2023 NEURODATA WITHOUT BORDERS

User Days
July 24-26, 2023

Developer Days
July 27-29, 2023

Report:
2023 NWB User Days and Developer Days at HHMI Janelia

15th and 16th NWB Hackathon: July 24 - 29, 2023

Benjamin Dichter, Ryan Ly, Jakob Voigt, and Oliver Rübel

Contents

2

1 Executive Summary

The Neurodata Without Borders (NWB) project is an effort to standardize the description and storage of neurophysiology data and metadata. NWB enables data sharing and reuse and reduces the energy barrier to applying data analytics both within and across labs. The 2023 NWB User Days training workshop and NWB Developer Days development hackathon events were held back-to-back at the HHMI Janelia Research Campus in Ashburn to facilitate interaction between users and developers and help users inform NWB developments and needs. During the two events, the **38 attendees from 28 different institutions (see Sec. 2) worked on 32+ different projects (see Sec. 4)** ranging from converting data to NWB, integrating tools with NWB, development of new features and extension of the NWB data standard and software, to data reuse and education.

Participants: 38 unique participants from 28 different main institutions attended the two workshops in person. 29 participants attended the User Days and 28 attended the Developer Days. Of the participants: i) 19 attended both events, ii) 10 attended only the User Days, and iii) 9 attended only the Developer Days. In addition, 3 additional participants presented virtually in select breakout sessions. Current positions of registrants were divided across: Engineer (14), Researcher (12), Post Doctoral Researcher (7), Ph.D. Student (6), Data Scientist (3), Research Assistant (3), and Undergraduate Student (1). For further details about participants see Sec. 2.

Program: Building on the success and experience from previous NWB workshops, the program during the User Days was designed to provide a balance between tutorials in the morning while providing dedicated time for hacking on projects in the afternoon. All presentations were optional and a second room for hacking was available to participants at all times. The program during the Developer Days then consisted of breakout sessions for joint discussions while emphasizing open time for collaboration and working on projects. Opportunities for social interaction and team discussions were available throughout the event during the open hacking sessions, during the provided group breakfast, lunch, and dinner, as well as during the reception and socials on Day 1 and 3 and after the event in Bob's bar.

Lodging & Travel Housing was provided onsite by HHMI Janelia for all participants. 10 registrants requested additional travel support for flights, with travel support being provided by HHMI Janelia.

Conclusion: The NWB User Days and Developer Days were successful in training new NWB users and developers. 67.5% of all participants attended an NWB workshop for the first time, indicating that the workshops continue to attract new users. The workshops were overall well-received (see Sec. 5) and helped further connectedness among participants both socially and in the interconnectivity of the data and software they are developing. Participants made significant progress on a large number (32+) of diverse projects across both events (see Sec. 4). This was also the first time that the NWB Technical Advisory Board (TAB) (est. fall 2022) participated in the workshops and was able to gather valuable feedback for guiding ongoing and future technical NWB developments.

Organizing Committee:

- **Site Chair:** Jakob Voigt (HHMI Janelia)
- **Program Chairs (User Days:** Oliver Rübel (LBNL) and Benjamin Dichter (CatalystNeuro)
- **Program Chairs (Developer Days:** Oliver Rübel (LBNL) and Ryan Ly (LBNL)
- **Janelia Event Support:** Janine Stevens and Alethea Vandamm

Related Documents:

- Event Website
- User Days Project GoogleDoc
- Developer Days Project GoogleDoc

Figure 1: NWB User and Developer Days 2023 group photo taken during the social on Day 3. (Photo by Megan Zipperer, HHMI Janelia).

2 Participants

38 unique participants from 28 different main institutions attended the two workshops in person, of which i) 19 attended both events, ii) 10 attended only the User Days, and iii) 9 attended only the Developer Days. The table in Figure 2 shows a summary of all participants as well as for the User Days and Developer Days workshop, respectively. 67.5% of all participants (i.e., including the organizers and tutors) attended a NWB workshop for the first time and 32.5% had attended previously.

- **User Days:** 29 participants from 24 different institutions attended the User Days. The breakdown of current positions of participants at the User Days is shown in Figure 3 (top left).
- **Developer Days:** 28 participants from 21 different institutions attended the Developer Days. The breakdown of current positions of participants at the User Days is shown in Figure 3 (top right).

All participants indicated their current position as part of the registration. Some (8 or less) participants indicated multiple positions, resulting in 46 responses from 38 participants. Current positions of registrants across the two events were divided across: Engineer (14), Researcher (12), Post Doctoral Researcher (7), Ph.D. Student (6), Data Scientist (3), Research Assistant (3), and Undergraduate Student (1).

The registration form also included optional questions for participants to indicate their gender and ethnicity. 37 of 38 participants responded. Figure 4 shows the distribution of gender and ethnicity of participants across both workshops. In terms of programming language, participants indicated to be most comfortable with Python ($\approx 92\%$) and MATLAB ($\approx 50\%$) (and $\approx 5\%$ indicated C/C++, R, and JavaScript).

Country	State	Institution	Count	User Days	Dev. Days
USA	CA	CatalystNeuro	2	2	2
	CA	Lawrence Berkeley National Lab	4	3	4
	CA	Cedars-Sinai Medical Center	1	1	1
	CA	UCSF	2	1	2
	CT	University of Connecticut, New York	1	1	
	DE	University of Delaware	1	1	
	GA	Georgia Institute of Technology	1	1	1
	GA	Emory University	1	1	
	IL	University of Chicago	1	1	
	IN	CatalystNeuro	1	1	1
	MA	MIT	2	1	1
	MA	Harvard Medical School	1		1
	MA	Harvard	1		1
	MD	Johns Hopkins University	1	1	
	MI	University of Michigan	1	1	1
	MN	Mayo Clinic	1	1	1
	NY	Simons Foundation	1	1	1
	NY	University of Rochester	1	1	
	RI	Brown University	1	1	
	TN	St Jude Children's Research Hospit	1		1
	TX	DataJoint	2	1	2
	TX	University of Houston	1	1	1
	VA	MBF Bioscience	1	1	1
	VA	HHMI Janelia	3	3	1
	WA	Allen Institute for Neural Dynamics	2		2
Hungary		CatalystNeuro	1	1	1
Germany		MPI for Neurobiology of Behavior	1	1	1
United Kingdom		Durham University	1	1	1
		Total	**38**	**29**	**28**

Figure 2: Number of participants per institutions. In addition, 3 participants presented virtually in select sessions during the User Days.

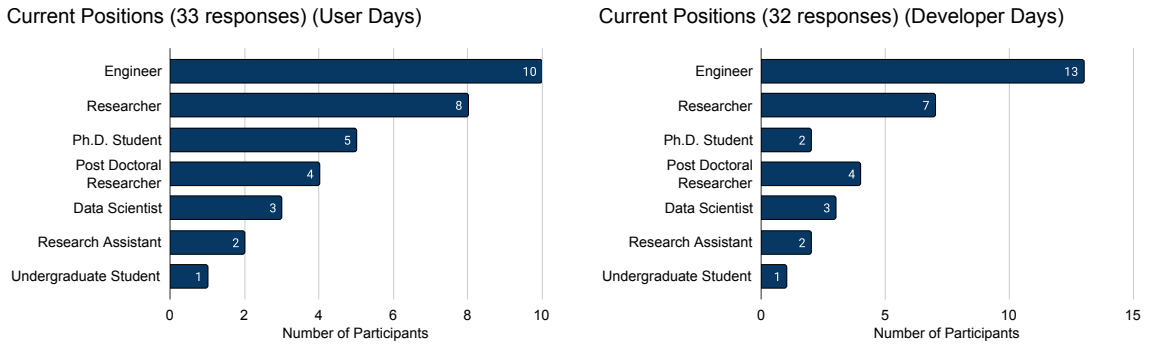


Figure 3: Current positions of participants of the NWB User Days (top left) and NWB Developer Days (top right). All participants indicated their current position as part of the registration, with some (8 or less) indicating multiple positions. As a result, there are 33 responses from 29 User Days participants and 32 responses from 28 Developer Days participants.

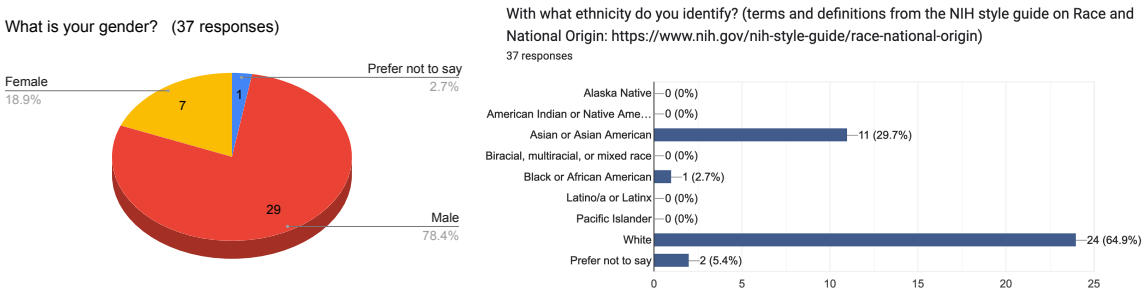


Figure 4: Distribution of gender and ethnicity of participants across both workshops.

3 Program

3.1 User Days Agenda

Color Legend				
Talks	Breaks	Hacking	Tutorials with hacking excercises (pick sessions that apply)	Group discussions

User Days

Day 0 (Sunday, July 23)			
6:00 PM	7:00 PM	1:00:00	Dinner (optional, Dining Room)

Day 1 (Monday, July 24)					
			Axon-Dendrite		Synapse
Start Time	End Time	Duration	Topic	Speaker	Topic
8:00 AM	9:00 AM	1:00:00	Breakfast (Dining Room)		
9:00 AM	9:05 AM	0:05:00	Welcome to Janelia	Janelia Team	
9:05 AM	9:20 AM	0:15:00	Welcome and overview of the hackathon	Ben Dichter	
9:20 AM	10:00 AM	0:40:00	Overview of NWB	Oliver Ruebel	
10:00 AM	11:00 AM	1:00:00	NeuroConv and NWB GUIDE	Cody Baker	Hacking on projects
11:00 AM	12:00 PM	1:00:00	Hacking on projects		
12:00 PM	1:00 PM	1:00:00	Lunch (Dining Room)		
1:00 PM	1:45 PM	0:45:00	Introduction to NWB APIs and Behavior	Ben Dichter, Ryan Ly	Hacking on projects
1:45 PM	2:30 PM	0:45:00	Introduction to Extracellular Electrophysiology		
2:30 PM	3:15 PM	0:45:00	Introduction to Optical Physiology		
3:15 PM	3:30 PM	0:15:00	Refreshments break		
3:30 PM	4:15 PM	0:45:00	Introduction to Intracellular Electrophysiology	Oliver Ruebel	Hacking on projects
4:15 PM	5:30 PM	1:15:00	Hacking on projects		
5:30 PM	6:30 PM	1:00:00	Reception / Ice Breaker (Reception)		
6:30 PM	8:00 PM	1:30:00	Dinner (Dining Room)		

Day 2 (Tuesday, July 25)					
			Axon-Dendrite		Synapse
Start Time	End Time	Duration	Topic	Speaker	Topic
8:00 AM	9:00 AM	1:00:00	Breakfast (Dining Room)		
9:00 AM	9:10 AM	0:10:00	Welcome to Day 2	Oliver Ruebel	
9:10 AM	9:40 AM	0:30:00	Processing and analysis of NWB data	Ben Dichter	Hacking on projects
9:40 AM	10:30 AM	0:50:00	Advanced data I/O (chunking, compression)	Oliver Ruebel, Ben Dichter	
10:30 AM	11:30 AM	1:00:00	TAB (Topic TBD)	Yaroslav O. Halchenko	
11:30 AM	12:00 PM	0:30:00	Hacking on projects		
12:00 PM	1:00 PM	1:00:00	Lunch (Dining Room)		
1:00 PM	2:00 PM	1:00:00	Using the DANDI Archive	Cody Baker	Hacking on projects
2:00 PM	2:25 PM	0:25:00	DataJoint	Kushal Bakshi	
2:25 PM	2:50 PM	0:25:00	pynapple	Guillaume Viejo	
2:50 PM	3:00 PM	0:10:00	Refreshments break		
3:00 PM	3:25 PM	0:25:00	SLEAP	Talmo Pereira	
3:25 PM	6:30 PM	3:05:00	Hacking on projects		
6:30 PM	8:00 PM	1:30:00	Dinner (Dining Room)		

Day 3 (Wednesday, July 26)					
			Axon-Dendrite		Synapse
Start Time	End Time	Duration	Topic	Speaker	Topic
8:00 AM	9:00 AM	1:00:00	Breakfast (Dining Room)		
9:00 AM	9:05 AM	0:05:00	Welcome to Day 3	Ben Dichter	
9:05 AM	9:50 AM	0:45:00	NWB Extensions	Ryan Ly	Hacking on projects
9:50 AM	10:35 AM	0:45:00	NWB Widgets	Ben Dichter	
10:35 AM	12:00 PM	1:25:00	Hacking on projects		
12:00 PM	1:00 PM	1:00:00	Lunch (Dining Room)		
1:00 PM	2:00 PM	1:00:00	Hacking on projects		Hacking on projects
2:00 PM	4:30 PM	2:30:00	Final project presentations		
4:30 PM	5:00 PM	0:30:00	Break		
5:00 PM	5:45 PM	0:45:00	NWB Users and Developer gather in the Axon-Dendrite room for closing discussions of the User Days	Ryan Ly	
5:45 PM	6:00 PM	0:15:00	Group photo (Lobby)		
6:00 PM	7:00 PM	1:00:00	Social reception (Lobby)		
7:00 PM	8:00 PM	1:00:00	Dinner (Dining Room)		

7/20/2023 13:46:39

3.2 Developer Days Agenda

Color Legend				
Talks	Breaks	Hacking	Tutorials with hacking excercises (pick sessions that apply)	Group discussions

Developer Days

Day 4 (Thursday, July 27)					
			Axon-Dendrite		Synapse
Start Time	End Time	Duration	Topic	Speaker	Topic
See the social at the end of day 3			Developers should plan to arrive in the afternoon on day 3 to be able to attend the social		
8:00 AM	9:00 AM	1:00:00	Breakfast (Dining Room)		
9:00 AM	9:30 AM	0:30:00	Welcome to the Developer Hackathon	Oliver Ruebel	
9:30 AM	10:15 AM	0:45:00	Project overviews. Each project provides 2 min overview of their project plan. Planning for new projects.		
10:15 AM	10:30 AM	0:15:00	Break		
10:30 AM	12:00 PM	1:30:00	Break out Discussion: Visualization	Cody Baker	Hacking on projects
12:00 PM	1:00 PM	1:00:00	Lunch (Dining Room)		
1:00 PM	1:15 PM	0:15:00	Developer Group Photo		
1:15 PM	2:15 PM	1:00:00	Break out Discussion: External resources & LinkML	Matthew Avaylon, Sujay Patil	Hacking on projects
2:15 PM	3:15 PM	1:00:00	Hacking on projects		
3:15 PM	3:35 PM	0:20:00	Break		
3:35 PM	6:30 PM	2:55:00	Hacking on projects		Hacking on projects
6:30 PM	8:00 PM	1:30:00	Dinner (Dining Room)		

Day 5 (Friday, July 28)					
			Axon-Dendrite		Synapse
Start Time	End Time	Duration	Topic	Speaker	Topic
8:00 AM	9:00 AM	1:00:00	Breakfast (Dining Room)		
9:00 AM	9:05 AM	0:05:00	Welcome to Day 5	Oliver Ruebel	
9:05 AM	9:25 AM	0:20:00	Morning check in. Review outcomes for day 4 and plans for day 5.	Oliver Ruebel	
9:25 AM	10:25 AM	1:00:00	Hacking on projects		Hacking on projects
10:25 AM	10:40 AM	0:15:00	Break		
10:40 AM	12:00 PM	1:20:00	TAB Discussion (TBD)	TAB	Hacking on projects
12:00 PM	1:00 PM	1:00:00	Lunch (Dining Room)		
1:00 PM	2:00 PM	1:00:00	Future of NWB and Community Building	Ryan Ly	
2:00 PM	3:00 PM	1:00:00	Hacking on projects		Video shoot
3:00 PM	3:20 PM	0:20:00	Break		
3:20 PM	6:30 PM	3:10:00	Hacking on projects		
6:30 PM	8:00 PM	1:30:00	Dinner (Dining Room)		

Day 6 (Saturday, July 29)					
			Axon-Dendrite		Synapse
Start Time	End Time	Duration	Topic	Speaker	Topic
8:00 AM	9:00 AM	1:00:00	Breakfast (Dining Room)		
9:00 AM	9:30 AM	0:30:00	Hacking on projects		Hacking on projects
9:30 AM	12:00 PM	2:30:00	Project presentations	Everyone	
12:00 PM	1:00 PM	1:00:00	Lunch (Dining Room)		
1:00 PM	3:00 PM	2:00:00	Group discussion, feedback, and roadmapping	Everyone	
3:00 PM	5:00 PM	2:00:00	Bonus hacking	Everyone	Hacking on projects

7/20/2023 13:46:39

4 Projects

Participants were encouraged to apply the skills they were learning to implement their own projects and to document their progress in the shared GoogleDocs. Participants reported on 32 unique projects (17 User Days projects and 15 Developer Days projects) and 5+ Breakout sessions. Some projects from the User Days (e.g., 4, 5, 12, and 14 listed below) also continued during the Developer Days. Below a brief summary of the projects reported by participants. For more details about the individual projects, please see the User Days Project GoogleDoc and Developer Days Project GoogleDoc, respectively.

4.1 User Days Projects

1. **NWB conversion of intrinsic physiological recordings across all layers of the rat barrel cortex:** I created a comprehensive dataset of morphoelectric properties across all layers of the rat barrel cortex. My plan is to release the morphology and intrinsic physiology of 300+ neurons. Thus, I'd like to convert the intracellular ephys to NWB. After adopting the NWB standard I'd like to publish the ephys data in the DANDI Archive to make it accessible to the community. For a single example the conversion was successful. The next steps are to replace placeholders in the metadata with actual information, and extend this approach to the other 300+ neurons. *Key Investigators: Felipe Yáñez (Max Planck Institute for Neurobiology of Behavior, Germany)*
2. **Creating spike sorting reference data using high-density extracellular probes:** We have collected data using Neuropixels 'Ultra' probes, which have much higher site density than conventional NP probes, over several brain regions. The high-density data has been sorted and then manually curated. The goal is to make this data useful for testing spike sorters. *Key Investigators: Jennifer Colonell (HHMI Janelia), Susu Chen, Tim Harris*
3. **Meta analysis of AP and EAP features for use as a NWB data format pipeline:** The scope of this project is to collect and analyze (pre- and post-process) neuropixels recordings in different mammalian species using the state-of-the-art tools (Kilosort and PHY) for spike sorting and then make use of in-house Python code algorithms for further spike features analysis. The findings are ought to be linked to current research results which show that based on the features of the intracellular and extracellular spike features one can classify the data into different clusters using machine learning algorithms (KMeans, SVM, Random Forest Classifiers). *Key Investigators: Vitalie Cervinschi (Cedars-Sinai Medical Center, CA, USA)*
4. **MEF3 to NWB data converter:** The goal is to convert electrophysiology data recorded in the MEF3 format to NWB. MEF3 is the Multiscale Electrophysiology Format 3. We have a lot of ECoG and SEEG data, and some microelectrode data from ¿500 human subjects. Being compatible with NWB will allow us to use new tools and to share data. *Key Investigators: Tom Richner (Mayo Clinic, MN, USA)*
5. **NWB + pynapple:** The project aims at redesigning the interface between NWB and the pynapple I/O. So far, pynapple loaders work by taking the output of specific preprocessing pipelines (i.e. Phy, Neurosuite, Minian, etc) and convert it into a NWB object that is then loaded into pynapple objects (i.e. time series and epochs). This approach requires constantly expanding to new pipelines and the package NeuroConv is better suited for this task. The idea is to develop a NWB interface class within pynapple I/O that facilitates reading and writing for users. *Key Investigators: Guillaume Viejo (Simons Foundation, NY, USA)*
6. **NWB Converter - Nemours DECODE lab:** The DECODE lab at Nemours focuses their research on rodent models of epilepsy and how that impacts prefrontal neuronal activity. I will be redesigning their workflow to implement NWB conversion procedures after data collection (behavior/ephys/calcium imaging), but before data analysis. If time permits, I will start implementation of caImAn for calcium imaging analysis or pynapple. *Key Investigators: John Stout (University of Delaware, DE, USA)*
7. **NWB format for high-density EMG arrays:** O'Connor Lab has collected motor unit activity from high-density, flexible EMG arrays made by Sober Lab. I will explore how to integrate EMG recordings and high-speed video tracking into publishable NWB format. The goal is to establish a widely compatible NWB format to generate and publish motor unit activity. *Key Investigators: JJ Kim (Johns Hopkins University, MD, USA)*
8. **Convert 2p imaging and corresponding behavior data to NWB format:** The Sheffield Lab has collected two-photon imaging data using scanbox in .sbx format, analysis is done using suite2p and postprocessing is done on .mat or .npy files. Behavior is collected alongside imaging which is generated as .mat files. The objective of the project is to generate a pipeline that can convert sbx data, suite2p and behavior data to NWB format from each experiment session. *Key Investigators: Seetha Krishnan (University of Chicago, IL, USA)*

9. **Convert multi-modality (e-phys, imaging, behavior) data from the Paz lab to NWB format:** The Paz lab records *in vivo / in vitro* electrophysiology, calcium imaging, fiber photometry, and various behaviors in mice. We need a way to automatically convert our data into NWB format. Ideally, if possible, any data conversion will also directly link, reference, or group data across formats but for the same animal. File types to convert include .mat, .csv, .s2rx, .tif, .sev, .mp4 . *Key Investigators: Jeremy Ford (Gladstone Institutes/UCSF, CA, USA)*
10. **Converting Inscopix 1P outputs to NWB :** The Fleischmann lab has some preliminary multi-session 1P recordings using Inscopix, with freely-moving animals. The goal is to convert the outputs of Inscopix and combine with data from DLC & microcontroller outputs into NWB format. The data potentially will have multi-session registration data (i.e. cell IDs being synced across days), which would also need a format to go in. *Key Investigators: Tuan Pham (Brown University, RI, USA)*
11. **Preliminary analysis of the diversity of NWB data on DANDI Archive:** This project aims to inspect the different aspects of diversity of the NWB (meta)data on DANDI Archive. For example, this could include their originating institutions, study interests, study subjects and techniques (e.g. ophys, electrophys, behaviors). This can be a preliminary effort to track the progress and diversity of neuroscience data, especially following the start of the "Year of Open Science" initiative in the US. *Key Investigators: Tuan Pham (Brown University, RI, USA)*
12. **Conversion of Depth MoSeq-Extract Data from Datta Lab to NWB:** I am helping the Datta Lab at Harvard convert their automatic motion sequencing (MoSeq) data to NWB. I will be adding this data to an ongoing conversion. The objectives of this project are to: i) fully convert one session of MoSeq-Extract output to NWB, ii) integrate this data into existing conversion architecture, and iii) clarify any questions about data structure with Datta Lab (Caleb Weinreb & Kai Fox) *Key Investigators: Paul Adkisson (CatalystNeuro, CA, USA)*
13. **Convert experimental 2P Ca-Imaging data collected from Haptics Lab into NWB:** The goal of this project is to convert the neurophys and behavioral data collected in my thesis work from a custom MATLAB framework (.im2p) into NWB. Currently the data consist of large (30GB per session) .tiff stacks, session metadata, behavioral responses, and subject video monitoring. Ideally all of these can be integrated into 1 file per session to facilitate my analysis pipeline. *Key Investigators: Michael Duhain (University of Rochester, NY, USA) (Affiliation)*
14. **Convert Intracellular EPhys Data to NWB:** My lab has several hundred gigabytes of patch-clamp data that we want to share with the scientific community. I would like to convert our recordings to NWB's format, from the current Axon Instruments ABF format, so that it can be uploaded to the DANDI archive. The objectives of this project are to: i) standardize a method to convert our various types of patch-clamp recordings—cell characterization, CRACM, pharmacology—to NWB and to ii) make the converted files loadable by our lab's intracellular ephys pipeline. *Key Investigators: Isaac Brooks (University of Michigan, MI, USA)*
15. **Convert multi-modality (e-phys, imaging, behavior) data from the Wang lab to NWB format :** Create conversion pipelines for various projects in the lab: i) Ec-ephys with silicone probes recorded on Blackrock acquisition system + whisker tracking + laser stimulation, ii) Ec-ephys with NPX probes recorded with Open Ephys + whisker tracking + DLC tracking + laser / electrical stimulations iii) 2P data recorded with Brucker system, processed with CaImAn + DLC tracking + additional analog signals. *Key Investigators: Vincent Prevosto and Fan Wang (MIT McGovern Institute for Brain Research, MA, USA)*
16. **Advancing Neuroscience Education: Implementing NWB in Undergraduate and Graduate Curriculum:** This educational initiative is aimed at integrating the use of NWB into neuroscience education. The project's focus is on equipping students with practical skills in data management, analysis, and open science practices while highlighting the importance of data standardization. Students will get a hands-on experience with NWB, from reproducing significant neuroscience findings to creating and conducting their own research projects. The ultimate goal is not only to enhance student learning but also to advocate for widespread adoption of NWB as a standard tool in neuroscience education, thereby fostering a culture of standardization and reproducibility in neuroscience research. *Key Investigators: Tommy Lee (Emory University, GA, USA)*
17. **Convert Behavior, Widefield Calcium Imaging, Ephys, and Pupil video data to NWB:** *Key Investigators: Bilal Haider*

4.2 Developer Days Projects

18. **Systems Neuro Browser (SNUB):** SNUB is a tool for exploring time-series data, such as neural recordings, behavioral videos, temperature, movement or other sensor signals, and higher-level annotations

derived from such data. The interface is divided into synchronized windows that each show a different data stream. The linked data views allow users to quickly inspect the relationships between experimental phenomena, such as the behaviors that occur during a particular pattern of neural activity. The objectives of this project are to build a set of functions that can ingest NWB-formatted neural/behavioral datasets and render them within SNUB and to preprocess a collection of datasets on DANDI to facilitate data exploration. *Key Investigators: Caleb Weinreb (Harvard Medical School), Mohammed Osman, Vitalie Cervinschi (Cedars-Sinai Medical Center)*

19. **Keypoint-MoSeq:** Motion Sequencing (MoSeq) is an unsupervised machine learning method for animal behavior analysis. Given behavioral recordings, MoSeq learns a set of stereotyped movement patterns and when they occur over time. The objectives of this project are to: build a loader for keypoint-MoSeq that ingests keypoint tracking data stored in NWB files; build an export function that saves the output of keypoint MoSeq in an NWB compatible format; and create a set of visualization tools and/or a pipeline/notebook/recipe for juxtaposing MoSeq output with other data modalities in NWB such as neural recordings (e.g. behavior-triggered averages, etc.,). *Key Investigators: Caleb Weinreb (Harvard Medical School), Kai Fox (Harvard), Mohammed Osman*
20. **Parallel Write Support for HDMF-Zarr:** Allow NWB files written using the Zarr backend for NWB available in hdmf-zarr to leverage multiple threads or CPUs on write to enhance speed of operation. *Key Investigators: Cody Baker (CatalystNeuro), Oliver Rübel (LBNL), Alessio Buccino (Allen Institute for Neural Dynamics), Vitalie Cervinschi (Cedars-Sinai Medical Center)*
21. **Neurodata Extension for Probeinterface:** Currently there is no established way to define information about neural probes in the core NWB schema. I plan to create an extension to address this issue. Specifically, the extension would be consistent with the way probes are defined in probeinterface, a mature package that is used to define neural probes as python objects. Using the probeinterface standard has advantages such as its interoperability with spikeinterface, a python package widely used for processing of electrophysiology data including spike sorting (e.g. you can attach a probeinterface Probe object to a spikeinterface Recording for spike sorting which requires knowing channel locations). There is already an NDX for this ndx-probe-interface but it is missing quite a number of features, which needs to be worked on. *Key Investigators: Kyu Hyun Lee (UCSF), Alessio Buccino (Allen Institute for Neural Dynamics), Vitalie Cervinschi (Cedars-Sinai Medical Center)*
22. **DataJoint to NWB for Calcium Imaging Data:** DataJoint develops and maintains standardized open source pipelines for neuroscience research. Currently, only the DataJoint pipeline for NeuroPixels electrophysiology supports converting and packaging data to NWB format. This project will develop a routine to convert calcium imaging data and relevant metadata to NWB format. *Key Investigators: Kushal Bakshi (DataJoint), Dimitri Yatsenko (DataJoint)*
23. **Educational notebook based on a Nobel Prize-Attributed dataset:** Create a lesson plan for undergraduate-level students using NWB and DANDI that could be dropped into computational neuroscience courses. The problem set will guide students through the process of accessing data and recreating key figures from the work of Sargolini et al. 2006, a Moser lab paper attributed in the 2014 Nobel Prize announcement. *Key Investigators: Anna Szonja Weigl (CatalystNeuro), Ben Dichter (CatalystNeuro), Dorota Jarecka (MIT)*
24. **DynamicTermset and ExternalResources:** The ExternalResources class provides a way to organize and map user terms from their data (keys) to multiple entities from the external resources. A typical use case for external resources is to link data stored in datasets or attributes to ontologies. We've integrated LinkML enumerations to create "TermSet", a way for users to create their own set of terms from brain atlases, species taxonomies, and anatomical, cell, and gene function ontologies. This is currently manual, but with the help of the LinkML team, we will integrate a dynamic way to pull this information from online resources/APIs. *Key Investigators: Matthew Avaylon (LBNL), Sujay Patil (LBNL)*
25. **NWB compatibility for Kilosort4:** Write code for reading raw data from NWB files (through SpikeInterface) in a manner compatible with Kilosort4. *Key Investigators: Jacob Pennington (HHMI Janelia)*
26. **NWB Project Analytics:** The NWB project maintains and contributes to a large number of codes related to NWB. The goal of this effort is to help developers to get a quick overview of the state of NWB code repositories by collecting metadata (e.g., from CodeCov and GitHub) and computing code statistics (e.g., using CLOC) in a central location and compiling the results in a consistent fashion. The code to collect and generate the statistics will be implementded in nwb-project-analytics and the pages will be rendered and released via the nwb-overview website. *Key Investigators: Oliver Rübel (LBNL)*
27. **Finalize ndx-events extension and NWB extension proposal:** The ndx-events extension for event and TTL data was recently reviewed by the TAB. Some work needs to be done on it before it can be merged into the core NWB standard. The objectives of this project are to finalize the NWB Extension

Proposal document and to finalize the extension and release it on PyPI. *Key Investigators: Ryan Ly (LBNL)*

28. **Test and finalize ndx-pose extension:** The ndx-pose extension for pose estimation data was recently extended to permit annotated data and model training parameters as applicable for both SLEAP and DeepLabCut. The objectives of this project are to collect and integrate feedback from users and developers who work with SLEAP and DLC data and to finalize the extension. *Key Investigators: Ryan Ly (LBNL)*
29. **Whole-cell EPhys Pipeline – NWB Integration:** In my lab, I am the developer of an in-house pipeline for batch analysis of patch-clamp recordings. We are planning to publish this pipeline with an associated GUI as a Python package later this year, but would like it to be able to interface with NWB data formats, as it currently only runs on Axon Instruments' .ABF files. The main objective of this project is to create a working version of the pipeline that interfaces with PyNWB's intracellular ephys data format. *Key Investigators: Isaac Brooks (University of Michigan)*
30. **Convert multi-modality (e-phys, imaging, behavior) data from the Paz lab to NWB format:** The Paz lab records in vivo/in vitro electrophysiology, calcium imaging, fiber photometry, and various behaviors in mice. We need a way to automatically convert our data into NWB format. Ideally, if possible, any data conversion will also directly link, reference, or group data across formats but for the same animal. File types to convert include .mat, .csv, .s2rx, .tif, .sev, .mp4 .. *Key Investigators: Jeremy Ford (Gladstone Institutes/UCSF)*
31. **Multiphoton Image Compression:** The first step is to reliably estimate the effective quantum flux. This computation will be included in the upcoming Methods paper. We worked through two-photon datasets from DANDi to compute noise-equivalent quantal size. The work will be featured on the compress-multiphoton repository. Having the knowledge of quantal sizes allows an efficient grayscale digitization that leads to efficient data compression. Our benchmarks show 4:1 to 10:1 compression. We plan to work with NWB to introduce a specialized codecs that efficiently and transparently compresses and decompresses multiphoton series. *Key Investigators: Dimitri Yatsenko (DataJoint)*
32. **Prompt exploration to identify relevant dandisets for a particular scientific question.** Given a scientific question, LLMs can also be helpful to search for potentially useful Dandisets. For this we would produce vector embeddings for each Dandiset. Such vectors will hold semantic values extracted from DANDI set abstracts and asset summaries, on top of which semantic search could be performed. In a first, naive approach, we will rank the results based on semantic similarity and return the first options as they come. In a more advanced approach, we plan to run a prompt consisting of the original question plus top results through a LLM asking it to further enquire and explain in which ways the best results would be relevant to the posed question. *Key Investigators: Felipe Yáñez (MPI for Neurobiology of Behavior, Germany), Ben Dichter (CatalystNeuro), Luiz Tauffer (CatalystNeuro).*

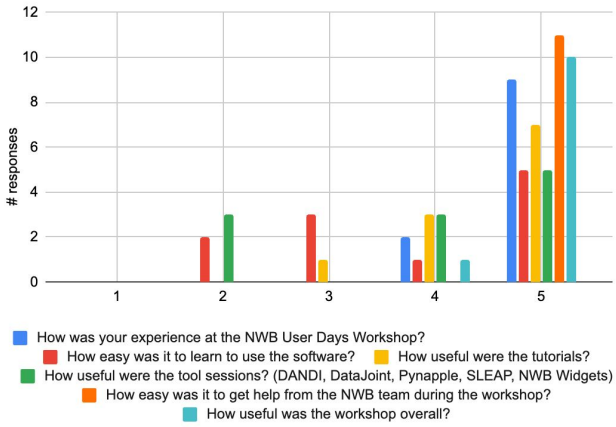
4.3 Breakout sessions

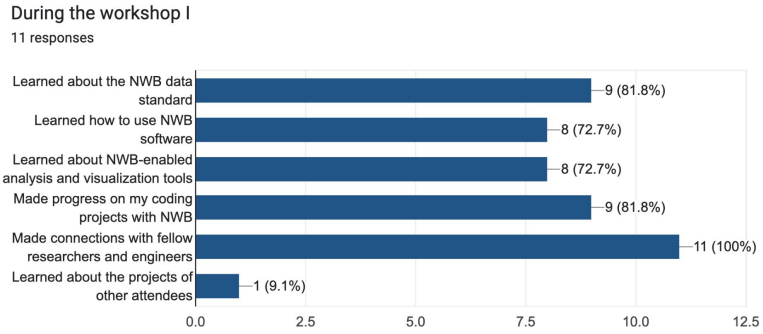
33. **NWB Technical Advisory Board (TAB)** The TAB hosted a breakout sessions during both the User Days and the Developer Days. The breakout was broadly focused on: 1) introduction of the TAB and recent progress by the TAB and 2) interaction of the TAB with the community to discuss important questions with community, e.g.,: *"What are high-level obstacles on adoption of NWB which need technical solutions?"*, *"What documentation, tools, specifications etc are missing in the NWB ecosystem?"*, or *"What other activities or responsibilities would NWB developers would want for TAB to get?"*. Materials: Slides (Google) Notes: GoogleDoc *Session Lead: Yaroslav O. Halchenko (Dartmouth), Session Scribe: Oliver Rübel (LBNL) TAB Participants Saskia de Vries (Allen Institute), Anna (Szonja) Weigl (CatalystNeuro), Lawrence Niu (MBF Bioscience), Alessio Buccino (Allen Institute), Yaroslav O. Halchenko (Dartmouth), Ryan Ly (LBNL), Oliver Rübel (LBNL)*
34. **Visualization Breakout:** Visualization is important to the whole data science pyramid, from collection, exploration, aggregation to learning. This breakout focused on discussions of challenges, needs, opportunities, and approaches for visualization in neurophysiology. Materials: slides (PDF) Notes: GoogleDoc *Session Lead: Cody Baker (CatalystNeuro), Session Scribe: Oliver Rübel (LBNL)*
35. **External Resources** & **LinkML:** This breakout session introduces the neuroscience external resources data standard (implemented in HDMF) and the use of term sets (implemented via LinkML) with NWB data. Materials: Slides (PowerPoint) *Session Lead: Matthew Avaylon (LBNL) and Sujay Patil (LBNL)*
36. **Future of NWB and Community Building:** This session focused on discussion with attendess on the future of NWB and growing and building the NWB community. Notes: GoogleDoc *Session Lead: Ryan Ly (LBNL), Session Scribe: Matthew Avaylon (LBNL)*

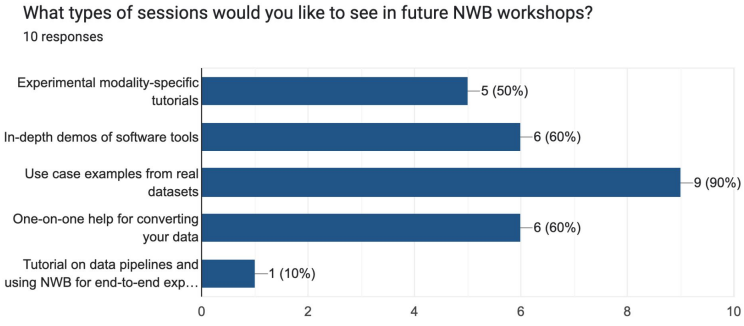
37. **Project Discussions and Presentations:** Attendees introduced their projects at the beginning of the workshop and presented and discussed progress on their projects on the last day of the User Days and Developer Days, respectively.

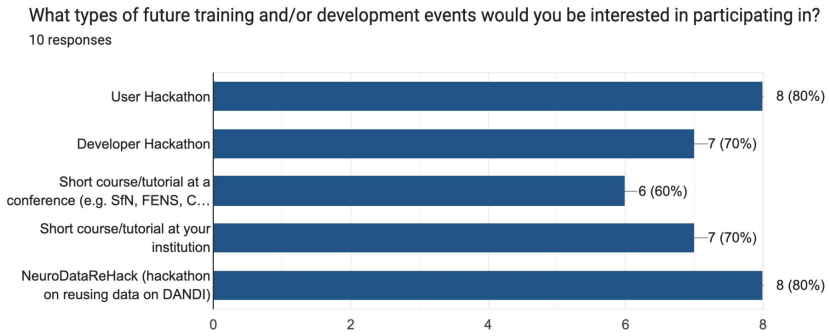
5 Exit Survey

5.1 User Days Exit Survey



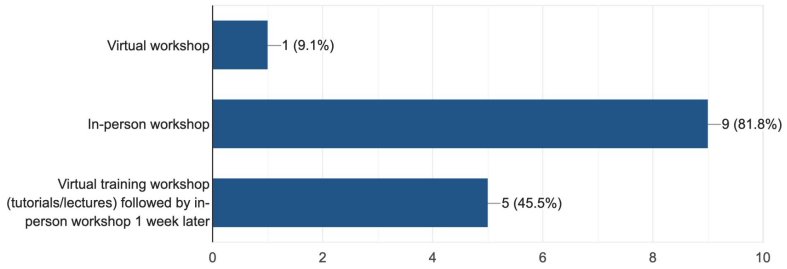






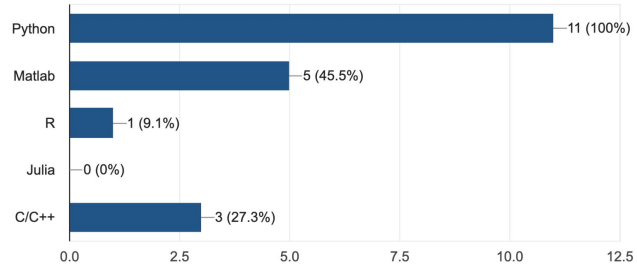
What format of event would you prefer if you were to attend another NWB event?

11 responses



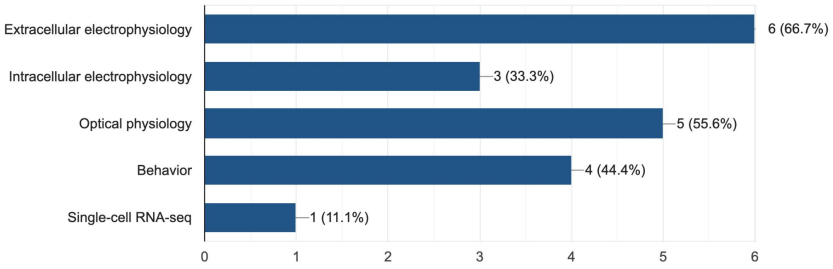
Which programming languages are you using to interact with your data?

11 responses



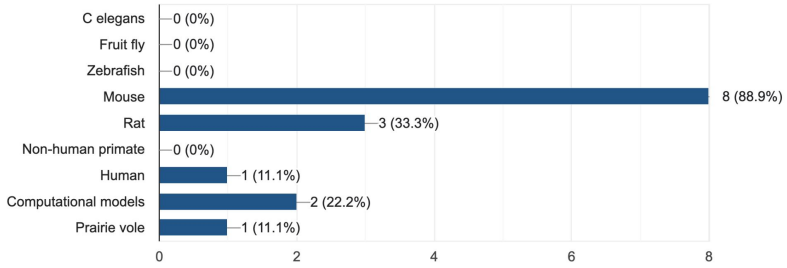
Which data modalities are you using

9 responses



Which subjects are you using in your research?

9 responses



What were the most helpful parts of the workshop? (7 + 2[1] responses): **1)** The hacking sessions. **2)** Getting pointers to specific tutorials for specific tools – I find it hard to absorb all the details in real time, so for me the most useful thing is knowing where to learn more later. **3)** One on one discussion and hacking time. **4)** Getting direct contact with other users and developers. **5)** Learning how to use NWB. **6)** Interacting with NWB developers. **7)** I focused on my coding projects with NWB, getting help from the team during the workshop was very useful. **8)** Time to work with the NWB team on projects. Free evenings to network. **9)** Direct access to core devs

What could have been improved at the workshop? (7 responses): **1)** More hands on hacking sessions. For novice users, it would have been nice to have an extra day where we spent converting some basic data to nwb. It was hard for me to pay attention to/understand lectures on all the extra resources available like visualization tools, extensions and widgets when I was still struggling with understanding the nwb format and data structure. **2)** Overall I think it went really well. **3)** None, please host more of these events! **4)** I would cut back on lecture. Multiple well-flushed and multifaceted tutorials would be better. The talks are important, but maybe after every talk, an hour hacking or something. **5)** more diverse pool of participants, though this is truly not easy; I like the idea of "NWB tour", which might address this **6)** I would suggest more hacking time, maybe at least 3 or 4 hours a day. And/or having a day 0 dedicated for first time users. **7)** Longer blocks of work time earlier in the conference.

How can we best reach other researchers and engineers in your lab or at your institution? (5 responses) **1)** Unsure... **2)** Workshops like this one are a good tool. **3)** Keep doing what you're doing. I think educational outreach is going to be an important piece **4)** NWB tour; have an ambassador? **5)** I will help with word of mouth around my institution, continue to produce quality documentation on the website.

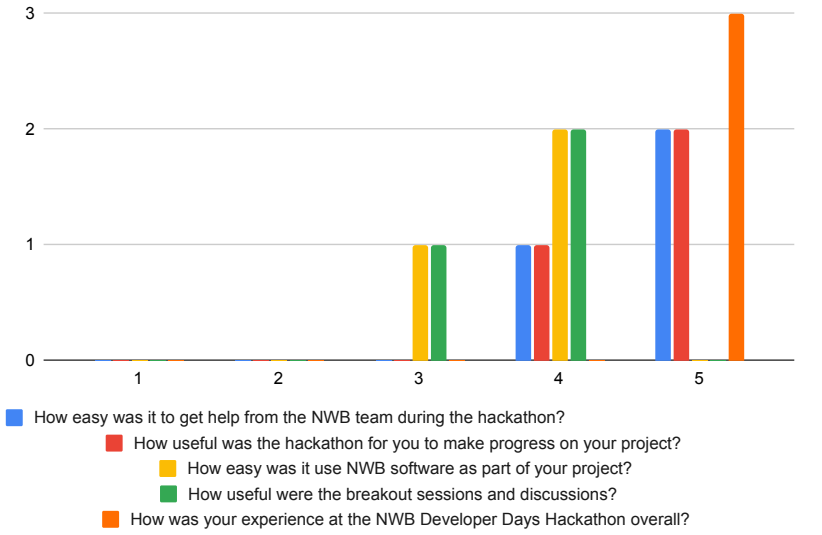
What tools would make it easier for you to adopt NWB? (3 + 1[1] responses): **1)** More .ipnyb notebook based tutorials using real datasets. **2)** neurosift, more analysis based like pynapple. **3)** NWB Guide. **4)** GUI based approaches for converting data. Additionally, a labview-like module that lets you build the data objects and connect them. Each block has fields to the required and optional entries. This could help people who have no idea how to code but have more complex data.

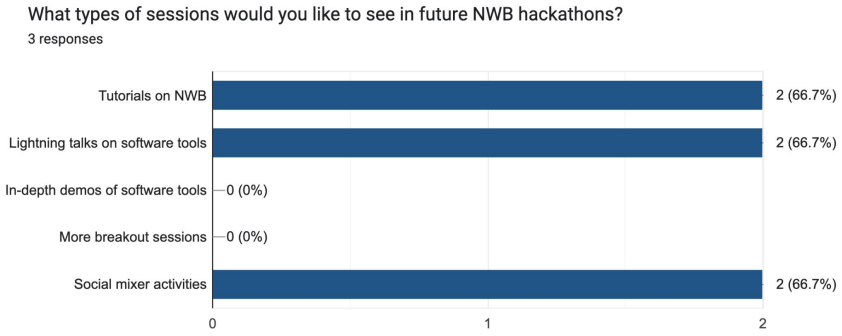
What existing software (data processing/analysis/management) and data acquisition systems would you like to see integrated with NWB? (3 responses): **1)** More complete support for neuralynx acquired data. **2)** inscopix. **3)** Molecular Devices' Axon pCLAMP Suite.

[1] "*+ X*" refers to additional responses to the corresponding question in the Developer Days Survey.

5.2 Developer Days Exit Survey

Only 3 attendees responded to the Developer Days survey, which is likely due to attendees who participated in both events only responding to the first (User Days) survey, but not the second (Developer Days) survey. All 3 respondents to the developer survey indicated that the length of the hackathon was "just right" and indicated potential interest in attending a 2/3day Brainhack-like hackathon for open science prior to COSYNE.





Testimonials

"The User Days event was incredibly helpful. The CatalystNeuro team was constantly providing support to attendees to close knowledge gaps and teach NWB formatting. Implementing NWB opens doors for analysis, lowers the energy barrier for data sharing, and if done correctly, supports experimental organization. This workshop is a must for anyone in the neuroscience community looking to move towards open science!" – John Stout, Nemours Children's Hospital.

"The Hackathon was a great place to learn about NWB and how to implement it for my data. Having direct access to the developers accelerated my ability to create a data conversion solution specific to my datasets. It was great to meet other scientists in the field that we're learning NWB alongside me." – Jeremy Ford, Gladstone Institutes, UCSF

"Participating in the NWB workshop was an enriching experience, providing invaluable insights into NWB's transformative potential in neuroscience education. Throughout the workshop, I was able to connect with like-minded researchers passionate about open science. These new connections are not just professional relationships, but future collaborations that will further the project's reach and impact. My project, "Advancing Neuroscience Education: Implementing NWB in Undergraduate and Graduate Curriculum," is designed to systematically incorporate NWB into our academic programs, thereby fostering vital skills needed in the ever-evolving field of neuroscience. Emphasizing diversity and inclusion, the project seeks to implement this innovative curriculum at a spectrum of institutional types, from minority-serving to high-research-activity to primarily undergraduate institutions. The initiative underscores data standardization and reproducibility's importance, prioritizing student-led research to foster hands-on learning and discovery. One notable challenge is distilling NWB into digestible, easy-to-follow guides featuring classical, high-impact experiments that can excite undergraduates in a computational neuroscience course. I'm confident, however, that through continual engagement with the NWB community, regular evaluations, and methodological adjustments, we can surmount this obstacle. Through broad-reaching initiatives and collaborations with other educators, with a specific focus on fostering inclusivity, I aim to extend our project's impact. I encourage anyone involved in neuroscience education to explore the vast resources and opportunities offered by the NWB community, which can fundamentally redefine how we teach and learn neuroscience. I highly recommend you to join the next NWB workshop and become an active part of this innovative community, shaping the future of neuroscience education." – Tommy Lee, NIH-IRACDA FIRST Postdoctoral Fellow, Emory University.

"Great job to everyone setting up and running a hackathon. I think it was a great success." Anonymous.

6 Photos from the Event

6.1 Photos from the User Days

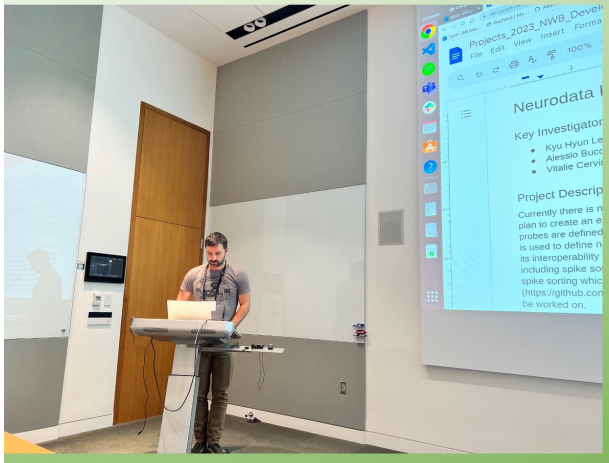
July 25, 2023: Day 2, User Days

July 26, 2023: Day 3, User Days

Final Project Presentations and Closing Discussions

6.2 Photos from the Developer Days

July 28 - 29, 2023:
Day 2-3, Developer Days



Ping Pong Social

Acknowledgements

We would like to thank the presenters and hosts of the breakout sessions for their contributions to the workshop (see Sec. 3). We would like to thank all participants for the great enthusiasm and for making the event a great success! We would like to thank the members of the NWB Technical Advisory Board (TAB) for their contributions to the events. Organization and operations of the event and NWB research activities were supported by the National Institute of Neurological Disorders and Stroke of the National Institutes of Health under Award Number U24NS120057 to O. Rübel (LBNL) and B. Dichter (CatalystNeuro). Housing, rooms, catering, and on-site organization was supported by HHMI Janelia. Additional travel support for participants was provided by HHMI Janelia. Community projects on NWB are also supported by the Kavli Foundation and Simons Foundation (see also https://www.nwb.org/projects/).

Disclaimer