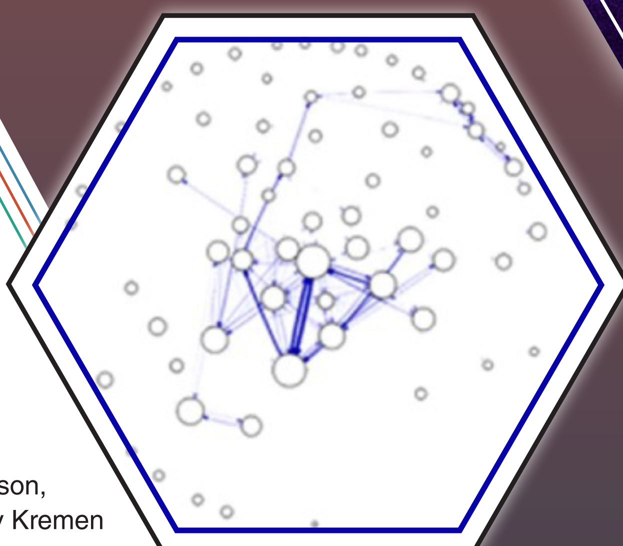
NOVEMBER 2022

Toolmakers Newsletter

Special Software Issue



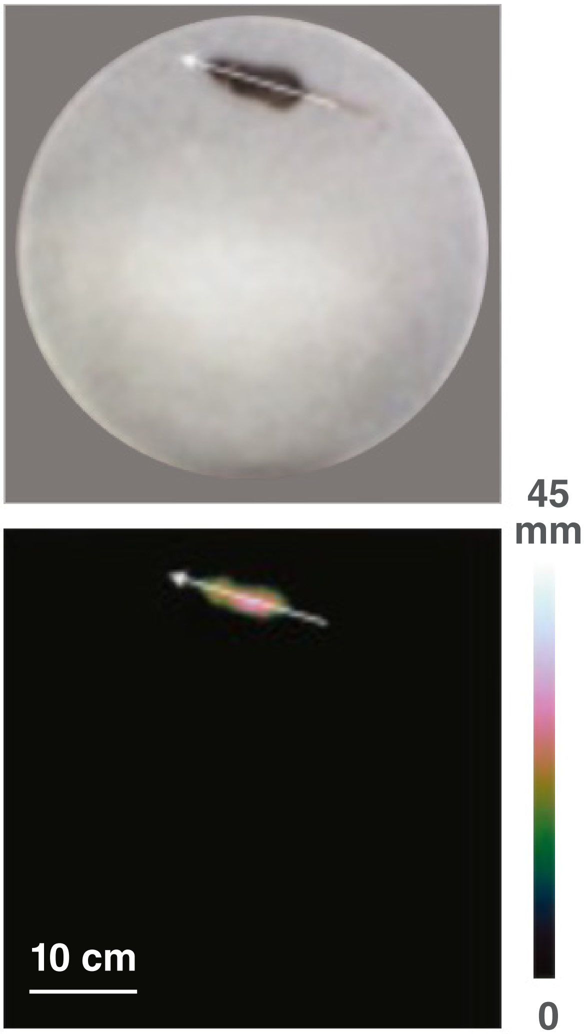
ISSUE 08

In this issue

BRAIN INITIATIVE ALLIANCE

www.braininitiative.org

Welcome Back to the Toolmakers Newsletter!



Welcome to the fourth and final *Brain Research Through Advancing Innovative Neuro-technologies®* (BRAIN) Initiative Alliance Toolmakers Newsletter of 2022!

In this special issue focusing on software tools, we are excited to highlight new installments and features of four types of neuroscience software: OpenMind Software Tools by Dr. Philip Starr, Dr. David Borton, Dr. Heather Dawes, Dr. Tim Denison, Dr. Jeffrey Herron, and Dr. Vaclav Kremen; DataJoint Elements by Dr. Dimitri Yatsenko; MoSeq by Dr. Bob Datta; and OpenNeuro by Dr. Russell Poldrack. Let's learn more details about these projects and some insights from the investigators!

Image: A mouse imaged with MoSeq in the circular open field with a standard RBG (top) and 3D depth camera (bottom) mm=mm above floor. The arrow indicates the inferred axis of the animal's spine. Credit: Wiltschko et al., 2015, *Neuron*.

On the front cover: Top Right Hexagon: An image still from a video of the *C. elegans* brain, including every nerve and muscle fiber, being reconstructed by serial-section electron microscopy. Credit: Daniel Witvliet, University of Toronto and Harvard University, 2020. Top Central Hexagon: Four-week-old rat cortical neurons labeled for dendrites (red), axons (green), and nuclei (blue). Credit: Karthik Krishnamurthy, Davide Trotti, Piera Pasinelli, Thomas Jefferson University, 2020. **Bottom Right Hexagon:** A pseudo-colored image of high-resolution gradient-echo MRI scan of a fixed cerebral hemisphere from a person with multiple sclerosis. Credit: Govind Bhagavatheeshwaran, Daniel Reich, National Institute of Neurological Disorders and Stroke, National Institutes of Health, 2016. **Bottom Central Hexagon:** Behavioral state maps generated by MoSeq depicting behavioral syllables (nodes, diameter is proportional to syllable usage) and transitions (edges, thickness is proportional to transition likelihood) encapsulate all behavior expressed within a given experiment captured by the camera and can be used to predict future behavior from present actions. Credit: Datta, S.R., 2019, *BMC Biol*.

OpenMind Software Tools – Dr. Philip Starr, Dr. David Borton, Dr. Heather Dawes, Dr. Tim Denison, Dr. Jeffrey Herron, and Dr. Vaclav Kremen

At the University of California, San Francisco and alongside an international collaboration of universities, Dr. Philip Starr and colleagues are making documented software tools and other resources available through the OpenMind Project, which is focused on the needs of clinical researchers in the field of advanced deep-brain stimulation. The distribution of OpenMind Software Tools, such as the "OMNI" platform shown in the image, is one of the activities overseen by OpenMind, which helps make research resources for brain-interfacing clinical research widely available. Software tools and other resources are sourced from across the OpenMind Consortium, which includes a community of over 40 academic research labs—along with industry, funding, and regulatory partners—devoted to developing new technologies to treat movement disorders such as Parkinson's disease, epilepsy, and neuropsychiatric conditions like major depression or chronic pain.

The software code that makes up the OpenMind Toolkit includes code for data collection, data visualization and analysis, deep-brain stimulation (DBS) device programming, and DBS device control using application program interfaces (APIs). For example, the SCBS-PatientFacingApp is a piece of patient-facing software that serves as a data collection tool to make clinical trials more manageable. The app makes it possible to collect data on patient neural recordings at home, allowing the patients to take an active approach in interacting with their implanted devices. Other popular pieces of software such as the Analysis-rcs-data tool

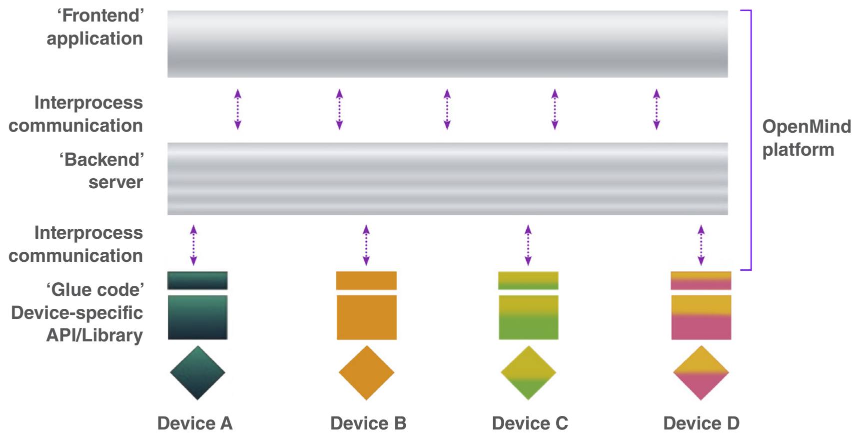


Image: OpenMind microservice architecture is a software platform that is agnostic to the neurotechnology hardware. Credit: Borton et al., 2020, *Neuron*.

extract raw data from a DBS device and package it for data analysis, facilitating data management and analysis processes.

In 2020, the first paper on OpenMind was published in *Neuron*, outlining how the academic consortium could promote community-supported neurotechnology platforms. Since then, the platform has grown to include a broad range of activities including software-based tools, know-how technical problem solving, and platform solutions. According to Dr. Heather Dawes, OpenMind's shared collection of knowledge can serve as a model to help other fields overcome technical, scientific, or operational challenges.

In the future, the team hopes that OpenMind's software capabilities will continue helping scientists overcome technical challenges. OpenMind's goal is to provide user-friendly tools and solutions for neuroscience devices used in clinical studies to cultivate a culture of shared purpose and problem-solving. In turn, this means that patients will get the treatments and devices they need, with faster timelines and lower costs.

> "There's often more to be gained by accessing knowledge, tools, and technical solutions from others—in a consortium that makes it easy to share out to others in return—than remaining in the classic siloed position, fated to invent wheels while others already have carts."
>
> *— Dr. Heather Dawes, on the advantages of OpenMind*

DataJoint Elements – Dr. Dimitri Yatsenko

DataJoint is a framework invented by Dr. Dimitri Yatsenko that creates and manages modular scientific workflows for shared data pipelines. The framework supports open science initiatives and includes graphical interfaces for data entry, ingestion of raw data, automated analysis, interfaces for data visualization, and it can query data across vast datasets.

Image: DataJoint Elements logo. Credit: DataJoint Elements, 2022.

> "We coordinate with other major BRAIN Initiative projects such as data standards, catalogs, and archives to make data and analysis transparent and reproducible."
>
> *— Dr. Dimitri Yatsenko*

Within DataJoint exists DataJoint Elements, an NIH-funded project that began in 2020 to develop open-source workflows for neurophysiology studies that organize data and automate computations from acquisition to visualization. DataJoint Elements helps neuroscience projects and tools work together easily and sustainably. The project aims to increase the scale, speed, and reproducibility of neuroscience research by disseminating automated workflows that organize data and execute the latest analysis tools. It contains curated modules that assist researchers in workflow assembly across neuroscience modalities. It combines manual steps, data management, and automated analysis to help neuroscientists collect, prepare, process, analyze, and model data.

DataJoint Elements continues to grow—55 labs use the DataJoint open-source framework for their data analyses, and the platform has been leveraged for 65 publications. There are a few new developments on the horizon for DataJoint Elements as well. The DataJoint team is currently adding new modules for cross-registration of 3D microscopy volumes, quality control metrics for neurophysiology recordings, and optogenetic stimulation. The team was also recently awarded additional funding, which is being used to develop a new cloud-based process platform for configuring automated workflows in neuroscience.

Researchers interested in using DataJoint Elements should start by exploring the sample workflows in the online Jupyter CodeBook to view common experiment types using the sample data. Next, teams can determine how to host the shared database and how to configure the computational pipeline and its interface, as described in the online documentation. After that, teams can schedule an Office Hour with the DataJoint team to review optimal configurations. Dr. Yatsenko also recommends that researchers sign up for the DataJoint Community Slack Channel to connect with other users.

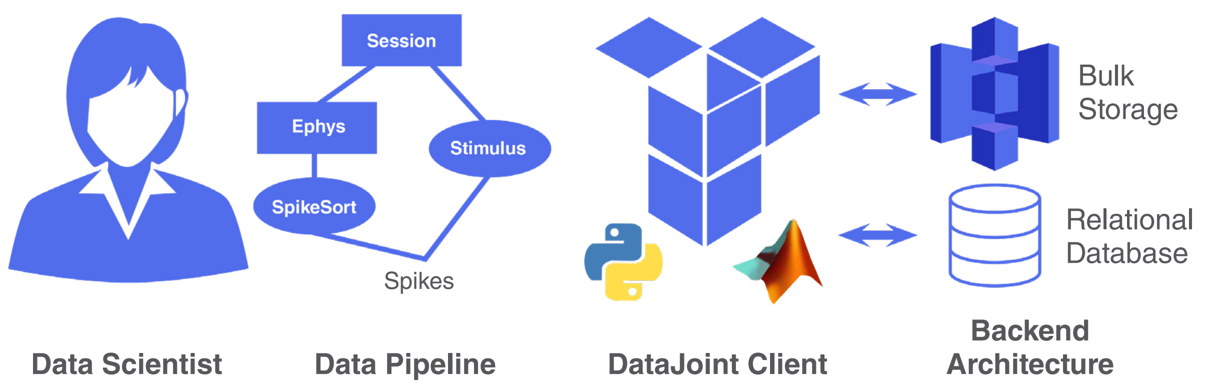


Image: DataJoint Elements capabilities breakdown. Credit: Dimitri Yatsenko, DataJoint Elements, 2022.

MoSeq –Dr. Bob Datta

At Harvard Medical School, Dr. Bob Datta and his lab developed Motion Sequencing, or MoSeq, in 2015 as a behavioral quantification technique to measure and understand how mice compose their behavior when exploring simple arenas. MoSeq works by processing 3D video of behaving mice to uncover the underlying structure behind mouse behavior. It uses quantified mapping and time-series statistical modeling techniques to translate the video into statistics and visualizations. These models help to identify the optimal set of behavioral symbols demonstrated during an experiment.

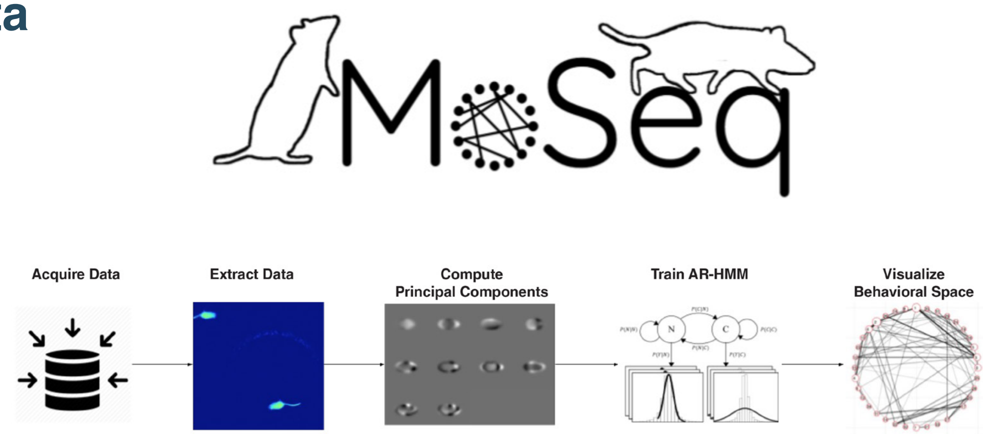


Image: MoSeq is an interactive pipeline and toolkit for unsupervised discovery of repeated motifs in 3D video of freely behaving mice. Credit: Bob Datta, MoSeq, 2022.

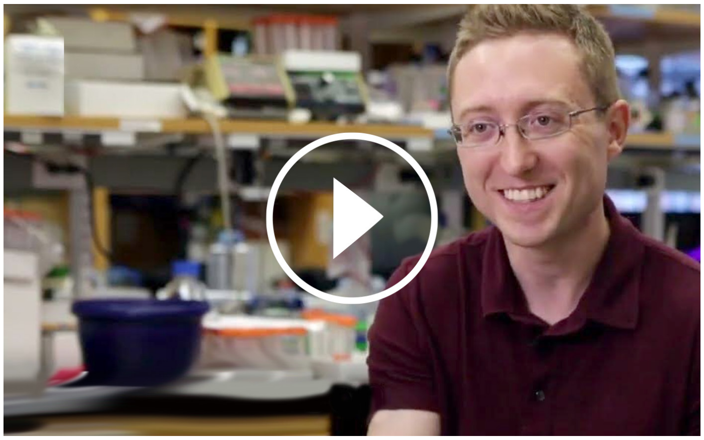
One of MoSeq's greatest advantages is reducing the amount of human labor required to explore mouse behavior. Whereas other behavioral classifiers may require manual labeling or motif specification, MoSeq requires no labeling or training data and uses machine learning (via probabilistic graphical modeling and clustering) to identify repeated motifs of behavior, such as pausing or a head bob.

In addition to reducing human labor, MoSeq also offers a high-resolution lens into behavior, which can facilitate the identification of genetic mutations and brain abnormalities in mice. "We and others have used it successfully to identify the impact of pharmacogenetic, optogenetic, genetic, pharmacological, and sensory manipulations on mouse behavior," says Dr. Datta. This technology carries the capacity to uncover hidden phenotypes or identify side effects or predictors of disease through behavioral patterns.

With MoSeq, Dr. Datta and his team can induce activity in various brain regions to map them and study the behavioral space. With the help of MoSeq's ability to give time stamps for behavior, Dr. Datta's lab is studying the role of dopamine in spontaneous behavior by linking time stamps of brain activity patterns to behavior.

The team at Harvard hopes to continuously evolve the MoSeq framework—they created a publicly available version of MoSeq and are working to incorporate new functionalities and different types of data into that platform. These capabilities include compatibility with point-tracking data and enabling MoSeq to work with new types of animals and more complex laboratory environments.

> "One of the main challenges in understanding spontaneous behavior is that it is hard to discern when one behavior stops and another begins. MoSeq gives you time stamps for behavior."
>
> *— Dr. Bob Datta*



Video: Researchers at the Datta Lab at Harvard Medical School explain what MoSeq is and how it works. Credit: MoSeq, 2022.

OpenNeuro – Dr. Russell Poldrack

OpenNeuro is a free and open platform created by Dr. Russell Poldrack and his lab at Stanford University that stores, shares, and analyzes neuroimaging data. The platform only accepts datasets that comply with Brain Imaging Data Structure (BIDS), a set of standards developed within the brain imaging community for organizing data and metadata for reuse. The goal is to make raw imaging datasets easy to share, reuse, and analyze.

Most of the datasets in the OpenNeuro repository are human data, but the platform also supports non-human data and a growing number of neuroscience data types. The most common data in the platform are magnetic resonance imaging (MRI) data. The platform also has positron emission tomography (PET), magnetoencephalography (MEG), electroencephalogram (EEG), and intracranial electroencephalography (iEEG) data.

The process to upload data onto OpenNeuro is robust, quick, and includes automatic data validation prior to upload, courtesy of the automated BIDS validator. Scientists can upload datasets of any size and generate a citable digital object identifier (DOI) to use in manuscripts. DOIs are generated for each version of every dataset that gets uploaded.

The upload process is just one of the many advantages that OpenNeuro offers researchers. There are no data use agreement requirements and there is minimally restrictive licensing for the datasets that get uploaded to OpenNeuro. This allows the platform to support multiple download options and researchers can obtain data in a variety of ways for large-scale reuse.

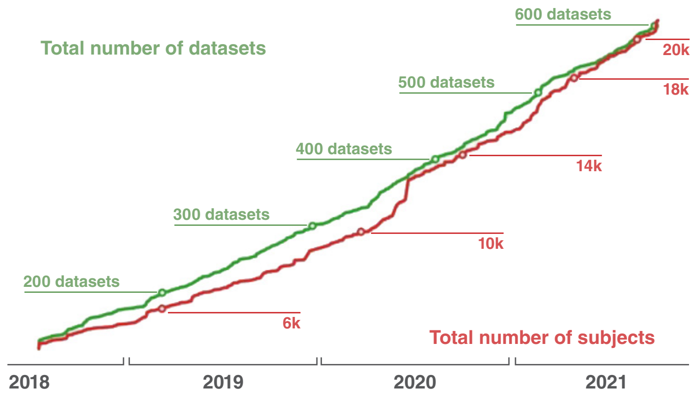


Image: A graph demonstraing OpenNeuro's growth in volume of data available since its opening started operations in 2017. Credit: Markiewics et al., 2021, *eLife*.

Image: OpenNeuro logo. Credit: OpenNeuro, 2022.

OpenNeuro allows researchers to search through datasets provided by other users, and the database is in the process of generating and sharing quality control results and preprocessed data for many of the datasets. The platform has grown exponentially over the last few years, with 750 datasets currently available and 28,544 participants.

Dr. Poldrack encourages researchers interested in using OpenNeuro to read his team's latest publication in *eLife* that outlines the archive. If you are a researcher wishing to share your data with OpenNeuro, you must first convert your data into the BIDS format. Datasets that cannot be shared openly are not currently supported. For more information on getting started with conversion, please view the BIDS Starter Kit.

> "Visits to the website have grown consistently over the last few years, now at 8-10K users/month."
>
> — *Dr. Russell Poldrack*

Excited by the potential of the tools in this issue?! Stay tuned for our next issue and explore more products of BRAIN Initiative discoveries in our Toolmakers' Resources page!