




Neurosift: DANDI exploration and NWB visualization in the browser

Jeremy Magland ¹¶, Jeff Soules ¹, Cody Baker ², and Benjamin Dichter ²

1 Center for Computational Mathematics, Flatiron Institute, United States of America 2 CatalystNeuro, United States of America ¶ Corresponding author

DOI: [10.21105/joss.06590](https://doi.org/10.21105/joss.06590)

Software

- [Review](#) 
- [Repository](#) 
- [Archive](#) 

Editor: [Marcel Stimberg](#)  

Reviewers:

- [@alexrockhill](#)
- [@jodeleeuw](#)
- [@sanjayankur31](#)

Submitted: 15 March 2024

Published: 27 May 2024

License

Authors of papers retain copyright and release the work under a Creative Commons Attribution 4.0 International License ([CC BY 4.0](https://creativecommons.org/licenses/by/4.0/)).

Summary

Neurosift, a browser-based visualization tool, is designed for the interactive exploration of Neurodata Without Borders (NWB) files, whether stored locally, on remote servers, or within the Distributed Archives for Neurophysiology Data Integration (DANDI). NWB ([Rübel et al., 2022](#); [Teeters et al., 2015](#)) is an open data standard for neurophysiology that enables the sharing, archiving, and analysis of various types of neurophysiology data. DANDI ([Rübel et al., 2022](#)) is a cloud-based platform that supports the storage, sharing, and analysis of neurophysiology data including NWB files. With Neurosift integration, users browsing DANDI can easily open any NWB file in the browser and explore its contents, including timeseries data, images, and more. Neurosift can also be used to browse the DANDI database or individual Dandisets. Overall, Neurosift simplifies the visualization and exploration of complex NWB file structures, making it a valuable tool for neuroscientists.

Statement of need

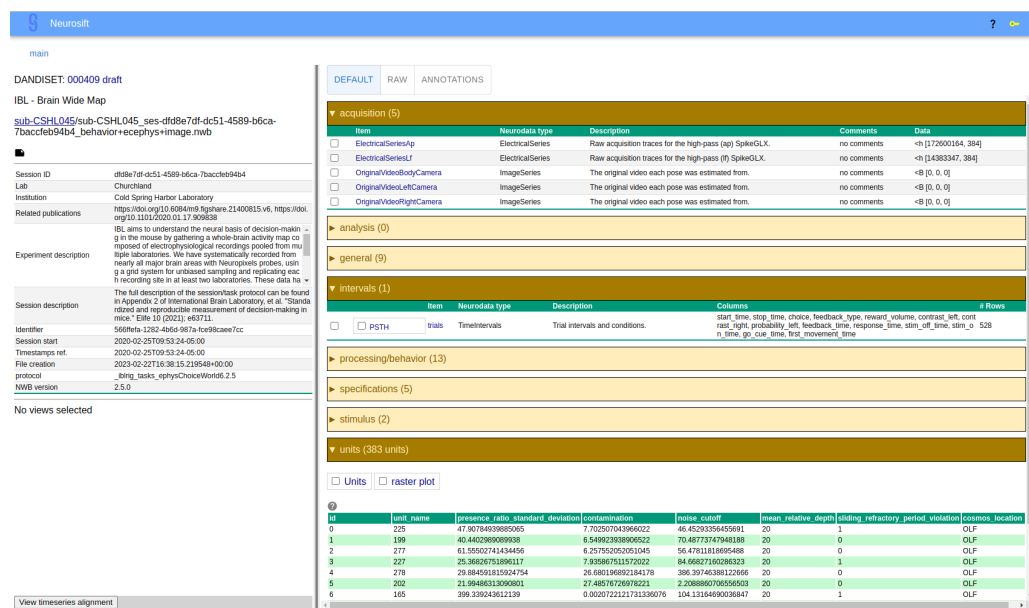
In the evolving field of neuroscience research, the ability to manage and share complex data sets is crucial. NWB has emerged as a standard for neurophysiology data, aimed at facilitating data sharing, storage, and analysis. However, the specialized nature of the NWB format necessitates tools that can provide intuitive interfaces for researchers to explore their data effectively. Neurosift is designed to address this need.

Because files found on DANDI can often be large and unwieldy, various tools have emerged to address this issue by streaming portions of the NWB file without the need to download the entire file. One such tool is NWB Widgets ([Dichter, 2022](#)), which provides a suite of interactive widgets for visualizing NWB data within Jupyter notebooks, enabling users to navigate the hierarchical structure of NWB files and directly visualize specific data elements. This package was a large part of the inspiration for Neurosift. The main difference is that NWB Widgets is a Python package that runs within interactive Python environments, while Neurosift is a browser-based tool that can be used without any installation. These two tools cater to different use cases, with Neurosift being more accessible to a wider audience, and being better suited for integration with DANDI.

Functionality

Neurodata Without Borders files are structured hierarchically and encapsulate various “neurodata” types that reflect different aspects of neurophysiological experiments. These types range from *BehavioralEvents*, which record discrete actions or occurrences within experiments, to data structures like *Fluorescence*, *ImageSegmentation*, and *RoiResponseSeries*, which are

key data types in optical neurophysiology. Other neurodata types include *ElectricalSeries* for electrophysiological signals and *Units* for spike times of neurons. Neurosift allows interactive navigation of this hierarchical structure (Figure 1) and provides plugin visualizations for many of these types (Figure 2). It also facilitates the creation of composite views by allowing users to select and synchronize multiple data types within the same interface (Figure 3). This synchronization extends to navigation actions such as zooming and panning, where different sub-windows, each displaying a different aspect of the data, maintain a shared time axis. These views can then be shared with others as a URL.



acquisition (5)

| Item | Neurodata type | Description | Comments | Data |
|---|------------------|---|-------------|---------------------|
| <input type="checkbox"/> ElectricalSeriesAp | ElectricalSeries | Raw acquisition traces for the high-pass (ap) SpikeGLX. | no comments | <h [172500164, 384] |
| <input type="checkbox"/> ElectricalSeriesL7 | ElectricalSeries | Raw acquisition traces for the high-pass (l7) SpikeGLX. | no comments | <h [14383347, 384] |
| <input type="checkbox"/> OriginalVideoBodyCamera | ImageSeries | The original video each pose was estimated from. | no comments | <B [0, 0, 0] |
| <input type="checkbox"/> OriginalVideoLeftCamera | ImageSeries | The original video each pose was estimated from. | no comments | <B [0, 0, 0] |
| <input type="checkbox"/> OriginalVideoRightCamera | ImageSeries | The original video each pose was estimated from. | no comments | <B [0, 0, 0] |

general (9)

intervals (1)

| Item | Neurodata type | Description | Columns | # Rows |
|-------------------------------|----------------|---------------------------------|---|--------|
| <input type="checkbox"/> PSTH | Intervals | Trial intervals and conditions. | start_time, stop_time, choice, feedback_type, reward_volume, contrast_left, contrast_right, probability_left, feedback_time, response_time, start_off_time, start_on_time, go_cue_time, first_movement_time | 528 |

processing/behavior (13)

specifications (5)

stimulus (2)

units (383 units)

Units raster plot

| ID | unit_name | presence_ratio | standard_deviation | contamination | noise_cutoff | main_relative_depth | sliding_refractory_period | violation | cosmos_location |
|----|-----------|--------------------|----------------------|--------------------|--------------|---------------------|---------------------------|-----------|-----------------|
| 0 | 225 | 47.90784939885065 | 7.702507043966022 | 46.45293356455961 | 20 | 1 | 0 | 0 | OLF |
| 1 | 199 | 40.4402999099395 | 6.549922938905522 | 70.48173747948188 | 20 | 0 | 0 | 0 | OLF |
| 2 | 277 | 61.55502741434456 | 6.257520202010465 | 56.4701311895468 | 20 | 0 | 0 | 0 | OLF |
| 3 | 227 | 25.36626751896117 | 7.935867511572022 | 84.66827160286323 | 20 | 1 | 0 | 0 | OLF |
| 4 | 278 | 29.884591815924754 | 26.680196892184178 | 386.39746388122668 | 20 | 0 | 0 | 0 | OLF |
| 5 | 302 | 23.894865313090961 | 27.486767295919221 | 2.2388969178650603 | 20 | 0 | 0 | 0 | OLF |
| 6 | 165 | 399.339243612139 | 0.002072212171336076 | 104.13164690036847 | 20 | 1 | 0 | 0 | OLF |

Figure 1: Screenshot of Neurosift displaying a NWB file from the 'IBL - Brain Wide Map' Dandiset (ID 000409), showcasing the tool's capability to navigate and visualize the hierarchical structure of neurophysiological data. Visible are expanded panels for *ElectricalSeries* and *ImageSeries* objects, alongside a *TimeIntervals* object and a *Units* table.

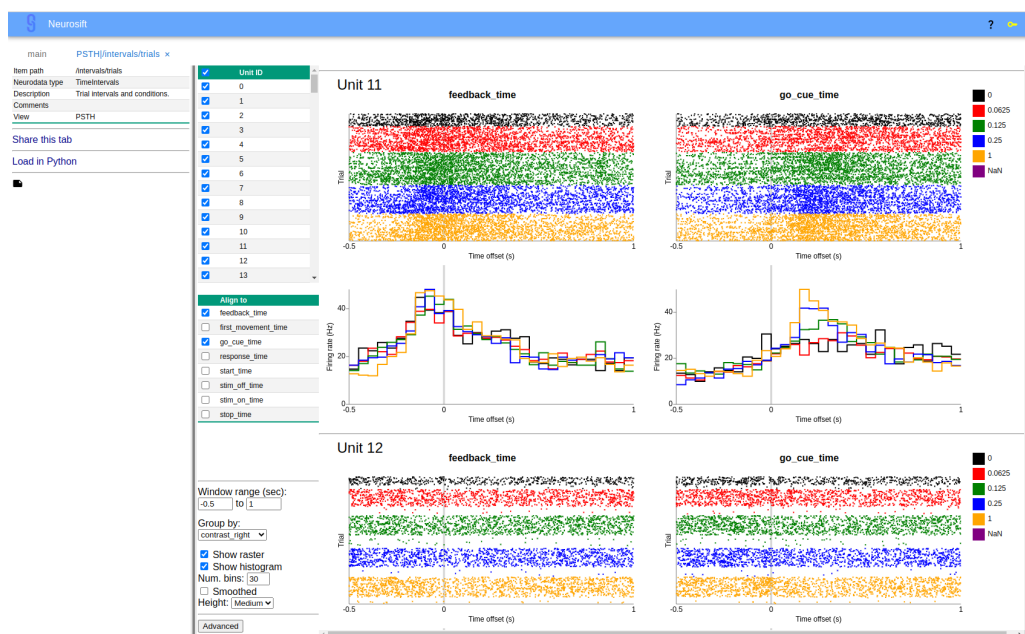


Figure 2: Interactive Peri-stimulus Time Histogram (PSTH) visualization in Neurosift, enabling users to select neural units, time alignment variables, and customize options like window range, histogram bin count, and grouping variables.

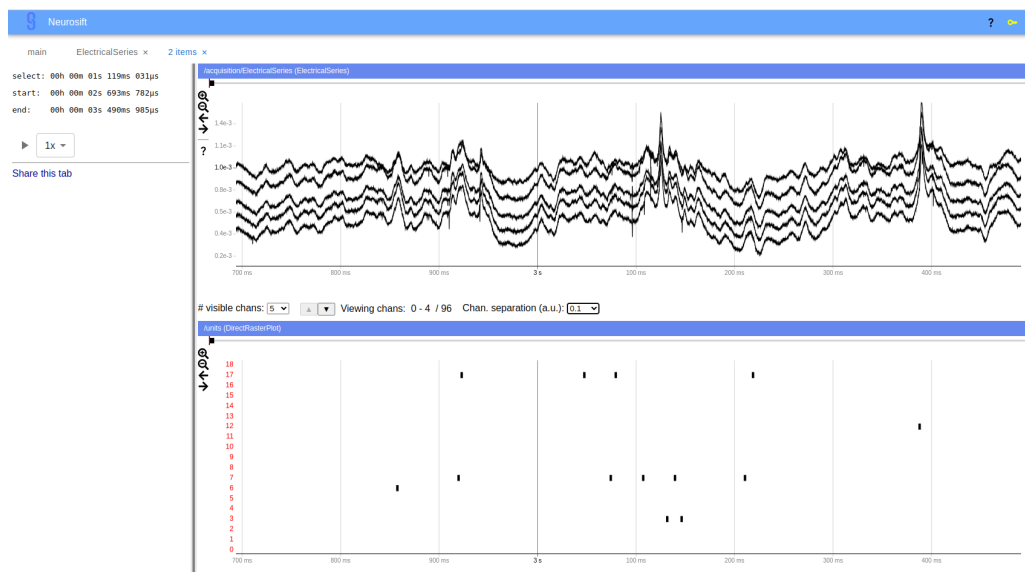


Figure 3: Synchronized view within Neurosift demonstrating the interactive alignment between an *ElectricalSeries* and a *Spike Raster Plot*. This feature allows users to seamlessly zoom and pan across both visualizations, maintaining a coherent temporal perspective across different data types.

Architecture and technical innovation

Neurosift is a *static* React/TypeScript website, meaning that it is delivered to the user's browser exactly as stored, without the need for dynamic server-side processing of requests. This approach simplifies deployment and maintenance; it can be deployed to any static hosting service.

The main technical challenge in developing Neurosift was the requirement to lazy-load data objects from remote NWB files that are built on the complex HDF5 format. While HDF5's efficient data organization is ideal for the large, multidimensional datasets typical in neurophysiology, its primary implementations are in the C language. This necessitates a creative solution to enable efficient web-based access to these files. To bridge this gap, Neurosift leverages WebAssembly to run compiled C code in the browser, specifically utilizing a modified version of the h5wasm (Maranville, 2022) library. Unlike the unmodified h5wasm, which primarily handles fully downloaded files, Neurosift's fork introduces an innovative approach to efficiently read data chunks from remote files. This allows for partial data reads without the need for a prior download of the entire file. This solution showcases the potential of WebAssembly in overcoming challenges associated with web-based data analysis tools.

Future directions

Looking forward, there is potential to expand Neurosift's capabilities with enhanced visualizations and support for additional data types. Additionally, we are adding support for Zarr (Miles, 2024), a cloud-friendly alternative to HDF5 as a storage backend for NWB files.

Conclusion

Neurosift makes neurophysiology data more accessible to scientists. By facilitating the exploration of complex datasets directly within a browser, it lowers the barrier to entry for data analysis and fosters collaborative research efforts.

References

- Dichter, B. K. (2022). Nwbwidgets: Explore the hierarchical structure of NWB 2.0 files and visualize data with jupyter widgets. In *GitHub repository*. GitHub. <https://github.com/NeurodataWithoutBorders/nwbwidgets>
- Maranville, B. B. (2022). h5wasm: A WebAssembly HDF5 reader/writer library. In *GitHub repository*. GitHub. <https://github.com/usnistgov/h5wasm>
- Miles, A. et al. (2024). Zarr-developers/zarr-python: v2.17.1. In *Zenodo*. Zenodo. <https://doi.org/10.5281/zenodo.10790679>
- Rübel, O., Tritt, A., Ly, R., Dichter, B. K., Ghosh, S., Niu, L., Baker, P., Soltesz, I., Ng, L., Svoboda, K., & others. (2022). The Neurodata Without Borders ecosystem for neurophysiological data science. *eLife*, *11*, e78362. <https://doi.org/10.7554/eLife.78362>
- Teeters, J. L., Godfrey, K., Young, R., Dang, C., Friedsam, C., Wark, B., Asari, H., Peron, S., Li, N., Peyrache, A., & others. (2015). Neurodata Without Borders: Creating a common data format for neurophysiology. *Neuron*, *88*(4), 629–634. <https://doi.org/10.1016/j.neuron.2015.10.025>