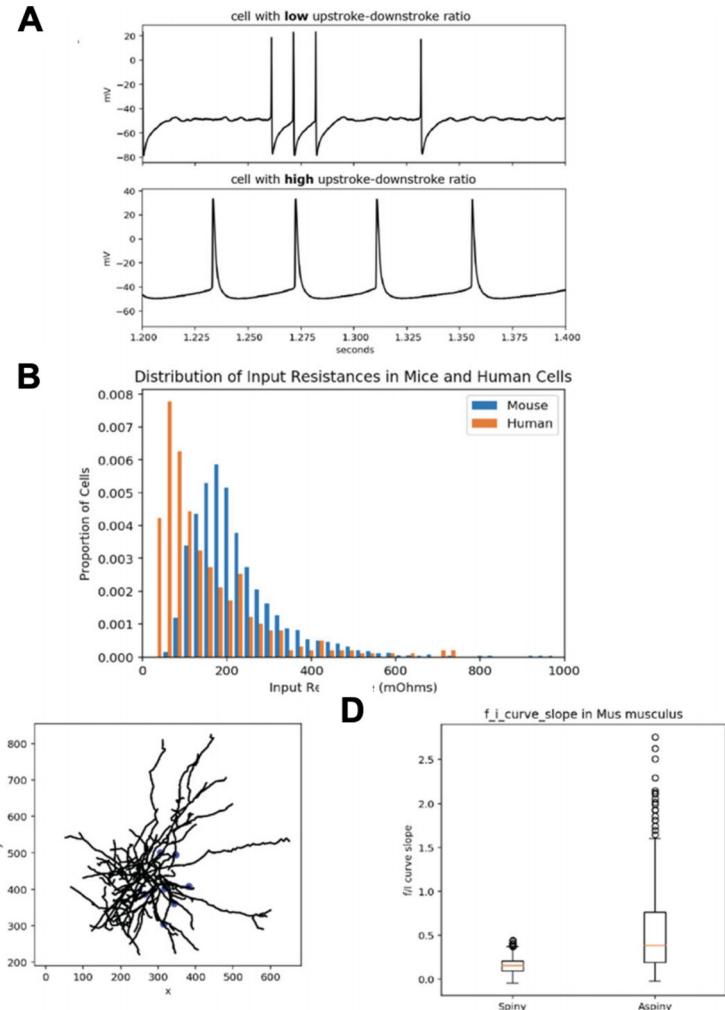
- Whole-cell patch clamp data from mouse and humans
- Many **pre-computed features**, such as tau, input resistance, adaptation ratio, average ISI, etc.
- Also includes **metadata** about cells: human/mice, transgenic line, cell size, layer of cortex



Example #1

Cell Types Lesson Objectives:

- Relate the diverse intrinsic electrophysiological features of neurons to their structure and function
- Compare electrophysiological characteristics of neurons in humans and mice
- Practice using coding notebooks to run and edit Python code
- Develop a sense of belonging and self-efficacy in coding and neuroscience research



Juavinett, "Learning to code while analyzing an open access dataset," JUNE, 2020

How do cell features correlate, and why?

Before this lesson, students should be able to:

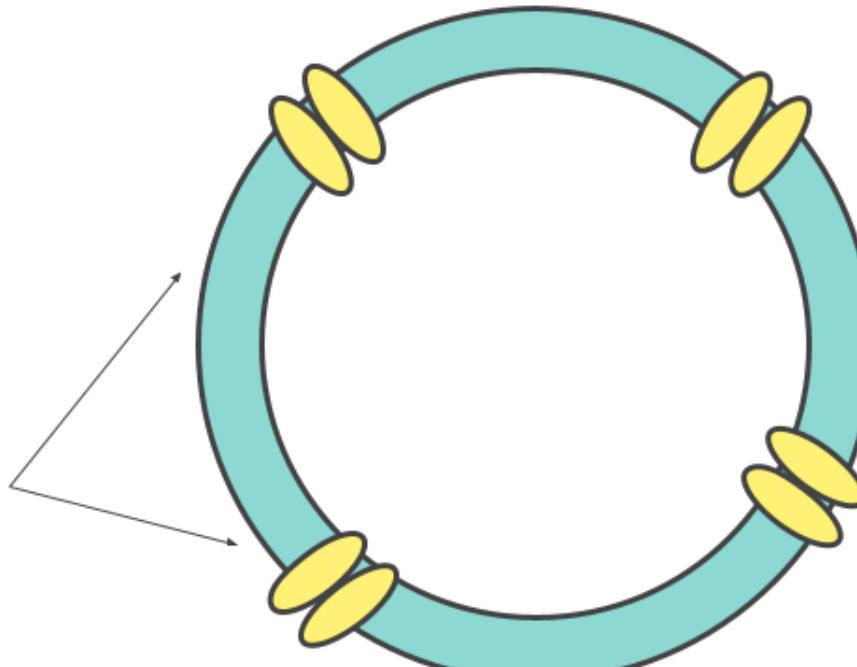
- Describe how patch clamp recordings enable us to obtain high fidelity signals from neurons
- Define time constant, input resistance, and rheobase
- Identify ways that cell types differ, structurally and physiologically



Either before or after the lesson, you can explain the biophysics behind how these properties relate



Assumption #1:
The membrane potential is uniform at all points of the sphere (the cell is **isopotential**)



Assumption #2:
The specific membrane resistance (R_m is constant and independent of voltage)

We'll treat the neural membrane as an **isopotential sphere**.

How do cell features correlate, and why?

Before this lesson, students should be able to:

- Describe how patch clamp recordings enable us to obtain high fidelity signals from neurons
- Define time constant, input resistance, and rheobase
- Identify ways that cell types differ, structurally and physiologically

After this lesson, students will be able to:

- Investigate relationships between surface area, rheobase, input resistance, and time constant
- Generate a scatterplot in Python
- Explain how and why cell features are correlated
- Appreciate the variability in real data and the importance of statistics



The screenshot shows a GitHub repository page for 'ajuavinett/FUN_2023'. The repository is public and contains 1 branch and 0 tags. The main branch has 19 commits. The repository description is 'Materials for Faculty for Undergraduate Neuroscience workshop (2023)'. It includes sections for Releases (no releases published), Packages (no packages published), and a README.md file.

ajuavinett / FUN_2023

Type ⌘ to search

Code Issues Pull requests Actions Projects Wiki Security Insights Settings

FUN_2023 Public

Pin Unwatch 1 Fork 0 Star 1

main 1 branch 0 tags

Go to file Add file Code

ajuavinett Merge branch 'main' of https://github.com/ajuavinett/FUN_2023 into main ee2d82e last week 19 commits

Data one more paper + human/mouse data 4 months ago

Related Papers add rubel paper 4 months ago

01-Introduction.ipynb rename and add spike sorting 4 months ago

02-CellTypes.ipynb clean up celltypes 4 months ago

03-SpikeSorting.ipynb remove references to textbook 4 months ago

04-CompareCellFeatures.ipynb revise compare cell features for sfn last week

FUN Workshop 2023.pdf add final slide 4 months ago

README.md Update README.md 4 months ago

README.md

About

Materials for Faculty for Undergraduate Neuroscience workshop (2023)

Readme Activity 1 star 1 watching 0 forks

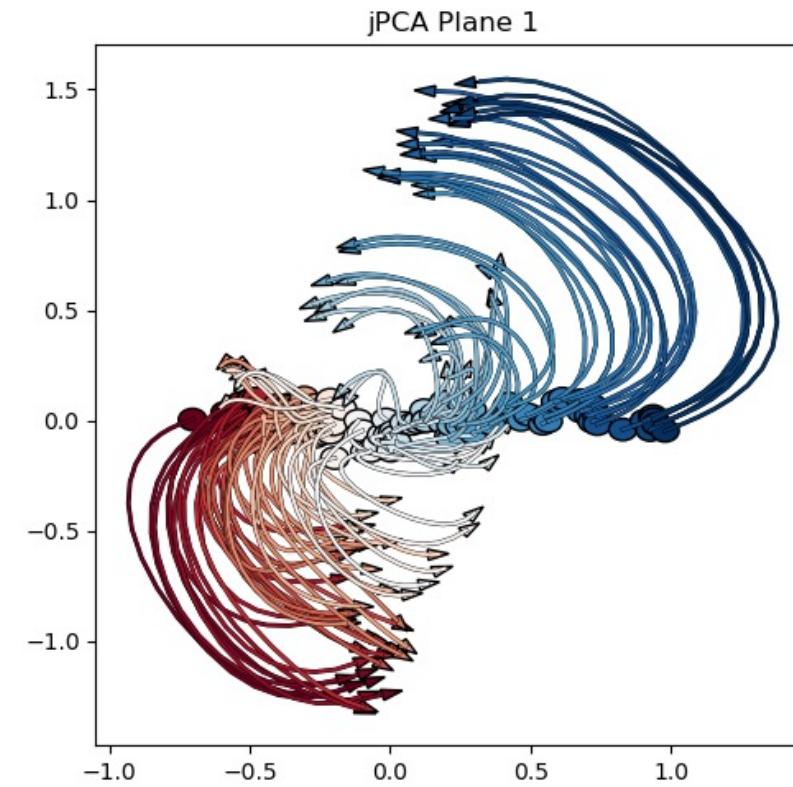
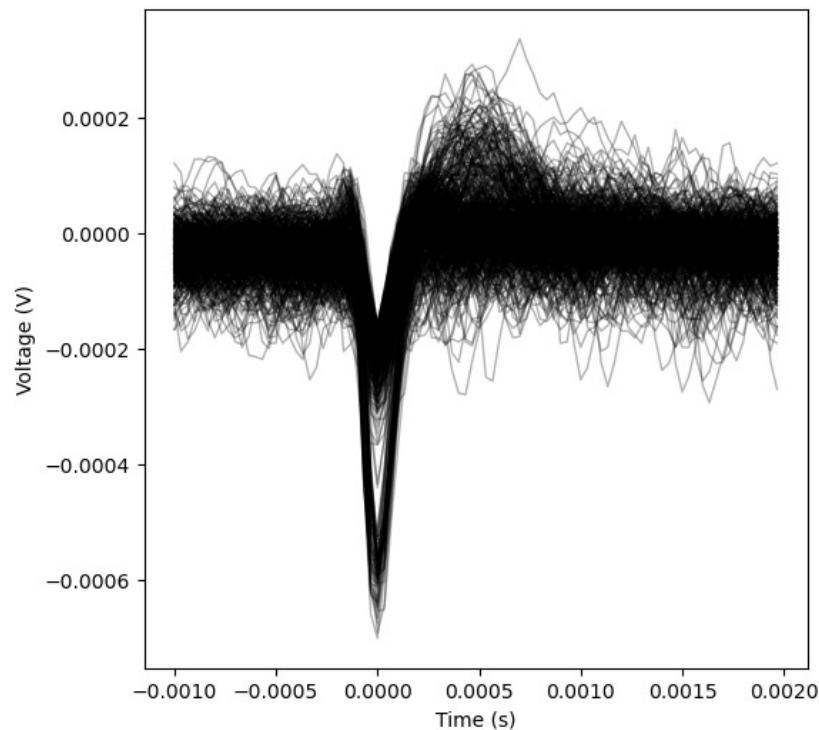
Releases

No releases published Create a new release

Packages

No packages published Publish your first package

Using Neurodata Without Borders datasets to teach more advanced topics, from spike sorting to dimensionality reduction



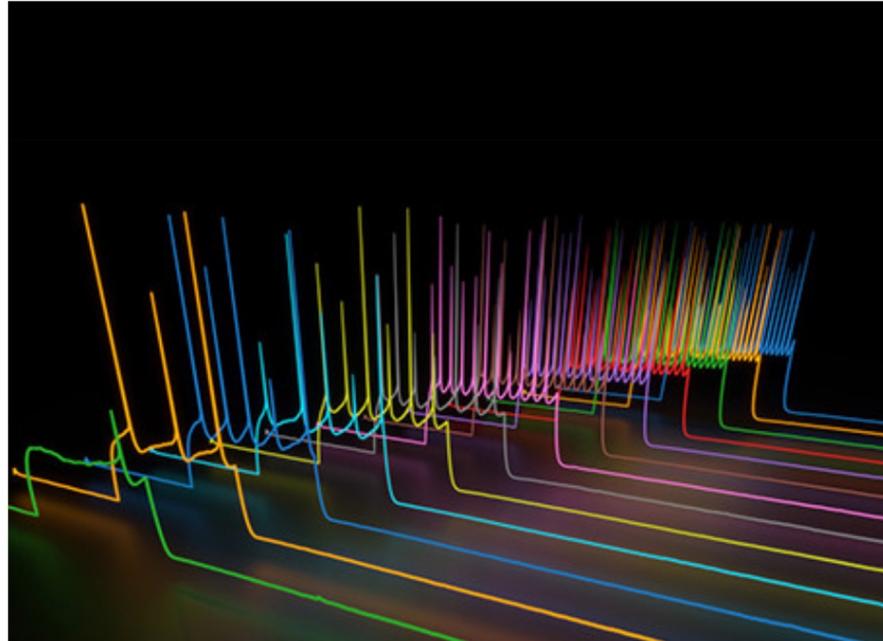
Neurodata Without Borders

Neurodata Without Borders (NWB) is a data standard for neurophysiology, providing neuroscientists with a common standard to share, archive, use, and build analysis tools for neurophysiology data. NWB is designed to store a variety of neurophysiology data, including data from intracellular and extracellular electrophysiology experiments, data from optical physiology experiments, and tracking and stimulus data.

The **NWB team** consists of neuroscientists and software developers who recognize that adoption of a unified data format is an important step toward breaking down the barriers to data sharing in neuroscience.



www.nwb.org



(Courtesy of the Allen Institute for Brain Science)

See also Rübel et al. (2022):
<https://elifesciences.org/articles/78362>

01-SpikeSort... (3) - JupyterLab X

https://hub.dandiarchive.org/user/ajuavinett/lab/tree/nwb4edu-dandi/Lesson_2/01-SpikeSorting.ipynb

File Edit View Run Kernel Tabs Settings Help

RotationalDynamics.ipynb 01-SpikeSorting.ipynb 000582_Sargolini2006_dei.ipynb

Code Open in... Python 3 (ipykernel)

Filter files by name / nwb4edu-dandi / Lesson_2 /

Name	Last Modified
00-LargeE...	4 months ago
01-SpikeS...	36 seconds ago
02-Visuali...	4 months ago
Lesson_2....	4 months ago
Lesson_2....	4 months ago
requireme...	4 months ago
spikesortin...	4 months ago

Spike Sorting

When you record extracellular electrophysiology data, one of the first data processing steps is figuring out which action potentials (or "spikes") came from which neurons. The process of doing this is called **spike sorting**.

Below, we'll work with [this dataset](#) from Lisa Giocomo's lab at Stanford to demonstrate the simplest form of spike sorting: thresholding, followed by feature extraction.

Note: The code below requires a different dataset than the one we interacted with in the last chapter. Because this dataset contains all of the raw recording data, it is much, much bigger. As a result, the best way to work with it is through the Dandihub. If you're not already running this book on the Dandihub, read "Using this Book for instructions on how to run this on the Dandihub."

Step 1. Inspect the Data

First, we need to find the correct URL for the dataset on the NWB's Amazon S3 storage system. There is a tool to do so within the dandiapi, which we'll use below to get the URL for one session within the dataset.

```
[1]: from dandi.dandiapi import DandiAPIClient  
  
dandiset_id = '000053' # giocomo data  
filepath = 'sub-npI5/sub-npI5_ses-20190414_behavior+ecephys.nwb'  
  
with DandiAPIClient() as client:  
    asset = client.get_dandiset(dandiset_id, 'draft').get_asset_by_path(filepath)  
    s3_path = asset.get_content_url(follow_redirects=1, strip_query=True)
```

Simple 0 3 Python 3 (ipykernel) | Idle Mode: Edit Ln 3, Col 10 01-SpikeSorting.ipynb 1



Online lessons using NWB datasets at nwb4edu.github.io

The screenshot shows a web page titled "Teaching and Learning with NWB Datasets". The page features a sidebar on the left with a navigation menu:

- Teaching and Learning with NWB Datasets** (highlighted)
- INTRODUCTION**
 - Introduction
 - Data Science in Python
- PYNWB**
 - Lesson #1: The Structure of NWB Datasets
 - Lesson #2: Analyzing Extracellular Recordings
- ALLENSDK**
 - Lesson #3: Single-Cell Electrophysiology

The main content area contains the following text:

For many, many years in neuroscience, the landscape of file formats has been the wild west. Everyone either uses a format either dictated by their lab or by the equipment they're using. This poses a huge challenge to data sharing: if you want to collaborate with the lab next door and share data with them, one of the first questions is, "what format is your data in?" While static images taken from a microscope and standard biology lab procedures like western blots are fairly straightforward to share, physiology and behavior data types are much more diverse.

That's where [Neurodata Without Borders \(NWB\)](#) comes in. NWB is a consortium of researchers who are aiming to *standardize neuroscience data on an international scale*. The use of this standard dataset format will make sharing and contributing to open source projects easier, ultimately accelerating discovery.

NWB has pioneered a data format called [Neurodata Without Borders: Neurophysiology \(NWB:N\)](#). NWB:N is a standard to share, store, and build analysis tools for neuroscience data. It is built on [HDF5 \(hierarchical data format\)](#), which was designed for large data storage and multi-language accessibility. In other words, it's good at doing exactly what we need it to do.

Standardized data for education

A standardized and accessible data format like NWB also opens possibilities for more easily integrating these cutting-edge datasets into the classroom. This online textbook aims to provide interactive tools, resources, and assessments for basic electrophysiology concepts.

THE KAVLI FOUNDATION



Teaching Neuroscience: New Approaches to Electrophysiology Labs

By the end of this talk you'll be able to:



- ✓ Identify multiple ways in which open access data can be used to teach basic principles of electrophysiology
- ✓ Access **Allen Institute** cell types data using Google Colab
- ✓ Expand your teaching (and research?) with **Neurodata Without Borders** datasets

Slides &
materials



Thank you!

Everything you need to get started is here:
https://github.com/ajuavinett/FUN_2023/



Ask me about teaching ~40
undergraduates to do sharp
intracellular recordings and
dye fills in leech ganglia!

