



# 2024 NWB Developer Days

## hosted by DataJoint!

April 17 – 19, 2024

Ryan Ly<sup>1</sup>, Kushal Bakshi<sup>2</sup>, Benjamin Dichter<sup>3</sup>, Oliver Rübel<sup>1</sup>

<sup>1</sup>Scientific Data Division, Lawrence Berkeley National Laboratory

<sup>2</sup>DataJoint

<sup>3</sup>CatalystNeuro



## Contents

1 Executive Summary	3
2 Participants	4
3 Program	6
4 Projects	8
5 Exit Survey	10
6 Photos from the Event	13
Acknowledgements	16
Disclaimer	16

2

# 1 Executive Summary

**Overview:** This event was focused on bringing together the developers of the NWB data standard and software ecosystem and developers of community tools who would like to integrate their tools with NWB. This hackathon enabled participants to work intensively on an NWB-related project with the assistance of the NWB development team and others in the community. Together, we developed and discussed ideas for solving technical problems that impact the broader community. Our goal was to foster collaboration and community among developers.

This was our first NWB hackathon event hosted by DataJoint. Working in the DataJoint building and with their team facilitated projects at the intersection between DataJoint and NWB, such as Spyglass and building NWB export for custom DataJoint pipelines. We also welcomed NWB projects that were not related to DataJoint. Many participants also attended the [DataJoint SciOps Summit](#) immediately preceding the NWB hackathon to learn about DataJoint and interact with DataJoint users.

During the event, the 19 attendees from 9 different institutions (see Sec. 2) worked on 16+ different projects ranging from improving the NWB software and website, to adding support for new data types to NWB and expanding on support for existing data types, to integrating NWB with community tools (see Sec. 4).

**Participants:** 26 people from 15 different institutions registered for the hackathon. 19 people from 9 different institutions attended the hackathon in-person. Most participants identified as staff researchers, research scientists, or engineers and came from a mix of research institutions and companies. For further details about participants, see Sec. 2.

**Program:** The program consisted of a combination of open sessions for joint coding on hacking projects and breakout sessions for coordination of efforts and discussions of development efforts with the community (see Sec. 4). The breakout sessions were optional but most participants attended and engaged in discussion. Opportunities for social interaction and team discussions were available throughout the event during the open hacking sessions, meals, a mixer at the beginning of the event, and an optional evening social.

**Conclusion:** The NWB Developer Hackathon was successful in connecting NWB users and developers and advancing the NWB software ecosystem. The hackathon was overall well-received (see Sec. 5). Participants made significant progress on their projects and appreciated getting to know each other and having interesting formal and informal discussions about NWB and the neurophysiology software ecosystem.

## Organizing Committee:

- **Program Chairs:**
  - Ryan Ly (LBNL)
  - Benjamin Dichter (CatalystNeuro)
- **Site Chair:** Kushal Bakshi (DataJoint)

3

## Related Documents:

- [Event Website](#)
- [Project GoogleDoc](#)

## 2 Participants

26 people from 15 different institutions registered for the NWB Developer Hackathon. 19 people from 9 unique institutions attended the hackathon in-person. The table in Figure 2 shows a summary of all in-person participants. 42% of participants attended a NWB workshop for the first time. The breakdown of current positions of participants at the User Days is shown in Figure 3. The registration form also included optional questions for participants to indicate their gender and ethnicity. 18 of 19 participants responded to each question. Figure 4 shows the distribution of gender and ethnicity of participants across both workshops. In terms of programming language, 18 of 19 participants indicated to be most comfortable with Python ( $\approx 95\%$ ) and MATLAB ( $\approx 26\%$ ) (and  $\approx 5\%$  indicated C++ or JavaScript).



Figure 1: Participants of the Developer Hackathon

Country	State	Institution	Count
United States	CA	CatalystNeuro	3
	CA	HHMI/UCSF	1
	CA	Lawrence Berkeley National Lab	5
	CA	University of California, San Francisco	1
	IN	CatalystNeuro	1
	IN	Orthogonal Research and Education Laboratory	1
	MA	MIT	1
	NJ	Princeton Neuroscience Institute/Princeton University	1
	NY	Flatiron Institute	1
	TX	DataJoint	4

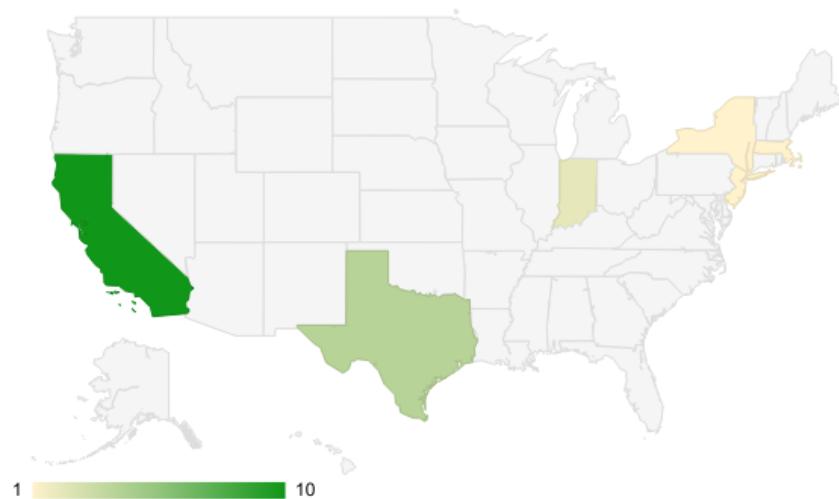


Figure 2: Institutions and locations of participants.

#### Current Position (19 Responses)

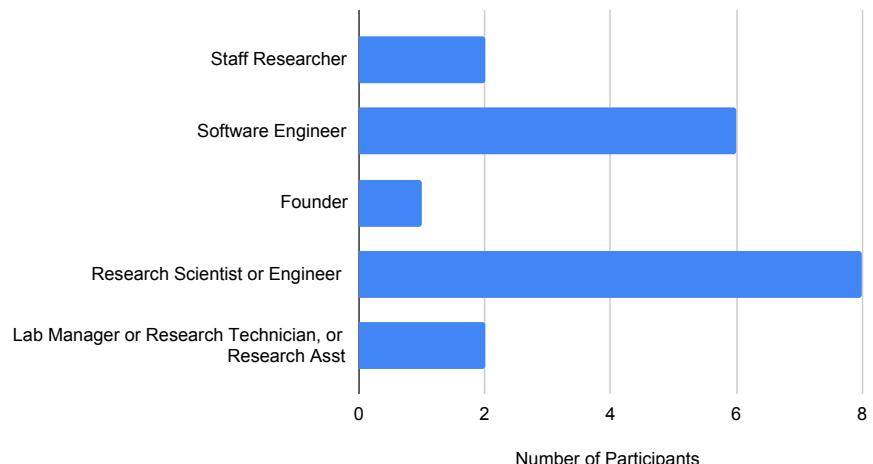


Figure 3: Current positions of participants.

### 3 Program

The agenda consisted of a combination of open sessions for joint coding on hacking projects and breakout sessions for coordination of efforts and discussions of development efforts with the community (see Sec. 4). Scheduled coffee breaks, lunch, and a optional evening social at *Activate*—an active gaming facility where you jump, climb, and solve problems in small groups—were used to encourage networking and social interactions. On the first day, a couple of talks set the stage for the event and provided and overview of recent developments in the NWB ecosystem. Breakout discussions covered topics of active development in the NWB ecosystem. The NWB Technical Advisory Board (TAB) also presented on its activities and solicited feedback on priorities from the developer community.

Color Legend			
Talks	Breaks	Hacking	Group discussions

## 2024 NWB Developer Hackathon

### Day 0 (Tuesday, April 16)

5:30 PM 6:30 PM 01:00 Social Mixer - DataJoint / NWB

### Day 1 (Wednesday, April 17)

Start Time	End Time	Duration	Topic	Speaker
8:50 AM	9:00 AM	00:10	Arrive at DataJoint HQ	
9:00 AM	9:05 AM	00:05	Welcome to DataJoint HQ	DataJoint Team
9:05 AM	9:15 AM	00:10	Welcome and overview of the hackathon	Ryan Ly
9:15 AM	9:45 AM	00:30	Developer updates on NWB	NWB Team
9:45 AM	10:30 AM	00:45	Project overviews. Each team will give a 2-min overview of their project plan	
10:30 AM	12:00 PM	01:30	Hacking on projects	
12:00 PM	1:00 PM	01:00	Go out to lunch	
1:00 PM	1:10 PM	00:10	Group photo	
1:10 PM	2:10 PM	01:00	Breakout Discussion: Scientific Workflows and NWB	Thinh Nguyen and Eric Denovellis
2:10 PM	3:30 PM	01:20	Hacking on projects	
3:30 PM	3:45 PM	00:15	Refreshments break	
3:45 PM	6:00 PM	02:15	Hacking on projects	

### Day 2 (Thursday, April 18)

Start Time	End Time	Duration	Topic	Speaker
8:50 AM	9:00 AM	00:10	Arrive at DataJoint HQ	
9:00 AM	9:05 AM	00:05	Welcome to Day 2	Ben Dichter
9:05 AM	11:00 AM	01:55	Hacking on projects	
11:00 AM	11:30 AM	00:30	NWB TAB and NWB Enhancement Proposals (NEPs)	NWB TAB
11:30 AM	12:00 PM	00:30	Hacking on projects	
12:00 PM	1:00 PM	01:00	Go out to lunch	
1:00 PM	2:00 PM	01:00	Breakout Discussion: Modular, Cloud-Friendly Data Storage and LINDI	Jeremy Magland, Ben Dichter, and Ryan Ly
2:00 PM	3:30 PM	01:30	Hacking on projects	
3:30 PM	3:45 PM	00:15	Refreshments break	
3:45 PM	6:00 PM	02:15	Hacking on projects	

### Day 3 (Friday, April 19)

Start Time	End Time	Duration	Topic	Speaker
8:50 AM	9:00 AM	00:10	Arrive at DataJoint HQ	
9:00 AM	9:05 AM	00:05	Welcome to Day 3	Ben Dichter
9:05 AM	9:45 AM	00:40	Breakout Discussion: LinkML & Neurophysiology Data	Ryan Ly
9:45 AM	12:00 PM	02:15	Hacking on projects	
12:00 PM	1:00 PM	01:00	Go out to lunch	
1:00 PM	2:00 PM	01:00	Hacking on projects	
2:00 PM	4:00 PM	02:00	Final project presentations	Ryan Ly
4:00 PM	4:15 PM	00:15	Refreshments break	
4:15 PM	5:30 PM	01:15	Group discussion, feedback, and roadmapping	Ben Dichter

## 4 Projects

As part of the hackathon, participants worked on a large variety of projects related to NWB. Participants documented their progress in the shared [Project GoogleDoc](#). Participants reported on 16+ unique projects and 4 Breakout sessions. Below a brief summary of the projects reported by participants:

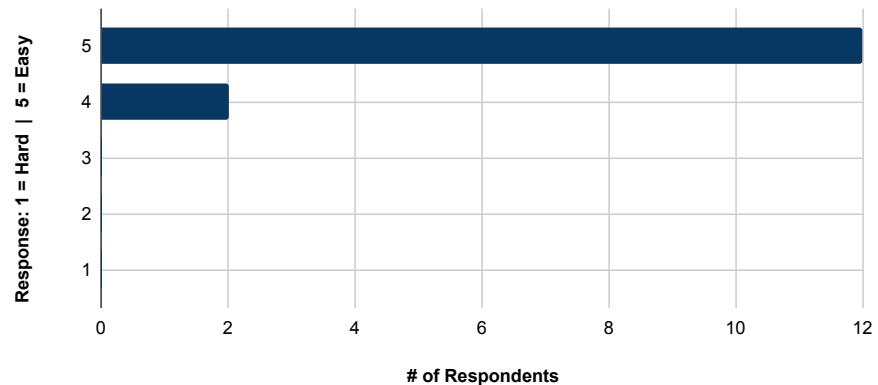
1. **Refine implementation of NEP002: Probes and channels:** Expand current representations of channels and electrodes with an emphasis on compatibility with ProbeInterface metadata. *Key Investigators: Ryan Ly, Matthew Avaylon, Cody Baker, Heberto Mayorquin*
2. **Finish implementation of dimension labels and Xarray support in HDMF:** There is renewed interest in using xarray for working with NWB/neurophysiology data and modeling neurophysiology data with labeled dimensions generally. The goal of the project is to implement support for dimension labels and Xarray in HDMF *Key Investigators: Ryan Ly*
3. **TermSet Schema and Integration with PyNWB:** Integrate term validation by having dedicated termsets for fields within the schema. *Key Investigators: Matthew Avaylon*
4. **NWB GUIDE User Testing:** Improve the NWB GUIDE through a series of rapid user tests. *Key Investigators: Garrett Flynn, Cody Baker, Ryan Ly*
5. **Upgrade pynwb validation methods:** The pynwb validate module validates an NWB file against a namespace. There are currently several validation-related issues in pynwb that should be upgraded for the next major release. *Key Investigators: Stephanie Prince*
6. **LINDI: Towards NWB and DANDI adoption:** LINDI provides a JSON/Zarr representation of NWB data where the large data chunks are stored separately from the main metadata. This enables efficient storage, composition, and sharing of NWB files on cloud systems such as DANDI without duplicating the large data blobs. *Key Investigators: Jeremy Magland, Ryan Ly*
7. **Neurosift: Expand support for additional NWB data types:** Neurosift is a browser-based tool designed for the visualization of NWB files, whether stored locally or hosted remotely. The goal of this project is to identify the most needed features and implement them! *Key Investigators: Jeremy Magland, Cody Baker, Ben Dichter*
8. **Converting NeuroData into NWB as part of a DataJoint Pipeline:** Learning the fundamentals of converting neuroscience data (ephys and calcium imaging) into NWB files using the Neuroconv. This is to help export the data from DataJoint and automated the process overall. *Key Investigators: Christian Tabedzki*
9. **NWB-Cloud Benchmarking:** Accessibility and storage issues in the cloud limits cloud-based analysis and exploration of large neurophysiology datasets stored in the cloud. This work helps evaluate and optimize reading of NWB neurophysiology data from cloud storage to enhance cloud-based analysis. *Key Investigators: Oliver Ruebel, Ryan Ly, Urjoshi Sinha, Cody Baker, Ben Dichter*
10. **AqNWB: C++ API:** Test and continue to develop a first prototype for a C++ acquisition API for NWB *Key Investigators: Stephanie Prince, Oliver Ruebel*
11. **Incorporate C. elegans extension into core schema and create robust and useful conversion tools:** I have developed an extension to the NWB schema that adds objects to support multi-channel volumetric imaging and metadata associated with C. elegans neuroscience. So far, 5 labs have converted their datasets to this format using the extension with multiple additional labs interested in converting in the near future. The goal of this project is to make the extension and data conversion broadly accessible for users. *Key Investigators: Daniel Sprague*
12. **Adding NWB export capabilities to the Movement package:** Tracking the detailed body pose of animals over time is becoming increasingly important to the field of systems neuroscience. *movement* is a new python package from the SWC/GCNU Neuroinformatics Unit that takes the output of popular pose estimation packages and standardizes their outputs for easy analysis. It would be useful if this package could export NWB objects utilizing the ndx-pose extension. *Key Investigators: Eric L. Denovellis*
13. **User interviews to improve the usability of the DANDI Hub:** The DANDI Hub is a NIH-sponsored compute resource for users to explore Dandisets and gain experience with the NWB ecosystem. The objective of this project is to define a few features to improve the discoverability and user experience of the DANDI Hub. *Key Investigators: Kabilar Gunalan*
14. **Add Tutorials to Export from DataJoint to NWB:** DataJoint Elements, DataJoint's open-source pipelines support exporting data to NWB files. DataJoint recently updated the tutorials within Elements to a more streamlined structure and standardized environment using devcontainers and GitHub codespaces. This project is intended to migrate the export to NWB tutorial for DataJoint Element for Extracellular Array Electrophysiology to this new structure. *Key Investigators: Kushal Bakshi*

15. **Improve html representation of containers with data attribute in hdmf:** Improve HTML representation for data attributes in hdmf containers for the different array implementations. *Key Investigators: Heberto Mayorquin*
16. **Create design document for NWB.org website migration:** The goal is to move the NWB.org to a more modern and accessible platform to promote community contribution and enhance maintainability. *Key Investigators: Ben Dichter*
17. **Breakout: Scientific Workflows and NWB:** Discuss how to manage scientific workflows using DataJoint and Spyglass and how to store provenance and workflow information about individual data objects within an NWB file. *Key Investigators: Thinh Nguyen and Eric Denovellis*
18. **Breakout: NWB TAB and NWB Enhancement Proposals (NEPs):** Update the developer community about recent activities of the NWB Technical Advisory Board, including improved processes for NWB Enhancement Proposals. Also solicited feedback from the developer community on what the NWB team should prioritize and how to improve usability, adoption, and outreach. *Key Investigators: NWB TAB*
19. **Breakout: Linked and Modular, Cloud-Friendly Data Storage and LINDI:** Discuss a new JSON-based representation of NWB data called LINDI and how it can be used to link source, raw, and processed data from different repositories, other potential applications, and potential pitfalls and concerns. *Key Investigators: Jeremy Magland, Ben Dichter, and Ryan Ly*
20. **Breakout: LinkML & Neurophysiology Data:** Discuss recent developments on adding support for n-dimensional arrays in LinkML with support for different backends and different in-memory representations, and prospects of translating and using a LinkML-based version of the NWB schema. *Key Investigators: Ryan Ly*

## 5 Exit Survey

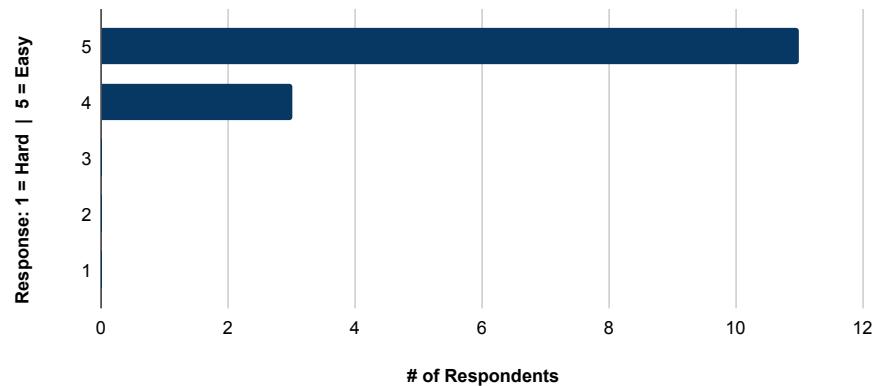
**How easy was it get help from the NWB / DANDI teams during the workshop?**

Responses: 14



**How useful was the hackathon for you to make progress on your project?**

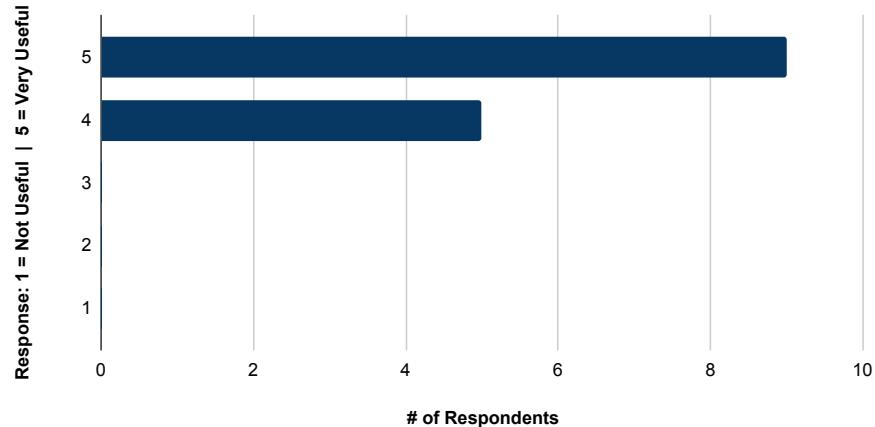
Responses: 14



10

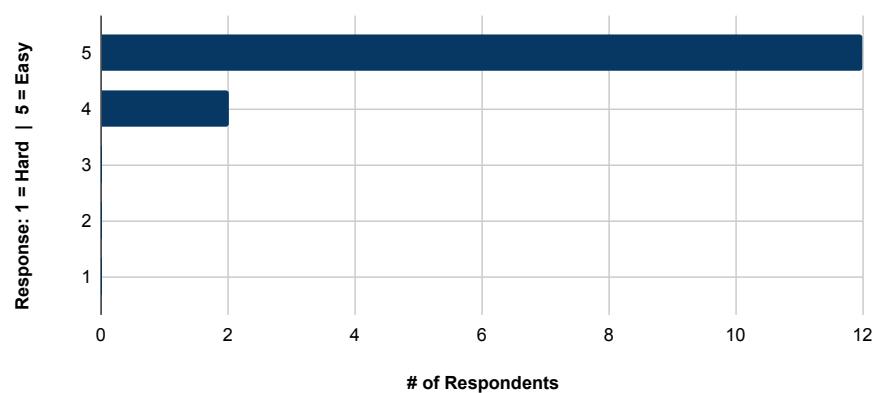
### **How useful were the breakout sessions overall?**

Responses: 14



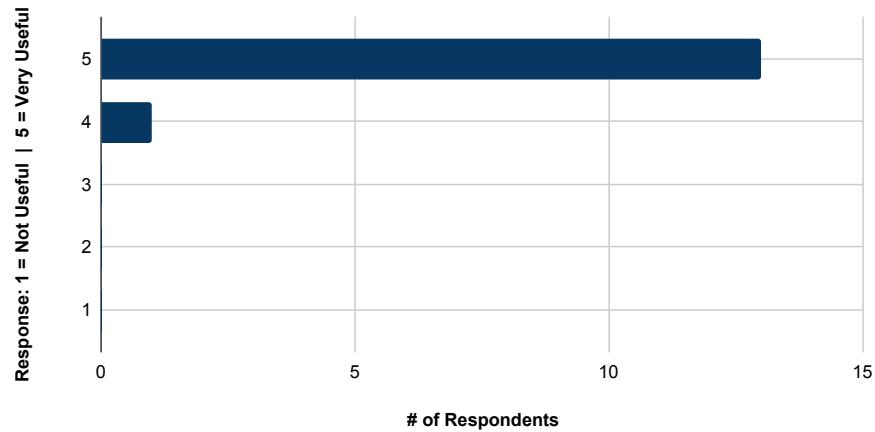
### **How useful was the hackathon for you to meet other developers?**

Responses: 14



### **How useful was the hackathon overall?**

Responses: 14



## Open-Ended Responses

### What did we do well at the hackathon?

1. Encouraging conversation
2. Great setup for connection and rich discussion on NWB-related topics
3. we are cool people
4. Inter communication
5. Giving unstructured time to work on specific projects. Supporting discussions and coworking. Planning fun lunches and social activities to get to know other developers.
6. organization, balance of coding / breakout sessions
7. Lots of time for interaction and collaboration. Overall good balance between updates, breakouts, and hacking.
8. The topics of the breakout sessions were very relevant.
9. Teambuilding through group activity, clear schedule, organized transitions to other locations
10. Good balance of different types of activities.
11. Great! It was a good opportunity to learn about projects that the broader NWB community is contributing to. It helped me understand overlaps and future direction on projects that each of us are individually contributing to.
12. The schedule was very organized from the beginning and there were a lot of opportunities to learn about what other developers were working on during the breakout sessions.

### What could we improve on at future hackathons?

1. Diversity
2. Maybe another day? So much to do, so little time
3. more opportunities for projects
4. Better dinner/lunch plans ahead of time
5. Maybe setting out specific time for small discussion groups on specific projects or topics
6. more non-core developers, open up projects for lower levels of experience
7. Broaden participation by tool developers outside of the core NWB tools
8. More, smaller rooms could have been helpful, as it seemed like there wasn't a lot of crossover after determining who worked on what
9. Having dedicated breakouts focused on possible long term collaborations or technology transfer (if possible)

### Any additional questions, comments or suggestions?

1. I would like tutorials / talks on existing but complicated parts of NWB by the developers (e.g. how talking to the backend during building works).
2. Thanks for the amazing workshop. I had a great time and appreciate all of the work by Ryan, Kushal, Ben, and the DataJoint team. Regarding planning future events... due to state anti-LGBTQ policies such as those denying healthcare to transgender youth, on principle I will not be able to attend any events in the future that take place in Texas or Florida. It would be great if you could take that into consideration.

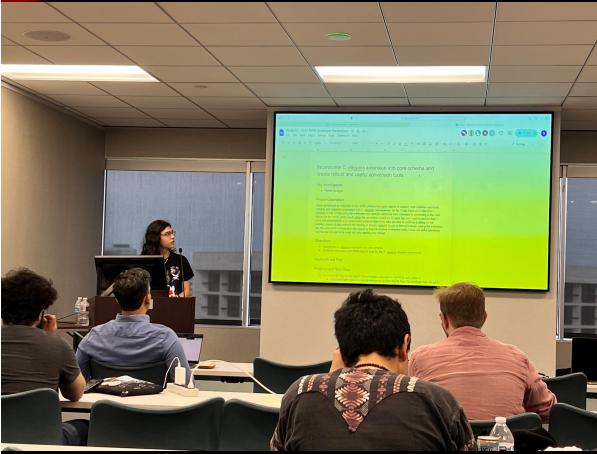
## Testimonials

*"Attending the NWB Hackathon was a great experience as a newcomer to NWB. I learned a great deal about the format and best practices from the people developing the standard. Everyone, from the developers to the other participants, was very helpful and willing to share their knowledge. I was able to make progress on my project and get a lot of help from the developers of the tools I was using, like NeuroConv. I would highly recommend the hackathon to anyone interested in working with NWB files."* Christian Tabedzki – Princeton Neuroscience Institute

## 6 Photos from the Event



# BREAKOUTS, TALKS & OTHER DISCUSSIONS



# 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 their great enthusiasm and for making the event a great success! 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). Thanks to DataJoint for welcoming us into their space, organizing local logistics, and coordinating the mixer event.

## Disclaimer

This document and related content were prepared as an account of or to expedite work sponsored at least in part by the United States Government. While we strive to provide correct information, neither the United States Government nor any agency thereof, nor The Regents of the University of California, nor any of their employees, makes any warranty, express or implied, or assumes any legal responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights.

Reference herein to any specific commercial product, process, or service by its trade name, trademark, manufacturer, or otherwise, does not necessarily constitute or imply its endorsement, recommendation, or favoring by the United States Government or any agency thereof, or The Regents of the University of California. Use of the Laboratory or University's name for endorsements is prohibited.

The views and opinions of authors expressed herein do not necessarily state or reflect those of the United States Government or any agency thereof or The Regents of the University of California. Neither Berkeley Lab nor its employees are agents of the US Government.

Links to external pages and documents included in the documents do not constitute an endorsement of the content or company and we are not responsible for the content of such links.