



# Report: 6th NWB:N Hackathon at HHMI Janelia

## User Days and Developer Hackathon

Oliver Rübel, Andrew Tritt, Ryan Ly, Benjamin Dichter eds.

Image courtesy of [www.nwb.org](http://www.nwb.org)

1

## Contents

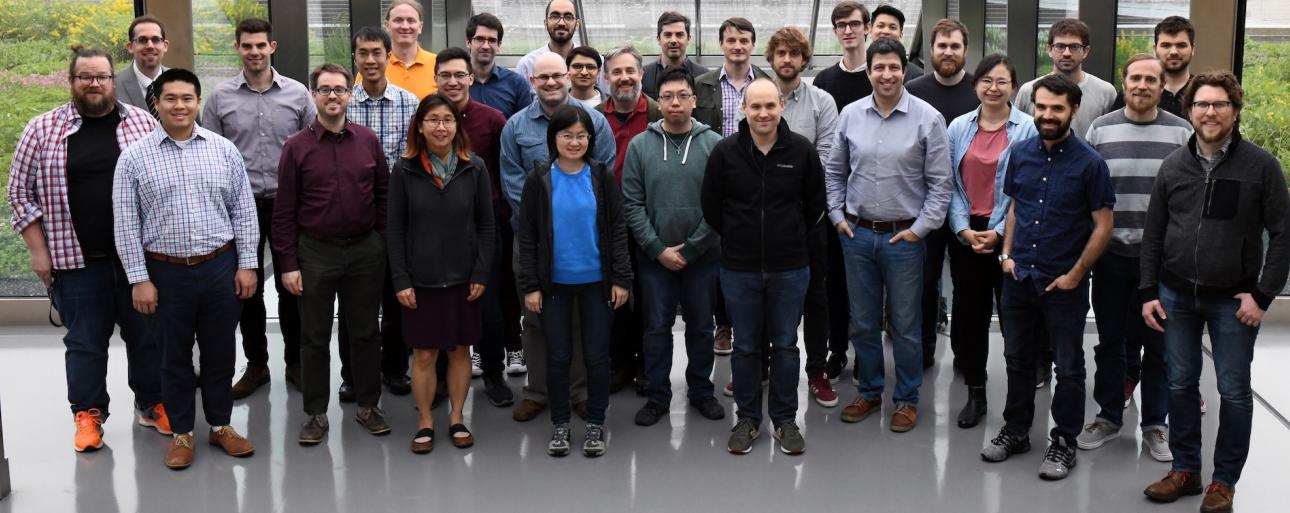
<b>Participants – NWB:N User Days</b>	<b>2</b>	<b>Code Activity (May 2019)</b>	<b>18</b>
<b>Participants – NWB:N Developer Days</b>	<b>3</b>	NWB:N Software Repositories . . . . .	18
<b>Executive Summary</b>	<b>4</b>	NWB:N Website . . . . .	19
Organization . . . . .	5		
Suggestions for Specific Developments . . . . .	6		
<b>Registration Survey</b>	<b>7</b>	<b>Event Photos</b>	<b>20</b>
Software Tools . . . . .	7		
Programming Languages . . . . .	7	<b>Acknowledgements</b>	<b>24</b>
Data Modalities . . . . .	8		
<b>Exit Survey</b>	<b>9</b>	<b>Disclaimer</b>	<b>24</b>
Hackathon Overall and Testimonials . . . . .	9		
Software . . . . .	10		
Talks and Tutorials . . . . .	10		
<b>Program</b>	<b>12</b>		
User Days Program . . . . .	12		
Developer Days Program . . . . .	13		
Time Breakdown . . . . .	13		
<b>Hackathon Projects</b>	<b>14</b>		
Data Import/Convert . . . . .	14		
Data Storage and Management . . . . .	15		
Data Analysis and Visualization . . . . .	16		
Core Development and Documentation . . . . .	16		

**NEURODATA:**  
WITHOUT BORDERS

THE  KAVLI FOUNDATION

 **BERKELEY LAB**  
Bringing Science Solutions to the World

  
**hhmi | janelia**  
Research Campus



Participants of the NWB:N User Days, May 13-14, 2019. Photo by Matt Staley, HHMI Janelia

## Participants – NWB:N User Days

1. Vyassa Baratham (University of California, Berkeley / Lawrence Berkeley National Laboratory)
2. Jim Berg (Allen Institute for Brain Science)
3. Zac Bowen (University of Maryland Biophysics)
4. Thomas Braun (byte physics)
5. Nand Chandravadia (Cedars-Sinai)
6. Nathan Clack (Vidrio Technologies LLC)
7. Tom Davidson (HHMI/University of California, San Francisco (Frank Lab))
8. Benjamin Dichter (Lawrence Berkeley National Laboratory / Stanford)
9. Nile Graddis (Allen Institute for Brain Science)
10. Michael Grauer (Kitware, Inc.)
11. Roman Huszar (Neuroscience Institute, NYU Langone Health (Buzsaki Lab))
12. Donghe Kang (The Ohio State University)
13. Adam Kepecs (Cold Spring Harbor Lab)
14. Seth Koehler (Decibel)
15. Andrew Landau (Harvard Medical School)
16. Jessie Liu (University of California, San Francisco (Chang Lab))
17. Ryan Ly (Lawrence Berkeley National Laboratory)
18. Lydia Ng (Allen Institute for Brain Science)
19. Thinh Nguyen (Vathes LLC)
20. Lawrence Niu (Vidrio Technologies LLC)
21. Peter Petersen (New York University)
22. James Roach (Cold Spring Harbor Lab)
23. Oliver Rübel (Lawrence Berkeley National Laboratory)
24. Kelson Shilling-Scriivo (University of Maryland)
25. Matthew Sit (Lawrence Berkeley National Laboratory (Bouchard Lab))
26. Andrew Tritt (Lawrence Berkeley National Laboratory)
27. Jochen Weber (Columbia University)
28. Oscar Woolnough (University of Texas Health Science Center at Houston)
29. Michael Wulf (Cold Spring Harbor Laboratory (Kepecs Lab))
30. Lingling Yang (University of Minnesota)
31. Dimitri Yatsenko (Vathes LLC)



Participants of the NWB:N Developer Days, May 15-16, 2019. Photo by Matt Staley, HHMI Janelia

## Participants – NWB:N Developer Days

1. Thomas Braun (byte physics)
2. Nicholas Cain (Allen Institute for Brain Science)
3. Nathan Clack (Vidrio Technologies LLC)
4. Tom Davidson (HHMI/University of California, San Francisco (Frank Lab))
5. Benjamin Dichter (Lawrence Berkeley National Laboratory / Stanford)
6. Matthew Earnshaw (University College London)
7. Mohammad Hossein Eybposh (University of North Carolina at Chapel Hill)
8. Jean-Christophe Fillion-Robin (Kitware inc.)
9. Andrea Giovannucci (University of North Carolina at Chapel Hill)
10. Nile Graddis (Allen Institute for Brain Science)
11. Michael Grauer (Kitware, Inc.)
12. Donghe Kang (The Ohio State University)
13. Christian Kellner (G-Node / Red Hat)
14. Mikkel Lepperød (University of Oslo)
15. Ryan Ly (Lawrence Berkeley National Laboratory)
16. Matt McCormick (Kitware Inc.)
17. Konstantinos Nasiotis (Montreal Neurological Institute)
18. Lydia Ng (Allen Institute for Brain Science)
19. Thinh Nguyen (Vathes LLC)
20. Lawrence Niu (Vidrio Technologies LLC)
21. Oliver Rübel (Lawrence Berkeley National Laboratory)
22. Will Schroeder (Kitware, Inc.)
23. Andrew Tritt (Lawrence Berkeley National Laboratory)
24. Zhengjia Wang (Baylor College of Medicine/Rice University)
25. Dimitri Yatsenko (Vathes LLC)

# Executive Summary

**Overview:** The Neurodata Without Borders: Neurophysiology (NWB:N) project is an effort to standardize the description and storage of neurophysiology data and metadata. NWB:N enables data sharing and reuse and reduces the energy barrier to applying data analytics both within and across labs. The hackathon consisted of a two-day User Days event and a two-day Developer Days event with a joint social on the second day to facilitate interaction between users and developers. For the User Days, we invited experts from the neuroscience research community to learn about NWB:N and to work on projects toward adoption of NWB:N for their lab's data sharing and analysis needs. For the Developer Days, we invited developers from the neuroinformatics community to integrate NWB:N with data analysis, visualization, storage, and management tools, as well as to discuss and design new features for NWB:N and build and improve core NWB:N infrastructure.

**Participants:** 43 users and developers from 29 major labs and research institutions attended the event (see page 2-3). The background in programming among participants at the User Days varied greatly, ranging from beginners to experts, whereas participants at the developer days generally had a strong background in programming. Most participants had investigated NWB:N prior to the event, but many were in the early stages of exploration and adoption. See also the [Registration Survey](#) for details.

**User Days:** The [User Days Program](#) consisted of a combination of training sessions in the morning and open hacking sessions in the afternoon. A breakout room for hacking was provided at all times to allow users to select the training programs most relevant to their projects. During the coding sessions, participants worked on their [Hackathon Projects](#), which they had designed and prepared prior to the event. The user projects generally focused on integration of neuroscience data with NWB:N and involved data from a broad range of [Data Modalities](#), including: a) electrophysiology (extracellular recordings, intracellular recordings, and electrocorticography), b) optophysiology (2-photon imaging, fluorescent wide-field images, etc.), c) behavioral data, d) processed data (e.g., from spike sorting), e) neural simulations (using NEURON software), and f) stimulus data. The user projects also included data from a number of different model species, such as mice, rats, monkeys, and humans.

**Developer Days:** The [Developer Days Program](#) focused on open hacking sessions and breakout sessions to facilitate interaction between developers and discussion on shared topics of interest and current developments, e.g., best practices, extensions sharing, ontologies, data management, and data visualization. The [Hackathon Projects](#) during the developer days focused on integration of NWB:N with data analysis, visualization, storage, and management software, and development and enhancement of core NWB:N technologies.

**Select Highlights:** During the user days, users worked on the integration of data from 14+ different labs and institutions with NWB:N. Importantly, all participants were able to make substantial progress during the hackathon towards integration of their data with NWB:N.

Several early adopters—e.g., the Allen Institute for Brain Science, the FrankLab (UCSF), and Buzsaki Lab (NYU)—have by now made first sets of NWB:N data files available to the public. Several other labs, e.g., the ChangLab (UCSF), have indicated plans for upcoming data releases in NWB:N 2.0. Given that NWB:N 2.0 was released in January 2019 (i.e., only  $\approx 3.5$  month before the hackathon), this fast move towards data releases is impressive and also very important to promote broad adoption of NWB:N. To enable users to easily find public NWB:N files, we have added an [Example Data](#) page to the NWB:N website.

A highlight here was the presentation by Tom Davidson in which he described time-based queries and visualization at the FrankLab. Incidentally, the Allen Institute had released data 1 week prior to the hackathon in NWB:N 2.0 and so they tried their tools on that data as well. “Because the data was in the NWB:N standard, we were able to very quickly and directly apply our tools to the Allen Institute’s data without requiring interaction with the Allen team” said Tom, “This showed us that the promise of data reuse and common analysis tools is not just a distant vision but because of NWB:N is within our reach.”

During the developer days, we were excited to see substantial progress toward integration of NWB:N with important data analysis tools, including, Brainstorm, CaImAn, RAVE, NWBWidgets, and the NWBExplorer as part of Open Source Brain. To enable users to more easily discover tools that support NWB:N, we created an [Analysis and Visualization Tools](#) page as part of the NWB:N website. Developers also explored the integration of Zarr and ExDir as alternative storage backends for NWB:N as well as compatibility between NWB:N and Nix and NWB:N and the DataJoint data pipeline management system. Developers also created tools to facilitate the import of data from common data acquisition devices and systems. Another main focus area was the development of core NWB:N technologies, e.g., for data query, extensions sharing, and documentation.

**Future Directions:** As part of the hackathon, participants suggested several areas that would greatly enhance the utilization and adoptability of NWB:N:

- **Extensions sharing:** The ability to share and use extensions across labs was recognized as a critical need. Formalizing and standardizing this process will be critical. This is part of current ongoing efforts towards development of the [NWB:N Extension Catalog](#).
- **Controlled Vocabularies and Ontologies:** The ability to associate controlled terms (e.g., enumerations) and complex ontologies with NWB:N data fields was recognized as a central need. This is already on the roadmap for NWB:N development as part of ongoing [NIH project](#) and we used the hackathon to interview participants and gather use-cases for the [NWB:N ontologies project](#).
- **Tools for visualization and analysis:** Tools for exploration and visualization of NWB:N was recognized as a general need. This was an active focus of the Developer Days and continues to be a critical thrust area in our outreach activities.
- **Data search and query:** Advanced tools for search, discovery, and query of NWB:N data are needed to facilitate advanced data analytics. This was also the topic of projects at the Developer Days but we need to expand our efforts in this area as part of future proposals.
- **Integration with Data Management:** The need for tools to manage large collections of NWB:N files and to facilitate integration of NWB:N with data management tools (e.g., LIMS or DataJoint) was recognized as a common need. While initial efforts in this area are underway, this is a critical area that should be the focus of future proposals.
- **Documentation, Tutorials, and Standard Import:** While we have made substantial progress over the last 2 years in creating documentation and tutorials, this area requires continued effort and improvement. In particular the need for additional high- and entry-level documentation was recognized and was also the focus of several Developer Days projects.

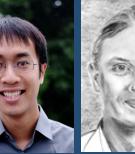
**Conclusion:** By the end of the hackathon, the vast majority of attendees had accomplished many of their goals, and several were well on their way to having NWB:N fulfill their labs needs. All attendees indicated an enthusiasm for the current software and were impressed with the (relative) ease of use given the generality. Overall, the hackathon was a great success. See also the [Exit Survey](#) for further details.

Based on discussions at the hackathon, we plan to have another hackathon in a year for both users and developers. In addition we are also hoping to have a developer-focused hackathon in November to facilitate continued development and close interaction between the developers, to discuss design and strategy, and engage more closely with user projects and use cases.

## Organization

The meeting was organized by Oliver Rübel (organizer), Benjamin Dichter (co-organizer), and Karel Svoboda (site chair) with logistics and organizational support by Janine Stevens (HHMI Janelia), AV support by Matt Staley (HHMI Janelia), and travel support by Stephanie Albin (Kavli Foundation). Travel support was provided by the Kavli Foundation. Housing, food and other meeting resources were sponsored by HHMI Janelia.

5

Site Chair	Organizing Committee		Logistics Support			Organizational Support		
								
Karel Svoboda HHMI Janelia	Oliver Rübel LBNL	Ben Dichter Data Science Consultant	Stephanie Albin Science Program Officer The Kavli Foundation	Janine Stevens Director, Scientific Programs HHMI Janelia	Matt Staley UC/AV Engineer HHMI Janelia			
Presenters								
								
Andrew Tritt LBNL	Nathan Clack Vidrio	Lawrence Niu Vidrio	Michael Grauer Kitware	Ryan Ly LBNL	Thomas Braun byte physics	Tom Davidson UCSF	Thinh Nguyen Vathes	Vyassa Baratham UC Berkeley

## Suggestions for Specific Developments

In addition to the more high-level direction for future development discussed in the [Executive Summary](#), we here describe additional, specific developments that were suggested by participants at the hackathon.

- Easily query within and across NWB:N files for particular data. For example, show me the recorded voltage traces after stimulus X for animals with manipulation Y that were tested between D1 and D2.
- Improve the tutorials to cater toward researchers with less programming experience. Users also mentioned that it would be very useful to include more graphics that depict how data is stored and have working boilerplate code where the user just needs to change a few variables to accommodate their data.
- Develop better tutorials to the MatNWB API, and add more convenience functions like in the PyNWB API.
- Add ability to define time points within a trial and re-reference timestamps and time series data relative to those time points. For example, users want to easily represent voltage traces aligned to reward time, across trials.
- Add ability to organize raw acquisition data rather than having them all in the same group. This is currently possible for processed data but not acquisition data.
- Create a plotting gallery for different neurodata types.
- Improve validation of NWB:N files.
- Add ability to dump state of PyNWB/MatNWB during read/write for better debugging.

# Registration Survey

As part of the registration we asked all participants about: which 1) software tools and 2) programming languages they are using to interact with their data as well as 3) which data modalities they are using. The questions were posed as multiple-choice questions with the option for users to provide custom other responses. 34 users registered for the User Days, of which 31 attended the User Days. 26 developers registered, of which 25 attended the Developer Days. In total, 45 unique people responded to the survey, i.e., some folks attended and responded to the survey for both the User- and Developer days.

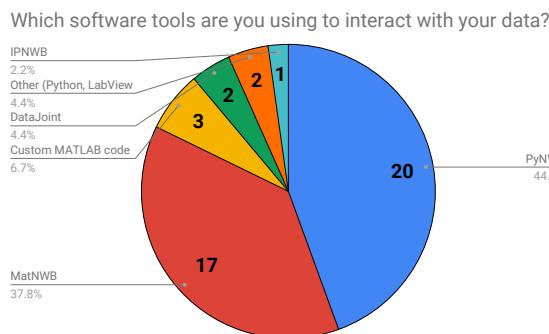
## Software Tools

**Survey Question:** Which software tools are you using to interact with your data?

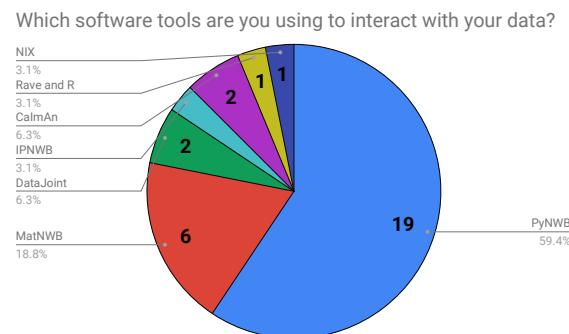
**Response Options:** 1) PyNWB, 2) MatNWB, 3) Other (user provides text)

**Number of User Days Responses:** 34 of 34

**Number of Developer Days Responses:** 25 of 25



(a) NWB:N User Days



(b) NWB:N Developer Days

Figure 1: Registration survey responses for the user days (a) and developer days (b) on: “Which software tools are you using to interact with your data?”

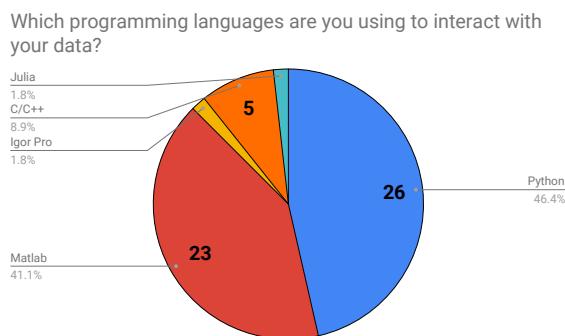
## Programming Languages

**Survey Question:** Which programming languages are you using to interact with your data?

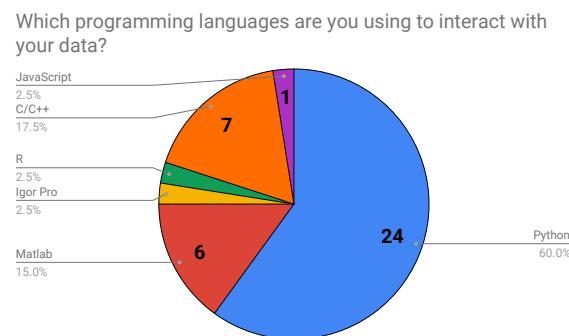
**Response Options:** 1) Python, 2) Matlab, 3) C/C++, 4) Julia, 5) Java, 6) Other (user provides text)

**Number of User Days Responses:** 34 of 34

**Number of Developer Days Responses:** 25 of 25



(a) NWB:N User Days



(b) NWB:N Developer Days

Figure 2: Registration survey responses for the user days (a) and developer days (b) on: “Which programming languages are you using to interact with your data?”

## Data Modalities

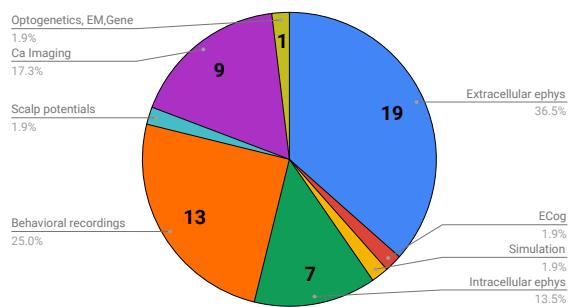
**Survey Question:** Which data modalities are you using?

**Response Options:** 1) Extracellular ephys, 2) Intracellular ephys, 3) Ca imaging, 4) Behavioral recordings, 5) Other (user provides text)

**Number of User Days Responses:** 28 of 34

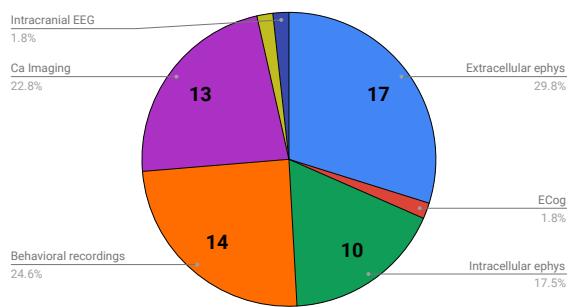
**Number of Developer Days Responses:** 22 of 25

Which data modalities are you using?



(a) NWB:N User Days

Which data modalities are you using?



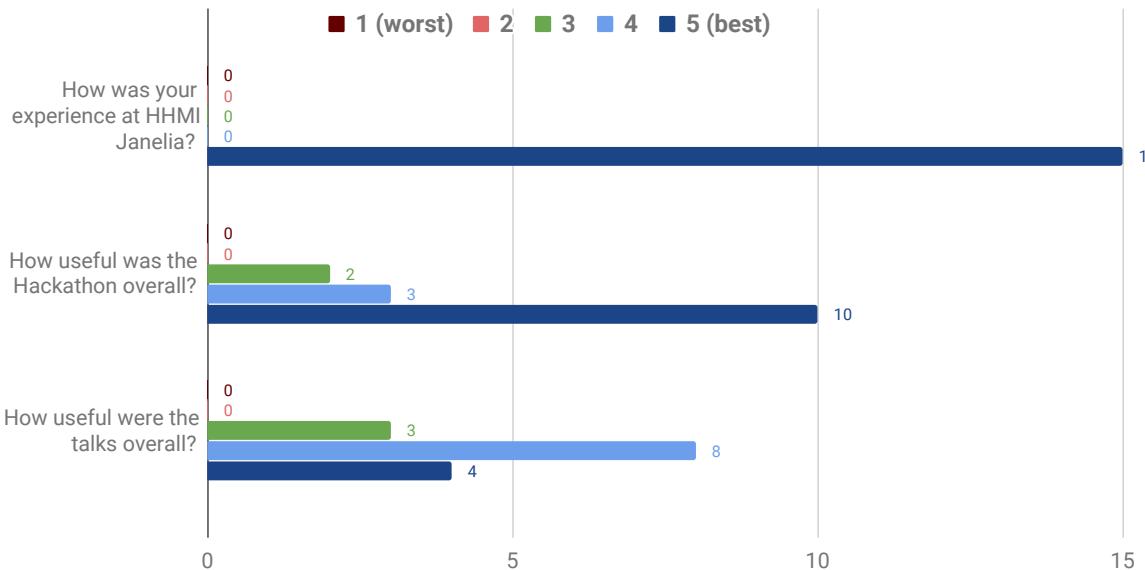
(b) NWB:N Developer Days

Figure 3: Registration survey responses for the user days (a) and developer days (b) on: “Which data modalities are you using?”

## Exit Survey

All participants were asked to participate in an exit survey. 15 of the 42 participants (i.e., 31%) submitted responses to the survey. 12 of the 15 respondents attended the user days.

### Hackathon Overall and Testimonials



On a scale of 1 (worst) to 5 (best) the average ratings for the hackathon were excellent, which is also reflected in the testimonials the participants provided:

- 5 : Experience at HHMI Janelia
- 4.5 : Usefulness of the hackathon overall
- 4.1: Usefulness of the talks overall

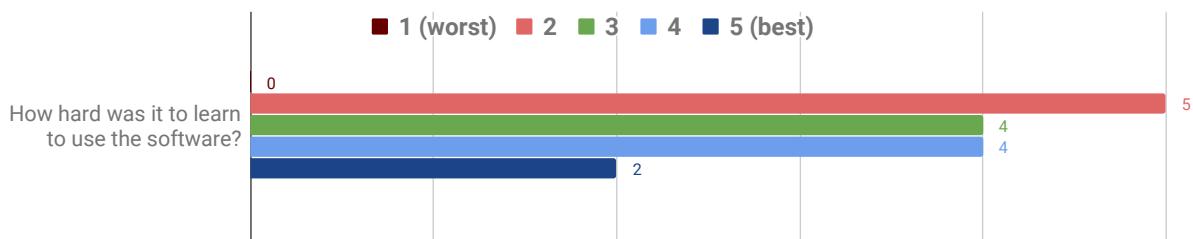
### Testimonials

As part of the exit survey we asked folks whether they would like to provide any testimonials. The exit survey was anonymous, i.e., testimonials are also anonymous unless participants submitted them via email.

9

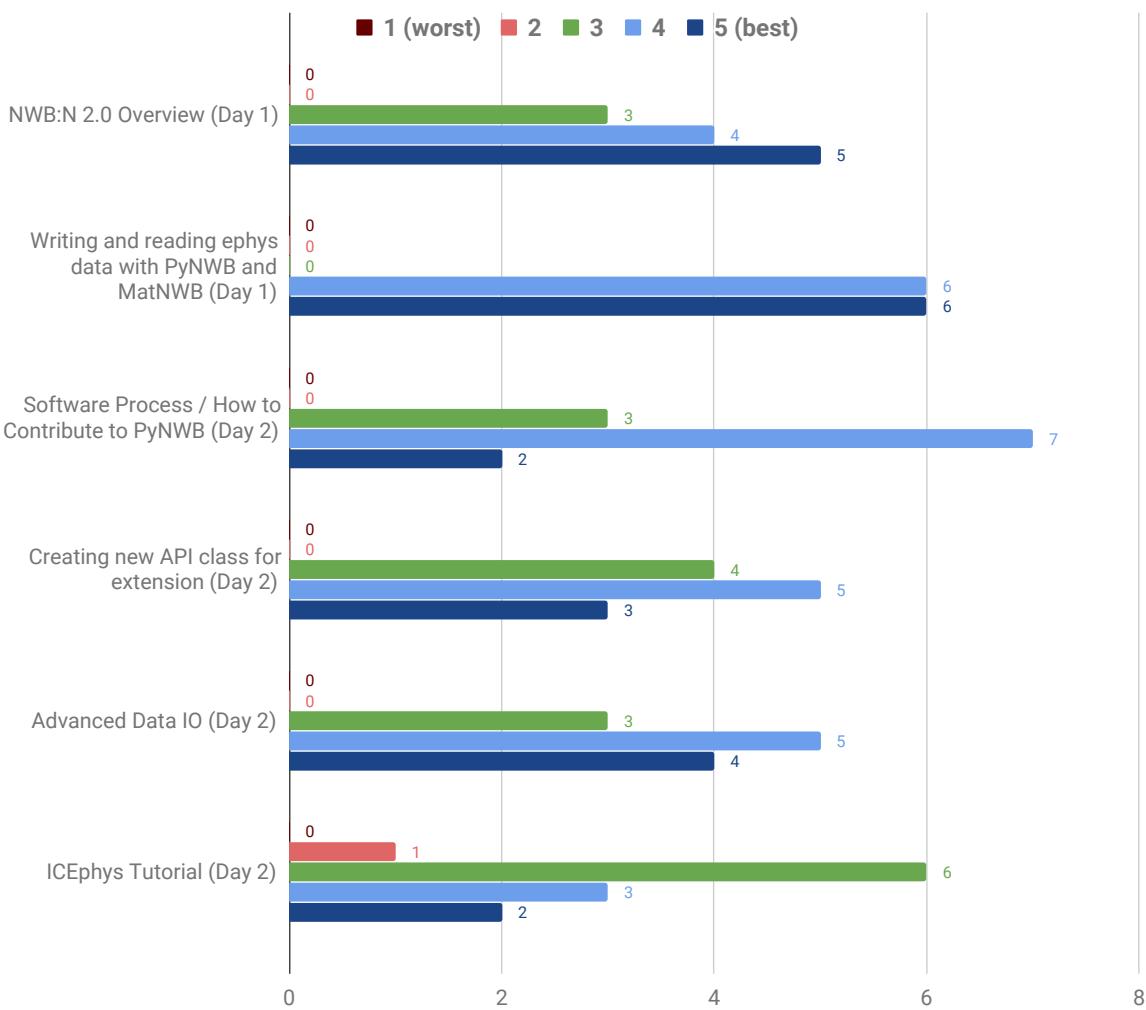
- The NWB Hackathon User Days was incredibly useful. Well organized, developers were always present and helpful with answering questions. They created an environment of users and developers coming together to help each other, it was refreshing and so conducive to all the good work that has come out of it!
- Coming from zero experience with NWB, the hackathon was very useful to have the developers there to aid with questions the whole time. The whole developer team was so receptive to feedback and quick to make changes to make the users experience better. It was nice to connect with other researchers who are facing similar data management issues and trying to move to NWB. The Janelia research campus was a great venue for the event.
- The organizers did an awesome job putting together a hackathon that provided (1) a needed refresher and overview, (2) opportunity to ask questions, and (3) space to work on the datasets in the hacking sessions. Overall close to perfect balance. Thank you!!
- Michael Grauer (Kitware): “I wanted to say thanks and congrats for putting on a fantastic hackathon. Personally I really enjoyed it, it was super productive for both me and the Kitware team. I think this gave the community a lot of cohesion and momentum, which is a great sign for things to come.”

## Software



On a scale of 1 (worst) to 5 (best) the average ratings for how hard it was to learn to use the software was **3.2**.

## Talks and Tutorials



Overall the talks were well-received, with average ratings of **4.2**, **4.5**, **3.9**, **3.9**, **4.1**, and **3.5**, respectively.

To further improve the ease of on-boarding as folks adopt NWB:N, survey participants suggested the creation of simple, reusable examples. Specific suggestions from the exit survey included:

- “The tutorials provided were very helpful but also have room for improvement (and they have been updated already which is nice). I think it needs to be more clear which aspect of the tutorials are or are not necessary to create a NWB file (i.e. I would start from the most basic bare-bones NWB file and

build up from there). I think some generalized examples of how to add in different types of data would go a long way, with detail explaining the intention of each step and why things are stored the way they are. Maybe these exist and I just missed them. It might be too much to ask, but usage examples would be helpful on the docs for each different type ([nwb-schema.readthedocs.io](https://nwb-schema.readthedocs.io))."

- "To simplify people getting started with, especially, their format conversion projects, having a collection of boilerplate codes (that is approved and tested by NWB contributors!) would be super helpful, IMO."

In response to specific user suggestion we have clarified and updated many of the existing online tutorials. We also started a project with the goal to create an [interactive online course for PyNWB](#). Improving and expanding documentation and tutorials is an important and perpetual effort. Our goal is to make NWB:N as easy-to-use as possible, and documentation and tutorials are key parts of that.

# Program

- **Event Website:** [https://neurodatawithoutborders.github.io/nwb\\_hackathons/HCK06\\_2019\\_Janelia/](https://neurodatawithoutborders.github.io/nwb_hackathons/HCK06_2019_Janelia/)
- **Dates:** May 13-16, 2019
  - **User Days:** May 13-14, 2019
  - **Developer Days:** May 15-16, 2019
- **Location:** Location: HHMI Janelia Research Campus, 19700 Helix Dr, Ashburn, VA 20147
- **Slides:** The slides from the talks and tutorials presented at the event are available online at: [https://neurodatawithoutborders.github.io/nwb\\_hackathons/HCK06\\_2019\\_Janelia/#resources](https://neurodatawithoutborders.github.io/nwb_hackathons/HCK06_2019_Janelia/#resources)

## User Days Program

Color Legend								
Talks	Breaks	Hacking	Tutorials with hacking excercises	Group discussions				
<b>Day 1 (Monday, May 13): User Days</b>								
Start Time	End Time	Duration	Topic	Speaker	Room	Topic	Speaker	Room
7:00 AM	8:00 AM	1:00:00	Breakfast	Everyone	Dining Room	No parallel session		
8:00 AM	8:15 AM	0:15:00	Welcome and Overview of the Meeting	Oliver Ruebel	Synapse			
8:15 AM	9:00 AM	0:45:00	NWB:N 2.0: Overview - Ecosystem components - Overview of spec language, data storage, format specification, PyNWB, MatNWB	Oliver Ruebel	Synapse	Hacking on projects	None. Developer in the room to help with questions	Spectrum
9:00 AM	9:15 AM	0:15:00	MatNWB Update	N. Clack / Lawrence Niu	Synapse			
9:15 AM	9:30 AM	0:15:00	PyNWB Update	Andrew Tritt	Synapse			
9:30 AM	9:45 AM	0:15:00	Break	Everyone	Synapse	Break	Everyone	Spectrum
9:45 AM	11:45 AM	2:00:00	Writing and reading ephys data with PyNWB and MatNWB + Overview of Simulation Extension	Ben Dichter	Synapse	Hacking on projects	None. Developer in the room to help	Spectrum
11:45 AM	12:00 PM	0:15:00	Group Photo (User days)	Everyone	Lobby	No parallel session		
12:00 PM	1:00 PM	1:00:00	Lunch	Everyone	Dining Room			
1:00 PM	1:30 PM	0:30:00	Overview of hackathon projects and use-cases	Everyone	Synapse			
1:30 PM	2:15 PM	0:45:00	NWB:N Early Adopter Experiences: NWB:N at the Bouchard Lab Interoperability between DataJoint and Case study: Importing legacy tetrode data, and working with time-based queries.	Vyassa Baratham Thinh Nguyen Tom Davidson	Synapse	Hacking on projects	None. Developer in the room to help with questions	Spectrum
2:15 PM	3:30 PM	1:15:00	Hacking on projects	Everyone	Synapse	Hacking on projects	Everyone	Spectrum
3:30 PM	3:50 PM	0:20:00	Break	Everyone	Synapse	Break	Everyone	Spectrum
3:50 PM	6:00 PM	2:10:00	Hacking on projects	Everyone	Synapse	Hacking on projects	Everyone	Spectrum
6:00 PM	7:30 PM	1:30:00	Dinner	Everyone	Dining Room	No parallel session		

Day 2 (Tuesday, May 14): User Days								
			Room 1 (Synapse)		Room 2 (Spectrum)			
Start Time	End Time	Duration	Topic	Speaker	Room	Topic	Speaker	Room
8:00 AM	9:00 AM	1:00:00	Breakfast	Everyone	Dining Room	No parallel session		
9:00 AM	9:15 AM	0:15:00	Welcome to day 2	Oliver Ruebel	Synapse			
9:15 AM	9:30 AM	0:15:00	Software Process / How to Contribute to PyNWB	Michael Grauer	Synapse	Hacking on projects	None. Developer in the room to help with questions	Spectrum
9:30 AM	10:30 AM	1:00:00	Creating new API classes for extensions	Andrew Tritt	Synapse			
10:30 AM	10:45 AM	0:15:00	Break		Synapse	Break	Everyone	Spectrum
10:45 AM	11:25 AM	0:40:00	Advanced Data I/O (PyNWB)	Ryan Ly	Synapse	Hacking on projects	None. Developer in the room to help with questions	Spectrum
11:25 AM	12:00 PM	0:35:00	ICEphys Tutorial	Thomas Braun	Synapse			
12:00 PM	1:00 PM	1:00:00	Lunch		Dining Room	Break		Dining Room
1:00 PM	2:30 PM	1:30:00	Hacking on projects	Everyone	Synapse / Spectrum	Hacking on projects	Everyone	Spectrum
2:30 PM	2:45 PM	0:15:00	Break	Everyone	Synapse	Break	Everyone	Spectrum
2:45 PM	4:30 PM	1:45:00	Hacking on projects	Everyone	Synapse / Spectrum	Hacking on projects	Everyone	Spectrum
4:30 PM	4:50 PM	0:20:00	NWB:N User and Developer Social (gather)	Everyone	Synapse	No parallel session		
4:50 PM	5:30 PM	0:40:00	User Days: Closing discussions (Developers in attendance)(Social group photo if possible)	Everyone	Synapse			
5:30 PM	7:00 PM	1:30:00	Social continues with dinner at 6pm	Everyone	Dining Room			

# Developer Days Program

Color Legend						
Talks	Breaks	Hacking	Tutorials with hacking excercises	Group discussions		
<b>Day 3 (Wednesday, May 15): Developer Hackathon</b>						
Start Time	End Time	Duration	Room 1 (Synapse)			Room 2 (Spectrum)
<b>See the social at the end of day 2</b>			Developers should plan to arrive in the afternoon on day 2 be able to attend the social			No parallel session
8:00 AM	9:00 AM	1:00:00	Breakfast		Everyone	Dining Room
9:00 AM	9:15 AM	0:15:00	Welcome to day 3		Michael Grauer / Oliver	Synapse
9:15 AM	9:35 AM	0:20:00	Review of outcomes from user days		Ben Dichter / Oliver Ruebel	Synapse
9:35 AM	10:35 AM	1:00:00	Project overviews. Each project provides 5min overview of their project plan. Planning for new projects.		Everyone	Synapse
10:35 AM	10:50 AM	0:15:00	Break		Everyone	Synapse
10:50 AM	12:00 PM	1:10:00	Hacking on projects		Everyone	Synapse
12:00 PM	1:00 PM	1:00:00	Lunch		Everyone	Dining Room
1:00 PM	1:10 PM	0:10:00	Developer Group Photo		Everyone	Lobby
1:10 PM	3:10 PM	2:00:00	Hacking on projects		Everyone	Synapse
3:10 PM	3:30 PM	0:20:00	Break		Everyone	Synapse
3:30 PM	6:00 PM	2:30:00	Hacking on projects		Everyone	Synapse
6:00 PM	7:30 PM	1:30:00	Dinner		Everyone	Dining Room
<b>Day 4 (Thursday, May 16): Developer Hackathon</b>						
Start Time	End Time	Duration	Room 1 (Synapse)			Room 2 (Spectrum)
8:00 AM	9:00 AM	1:00:00	Breakfast		Everyone	Dining Room
9:00 AM	9:10 AM	0:10:00	Welcome to day 4		Oliver Ruebel	Synapse
9:10 AM	9:35 AM	0:25:00	Morning check in. Review of outcomes from day 3 and plans for day 4.		Everyone	Synapse
9:35 AM	10:35 AM	1:00:00	Hacking on projects		Everyone	Synapse
10:35 AM	10:50 AM	0:15:00	Break		Everyone	Synapse
10:50 AM	12:00 PM	1:10:00	Hacking on projects		Everyone	Synapse
12:00 PM	1:00 PM	1:00:00	Lunch		Everyone	Dining Room
1:00 PM	2:30 PM	1:30:00	Hacking on projects		Everyone	Synapse
2:30 PM	2:45 PM	0:15:00	Break		Everyone	Synapse
2:45 PM	3:30 PM	0:45:00	Developer days: Closing discussion		Everyone	Synapse
3:30 PM	6:00 PM	2:30:00	Hacking on projects		Everyone	Synapse

## Time Breakdown

User Days: Time per Category						
	Day 1 (Main)	Day 1 (Breakout)	Day 2 (Main)	Day 2 (Breakout)	Total (Main)	Total (Breakout)
<b>Talks</b>	2:15:00	0:00:00	0:30:00	0:00:00	<b>2:45:00</b>	<b>0:00:00</b>
<b>Tutorials with hacking excercises</b>	2:00:00	0:00:00	2:15:00	0:00:00	<b>4:15:00</b>	<b>0:00:00</b>
<b>Hacking</b>	3:25:00	7:25:00	3:15:00	5:45:00	<b>6:40:00</b>	<b>13:10:00</b>
<b>Breaks (excluding breakfast and dinner)</b>	1:35:00	1:35:00	1:30:00	1:30:00	<b>3:05:00</b>	<b>3:05:00</b>
<b>Group discussions</b>	0:45:00	0:00:00	2:30:00	0:00:00	<b>3:15:00</b>	<b>0:00:00</b>
<b>Totals time per day</b>	<b>10:00:00</b>		<b>10:00:00</b>		<b>20:00:00</b>	

Developer Days: Time per Category						
	Day 3 (Main)	Day 3 (Breakout)	Day 4 (Main)	Day 4 (Breakout)	Total (Main)	Total (Breakout)
<b>Talks</b>	0:35:00	0:00:00	0:10:00	0:00:00	<b>0:45:00</b>	<b>0:00:00</b>
<b>Tutorials with hacking excercises</b>	0:00:00	0:00:00	0:00:00	0:00:00	<b>0:00:00</b>	<b>0:00:00</b>
<b>Hacking</b>	5:40:00	2:30:00	6:10:00	2:30:00	<b>11:50:00</b>	<b>5:00:00</b>
<b>Breaks (excluding breakfast and dinner)</b>	1:35:00	1:20:00	1:30:00	1:30:00	<b>3:05:00</b>	<b>2:50:00</b>
<b>Group discussions</b>	1:10:00	3:10:00	1:10:00	3:40:00	<b>2:20:00</b>	<b>6:50:00</b>
<b>Totals time per day</b>	<b>9:00:00</b>		<b>9:00:00</b>		<b>18:00:00</b>	

# Hackathon Projects

Prior to the event, each participant was asked by the organizers to formulate a specific project that she/he will work on. The participants created project pages on the [hackathon GitHub repository](#). Participants then updated their pages during (and after) the event to document progress. In the following we provide a brief overview of the various projects. The project headings include links to the specific project pages for further details.

## Data Import/Convert

1. [Integrate neuron model optimization and simulation data with NWB:N](#) The goal of the project is to use NWB:N to store simulated neuron model data, as well as stimuli and parameter sets, instead of the current method of storing data in csv files. The project has made significant progress in identifying native NWB:N data structures to use, determining how best to organize data in NWB:N, learning how to use extensions for the simulation output, and integrating NWB:N with the stimulus generation code.
  - **Key Investigators:** Matthew T. Sit and Roy Ben-Shalom
  - **Key Institutions:** University of California Berkeley (UCB), University of California San Francisco (UCSF), and Lawrence Berkeley National Laboratory (LBNL)
2. [Convert data in Abf and Dat formats to NWB:N](#) The `x_to_nwb` package has been developed to convert two of the most common formats for *in vitro* physiology data – `.abf` (pClamp, Molecular Devices) and `.dat` (PatchMaster, HEKA) – into the NWB:N format. These tools are new and this project will be focused on investigating their compatibility. The objectives of this project include to 1) determine if there are any lingering issues to be addressed in the latest version of the abf and dat file to nwb converter and to 2) brainstorm strategies for increasing visibility of abf\_and\_dat\_to\_nwb functionality. The project has made significant progress in that: 1) the latest version of the converter works well with several recent example abf and dat files and 2) a list of ways to increase outreach has been started, which includes pulling the converter out of the ipfx package in the Allen SDK and giving it its own repo as well as making sure that the code is clear on the expectations for maintenance.
  - **Key Investigators:** Jim Berg and Thomas Braun
  - **Key Institutions:** Allen Institute for Brain Science (AIBS) and byte physics
3. [Integrate ChangLab ECoG, behavioral, and imaging data with NWB:N](#) The Chang Lab records from a TDT acquisition system and has an in-house preprocessing pipeline for converting the raw TDT files into raw and preprocessed HTK files. The goal is to update this pipeline to be completely in Python and to utilize the NWB:N format. The objectives of this project include to: 1) Gain experience integrating the most common types of lab data (TDT/HTK, behavioral logs, imaging) into .nwb files and 2) be able to help lab members convert their data/be familiar with existing efforts to convert lab data to NWB:N format. Progress for this project include: 1) developed Python script for conversion of TDT to NWB:N files (including cortical surface using an extension), 2) stored frequency decomposition as DecompositionSeries, 3) used visualization tools with NWB:N files, 4) verified files.
  - **Key Investigators:** Jessie R. Liu
  - **Key Institutions:** Chang Lab, UCSF
4. [Integrate data from the Neuromodulation Research Center \(UMN\)](#) The goal of this project is to integrate all types of data from the Neuromodulation Research Center with NWB:N, including 1) raw data from TDT recording system, 2) raw behavioral and motion analysis system data (e.g., touch pad and eye movement data), 3) pre-processed data and 4) general experiment metadata. Accomplishments as part of the hackathon include: 1) converted raw data from TDT, 2) converted data from motion analysis system, and 3) learned details about the structure and usage of NWB:N.
  - **Key Investigators:** Lingling Yang
  - **Key Institutions:** University of Minnesota
5. [Integrate Thorlabs 2-photon calcium imaging data with NWB](#) The Kanold lab at UMD uses 2-photon calcium imaging to study the auditory cortex. The goal is to convert 2-photon calcium imaging data, both raw and processed, into the NWB:N format in order to have a generalized output from the Thorlabs microscope. Accomplishments as part of the hackathon include: 1) converted raw and registered imaging data as well as experimental metadata to NWB, 2) learned about structure and usage of NWB:N for imaging and behavioral data.
  - **Key Investigators:** Zac Bowen and Kelson Shilling-Scrivo
  - **Key Institutions:** University of Maryland
6. [Convert Tandon lab ECoG data to NWB](#)
  - **Key Investigators:** Oscar Woolnough

- **Key Institutions:** University of Texas
7. [Convert Rutishauser lab data to NWB](#) The goal of the project is to convert native Rutishauser lab data – electrophysiology and behavior – into the NWB:N format and modify the existing analysis pipeline to use NWB.
- **Key Investigators:** Nand Chandravadia
  - **Key Institutions:** Cedars-Sinai
8. [Create demonstration code to convert data from Zuckerman Institute to NWB](#) The goal of the project is to demonstrate to the Zuckerman Institute's faculty and their labs, associated with one of the two U19 projects (Costa and Miller), that data routinely used in data processing pipelines in the labs can be readily converted into and stored as NWB:N files. The objectives include identifying the mapping between existing metadata and NWB:N fields, creating a first-pass converter each of five datasets, and testing that NWB:N data can be read back and interchanged between Python and Matlab environments.
- **Key Investigators:** Jochen Weber
  - **Key Institutions:** Columbia Zuckerman Institute
9. [Integrate simultaneously recorded single-cell physiology and imaging data with NWB](#) The goals of the project are to integrate single-cell physiology and imaging data in NWB:N and create an extension for kymograph data.
- **Key Investigators:** Andrew Landau
  - **Key Institutions:** Harvard Medical School
10. [Convert Buzsaki lab data to NWB](#) The goal of the project is to generate scripts that load data stored in an internal lab data format (“buzcode”) to NWB.
- **Key Investigators:** Roman Huszar and Peter Petersen
  - **Key Institutions:** New York University
11. [Convert Mesoscale Activity Project data to NWB](#) The goal of the project is to design export tools for the [Mesoscale Activity Project](#) into NWB. This involves first mapping DataJoint schema into the NWB:N format and then developing automatic export routines.
- **Key Investigators:** Dimitriy Yatsenko and Thinh Nguyen
  - **Key Institutions:** Vathes LLC
12. [Convert Brain Modeling Toolkit output to NWB](#) The goal of the project is to use NWB:N as an output format for the Brain Modeling Toolkit to record cell variables, utilizing the NWB:N simulation\_output extension produced at a previous hackathon.
- **Key Investigators:** Vyassa Baratham
  - **Key Institutions:** LBNL
13. [Import data from acquisition systems](#) The objectives of this project include adding support and writing conversion scripts for the following acquisition systems: TDT, Plexon, Intan, Ripple, and Blackrock.
- **Key Investigators:** Konstantinos Nasiosit
  - **Key Institutions:** McGill University

## 15 Data Storage and Management

14. [Integrate ZARR with HDMF/PyNWB](#) The goal of this project is to integrate ZARR as an I/O backend with HDMF.
- **Key Investigators:** Donghe Kang and Oliver Rübel
  - **Key Institutions:** Ohio State University and LBNL
15. [Implement Exdir as an alternative backend to HDF5](#) Exdir is a hierarchical data structure similar to HDF5 which operates on the filesystem instead of a HDF5 file. Having Exdir as a backend in NWB:N would enable users to operate directly on the filesystem, which would facilitate human readability of metadata and may work better with version control systems such as git and sharing of partial data sets. The goals of this project are to implement object and region references in Exdir and to then implement a working version of Exdir in HDMF.
- **Key Investigators:** Mikkel Elle Lepperød
  - **Key Institutions:** CINPLA, University of Oslo, Norway
16. [Evaluate possible integration between NIX and NWB:N](#) The goal of the project is to explore possibilities for integration between the NIX and NWB:N standards and software ecosystems, including conversion of select data types between NIX and NWB:N and identification of possible avenues for future collaboration between NIX and NWB.
- **Key Investigators:** Christian Kellner and Oliver Rübel
  - **Key Institutions:** G-Node, Red Hat, and LBNL

## Data Analysis and Visualization

17. [Integrate NWB:N with RAVE](#) The goal of this project is to create “R” code that will allow our ECOG package, RAVE, to read in (as a start) voltage traces from NWB:N files. Progress for this project include development of a working converter script for reading NWB:N files into RAVE and a [working visualization demo](#) with ECoG data stored in NWB.
  - **Key Investigators:** Zhengjia Wang
  - **Key Institutions:** Baylor College of Medicine
18. [NWB:N Jupyter widgets](#) The [nwbwidgets project](#) allows for interactive exploration of an NWB:N file with built-in plotting of recognized data types within Jupyter. The goal of the project is to detect when an ECoG cortical surfaces is stored in an NWB:N file and create a 3D, interactive visualization of the surface. Considerable progress was made in [visualizing behavioral data](#) from the Allen Institute, and some progress was made in visualizing a cortical surface.
  - **Key Investigators:** Matt McCormick and Ben Dichter
  - **Key Institutions:** Kitware and LBNL
19. [Visualizing contents of NWB:N files with NWB Explorer and Open Source Brain Open Source Brain](#) is an open source, web-based resource for visualizing, simulating and disseminating standardized models of neurons and circuits from many brain regions. The primary goal at this hackathon is to gather existing NWB:N datasets to check and improve compatibility with the basic parser of NWB Explorer, determine scientifically useful workflows and the components necessary to support them, and develop a basic widget for exposing in a readable fashion the structure of an NWB:N file. The approach includes collecting example datasets and conversion scripts for the [NWB Showcase](#), and assessing attendee demand for functionality in NWB Explorer.
  - **Key Investigators:** Matt Earnshaw, Padraig Gleeson, and Matteo Cantarelli Gleeson
  - **Key Institutions:** University College London and MetaCell

## Core Development and Documentation

20. [Implement lazy cross-file links](#) The Allen plans to release large data sets in the NWB:N format. The goal of the project is to enable downloading of components of an NWB:N file, e.g. only the LFP data from multiple neuropixels probes.
  - **Key Investigators:** Nile Graddis
  - **Key Institutions:** AIBS
21. [Use case gathering](#) The goal of the project is to collect experimental and analysis use cases from participants of the User Days to inform community needs in the development of controlled vocabularies and ontologies. Progress for this project include interviews of five groups of users and creation of a [GitHub repo](#) to store the interview summaries.
  - **Key Investigators:** Lydia Ng, Michael Grauer
  - **Key Institutions:** AIBS, Kitware
22. [Develop infrastructure for extension sharing](#) The goal of the project is to evaluate plans for infrastructure for web-based sharing, deployment, and evaluation of extensions, and also develop the infrastructure for the online archive of NWB:N extensions. Progress for this project include receiving feedback from developers on the user flowchart for developing, sharing, and deploying extensions, adding continuous integration using Azure Pipelines to the [ndx-template](#) repo and related repos, consolidating documentation, reviewing conda-forge infrastructure, and planning continuous integration and automation for the [staged-extensions](#) repo.
  - **Key Investigators:** Ryan Ly, Jean-Christophe Fillion-Robin, Oliver Rübel, Andrew Tritt, and Ben Dichter
  - **Key Institutions:** LBNL, Kitware
23. [DataJoint-NWB-Interoperability](#) DataJoint is a data processing pipeline management system that has been successful in various large-scale neuroscience collaborations. The objectives of this project include discussing how to integrate DataJoint with NWB:N to make them interoperable; that is, how to allow data stored in NWB:N to be used with DataJoint, and how to export data from DataJoint to NWB. Progress includes discussion of four possible routes for DataJoint-NWB integration: 1. through DataJoint canonical pipelines, 2. by using NWB:N datatypes in DataJoint and implementing an archival adapter, 3. by inferring DataJoint schemas from NWB:N types and using DataJoint as a back-end, and 4. by building DataJoint schemas using a DataJoint schema template inferred from the NWB:N YAML schema files plus human intervention. Of these, options 4 and 1 were the most promising and will be explored further.

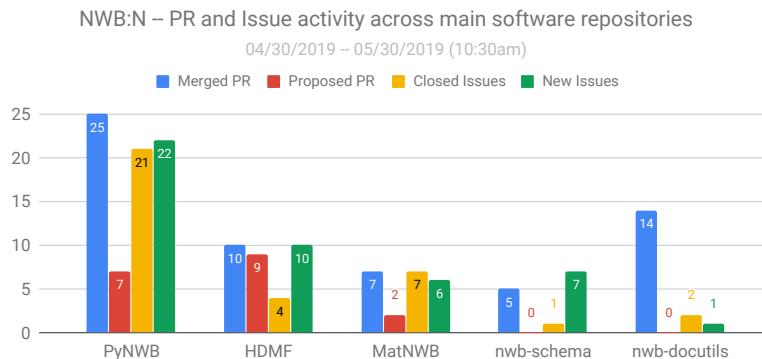
- **Key Investigators:** Dimitri Yatsenko, Thinh Nguyen, Andrew Tritt, and Oliver Rübel
  - **Key Institutions:** Vathes LLC, LBNL
24. [NWBIndex](#) The NWBIndex project is intended to help with the problem of sharing NWB:N files, to allow for discoverability, browsing, search, and accessibility. It is a long-range project intended to build tools to index existing NWB:N files and provide guidance to groups wanting to share their data. The primary goal for this hackathon, similar to the OSB NWB Explorer Project, is to gather examples of NWB:N datasets and forge collaborations with groups producing and consuming NWB:N data. The approach is to meet with participants to ask about NWB:N data for sharing and data sharing needs; collect public datasets along with metadata, and ideas about useful visual summaries; discuss with participants current approaches and pain points to sharing data, and find commonalities; and understand participants' data release plans, build collaborations to improve usability and discoverability of that data.
- **Key Investigators:** Mike Grauer
  - **Key Institutions:** Kitware
25. [Query and complex indexing for NWB:N data](#) The objectives of this project is to develop an implementation plan for query/complex slicing in NWB. It was discussed how to enhance or rewrite the Query class of HDMF to take and perform slices on a target, e.g. TimeSeries, add defined-over/domain/support/observation-intervals to TimeSeries class, add an abstract Mask class to HDMF, and add a class to represent list of non-overlapping intervals. The next step is to move these into GitHub as issues, and start writing code.
- **Key Investigators:** Andrew Tritt, Tom Davidson, and Nile Graddis
  - **Key Institutions:** LBNL, UCSF, AIBS
26. [Greatly enhance documentation](#) The documentation currently does not link to internal types and some external types accessible via intersphinx. Progress for the project include [major fixes](#) to this issue, which have since been merged into the main codebase.
- **Key Investigators:** Thomas Braun
  - **Key Institutions:** byte physics
27. [Enhance the validation tool](#) The validation tool currently does not support checking against a stored specification. In addition, it does not warn about unknown entries in the NWB:N file. Progress for the project include discussing how to best handle various issues with the validator.
- **Key Investigators:** Thomas Braun
  - **Key Institutions:** byte physics
28. [Interactive PyNWB Course](#) Users described a large learning curve for understanding the NWB:N format and using the APIs. The goal of the project is to create a more beginner-friendly, interactive tutorial/mini-course on NWB, starting with PyNWB. Open-source software was used to develop the interactive course, and an existing tutorial was adapted to become Chapter 1. A live demo of the course is linked [here](#).
- **Key Investigators:** Ryan Ly and Ben Dichter
  - **Key Institutions:** LBNL
29. [Develop list of best practices](#) To enable NWB:N to accommodate the needs of the diverse neuroscience community, NWB:N provides a great degree of flexibility. While this flexibility is essential to enable coverage of a broad range of use-cases, it can also lead to ambiguity. Developers and experienced users came together during a Developer Days breakout session to create a [list of best practices](#), which outlines some of the pitfalls to avoid and common usage patterns to emulate.
- **Key Investigators:** Oliver Rübel, Andrew Tritt, Ben Dichter, Ryan Ly, Thomas Braun, Tom Davidson, Nicholas Cain, Lydia Ng
  - **Key Institutions:** LBNL, byte physics, UCSF, Allen Institute

# Code Activity (May 2019)

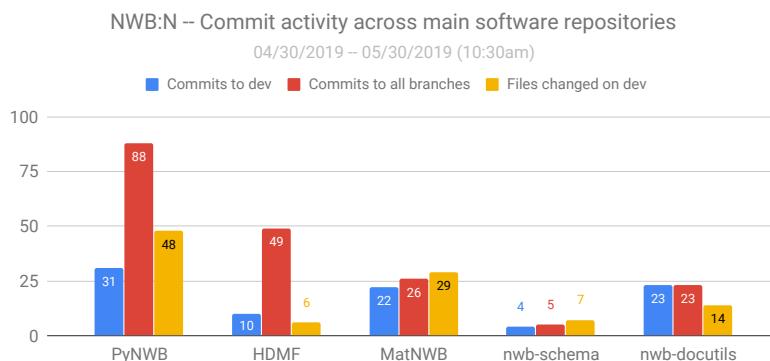
The hackathon has helped to create a lot of energy around the NWB:N main repositories. Here we report GitHub code statistics for the main NWB:N repositories<sup>1</sup> for the period of 04/30/2019 – 05/30/2019 (10:30am), with the User Days (May 13-14) and Developer Days (May 15-16) being roughly in the middle of that period.

## NWB:N Software Repositories

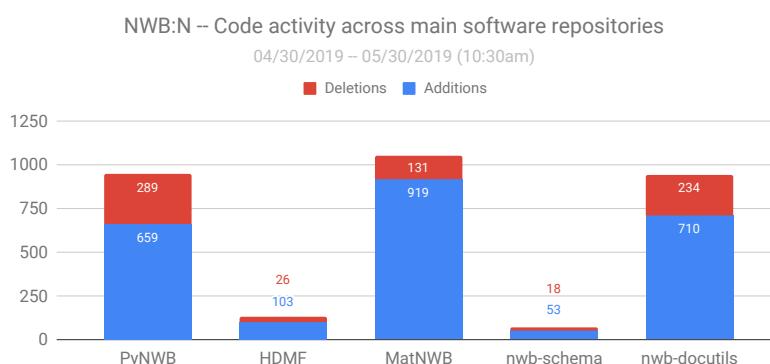
The charts below summarize activity on the main NWB:N code repositories for 1) [PyNWB](#), 2) [HDMF](#), 3) [MatNWB](#), 4) [NWB:N Schema](#), and the 5) [Documentation Utilities](#). Across these NWB:N codes we have seen significant progress on features and identification and resolution of issues, with **61** merged and **18** proposed pull requests and **35** closed and **46** new issues in the 30 day period reported here.



Pull request (PR) and issue activity across the main NWB:N software repositories in May 2019. In total there were **61** merged PR, **18** proposed PR, **35** closed issues, and **46** new issues.



Number of commits and files changed across the main NWB:N software repositories in May 2019. In total there were **191** commits and **104** files changed.



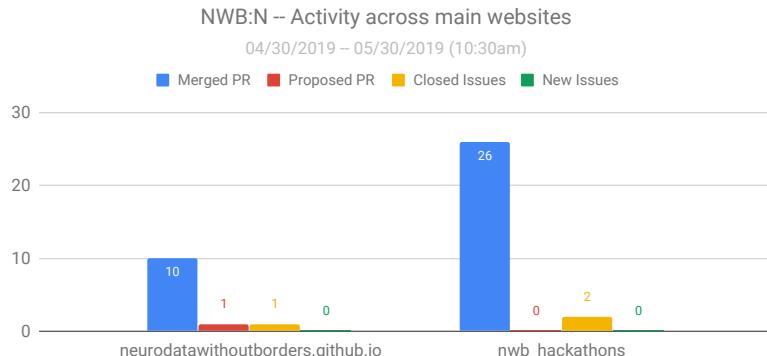
Code activity across the main NWB:N software repositories in May 2019. In total there were **2,444** additions and **698** deletions.

18

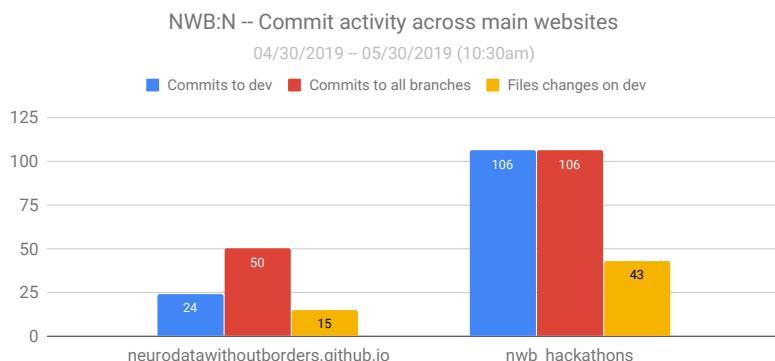
<sup>1</sup>**Note:** The statistics reported here focus on the main NWB:N repositories only. A majority of hackathon projects involved external data analysis, visualization, storage, and management tools and custom user codes (e.g., for data conversion) as well as new NWB:N code repositories, e.g., for the 1) [PyNWB interactive course](#), 2) [ontology project](#), 3) NWB:N jupyter widgets, and the 4) [NWB:N extension catalog](#), which are not reflected here.

## NWB:N Website

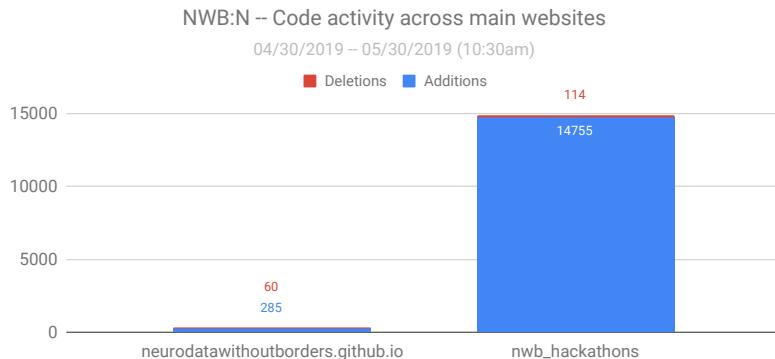
The charts below summarize activity on the main [neurodatawithoutborders.github.io](https://neurodatawithoutborders.github.io) as well as the [hackathon](#) event page. As a result of the User Days and Developer Days we have seen significant improvements to the main website, in particular the addition of 1) [Best Practices](#), 2) [Example NWB:N Data Files](#), and 3) [Analysis and Vis Tools](#) pages as well as 4) 30 different [hackathon project](#) pages.



Pull request (PR) and issue activity on [neurodatawithoutborders.github.io](https://neurodatawithoutborders.github.io) in May 2019. In total there were **36** merged PR, **1** proposed PR, **3** closed issues, and **0** new issues.



Number of commits and files changed on [neurodatawithoutborders.github.io](https://neurodatawithoutborders.github.io) in May 2019. In total there were **156** commits and **58** files changed.



Code activity on [neurodatawithoutborders.github.io](https://neurodatawithoutborders.github.io) in May 2019. In total there were **15,040** additions and **174** deletions.

## Event Photos



Figure 4: NWB:N User Days, Day 1 (Synapse Room). (Photos by Oliver Rübel, LBNL)



Figure 5: NWB:N User Days, Day 1-2 (Spectrum Room). (*Photos by Oliver Rübel, LBNL*)



Figure 6: NWB:N User Days, Day 2 (Synapse Room). (*Photos by Oliver Rübel, LBNL*)



Figure 7: NWB:N Developer Days, Day 3 and 4. (Photos by Oliver Rübel, LBNL)

## Acknowledgements

We would like to thank Andrew Tritt, Nathan Clack, Lawrence Niu, Michael Grauer, Ryan Ly, Thomas Braun, Tom Davidson, Thinh Nguyen, and Vyassa Baratham, Benjamin Dichter, and Oliver Rübel for their contributions as presenters and facilitators at the event. Special thanks also to Stephanie Albin (Kavli Foundation) for travel support, Janine Stephens (HHMI Janelia) for onsite logistic support, and Matt Staley (HHMI Janelia) for AV support and photography. We would like to thank all the participants for the great enthusiasm and commitment and for making the hackathon a great success! We thank the Kavli foundation and HHMI Janelia for financial support of the hackathon.

## 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.