# NEMAR: A BRAIN INITIATIVE SCIENCE GATEWAY TO NEUROELECTROMAGNETIC DATA, TOOLS, AND COMPUTATION

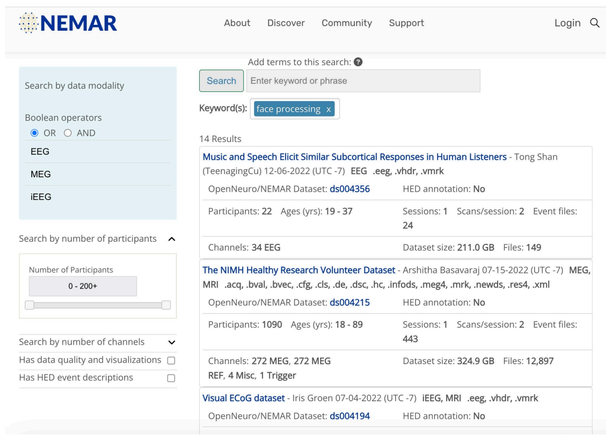

Youn, C.[3], Delorme, A.[1,2], Truong, D.[1], Salazar, A [1]. Sivagnanam, S.[3], Stirm, C.[3], Yoshimoto, K.[3], Poldrack, R.A.[4], Majumdar, A.[3], Makeig, S.[1]

[1]Swartz Center of Computational Neuroscience, INC, UCSD, La Jolla, CA, USA
[2]CERCO, CNRS, Paul Sabatier University, Toulouse, France
[3] San Diego Supercomputer Center, UCSD, La Jolla, CA, USA
[4] Stanford University, Stanford, CA, USA


**MOTIVATION** To take advantage of recent and ongoing advances in intensive and large-scale computational methods, and to preserve the scientific data created by publicly funded research projects, public data archives must be created as well as standards for specifying, identifying, and annotating the deposited data. The value of and interest in such archives among researchers can be greatly increased by connecting them to active computational capabilities and a framework of analysis and search tools that support further analysis as well as larger scale meta-analysis and data mining.

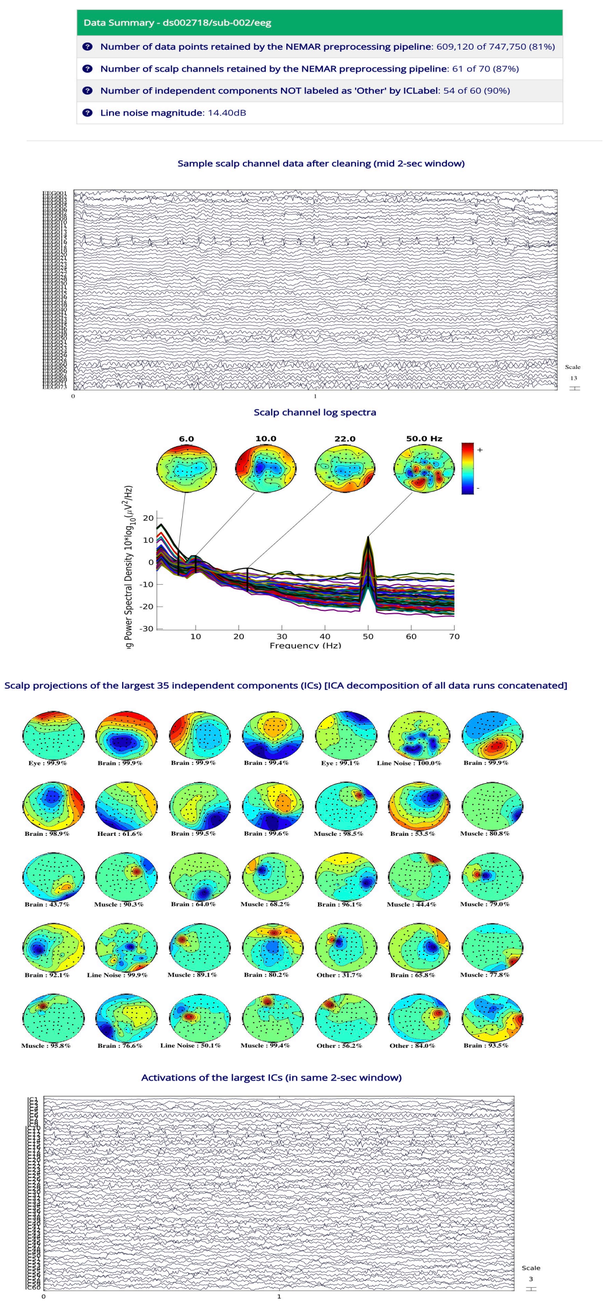
NEMAR, the NeuroElectroMagnetic Archive for Research, serves as a gateway to BIDS-formatted NEM (EEG, MEG, iEEG) imaging data uploaded and shared on OpenNeuro.org. NEMAR offers a range of functions, including data search and visualization, and operates as a freely accessible integrated resource known as DATCOR (Data, Tools, and Compute Resource) dedicated to human NeuroElectroMagnetic brain imaging data. The construction of NEMAR leverages the HubZero framework, an open-source software platform renowned for empowering gateway development, data publication, resource sharing, collaboration, and community-building within a unified web-based ecosystem.

## Search and visualize available EEG, MEG, and iEEG (NEM) data




Search interface with overview of dataset information and filtering options

View and download dataset




Data summary and visualization

## An integrated data, tools and compute resource

Identified NEMAR datasets can be processed using high-performance computing (HPC) resources via the Neuroscience Gateway (www.*nsgportal.org*), a freely available, easy-to-access portal to HPC resources that serves several computational environments and packages (MATLAB, Python, R, EEGLAB, Freesurfer, etc.). Users can submit custom analysis scripts operating on NEMAR data with no need for inefficient data downloads and reuploads.

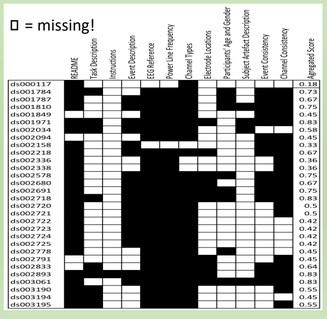
Sample MATLAB script

```
eeglab; % start EEGLAB and add paths
[~,EEG] = pop_importbids('/expanse/projects/nemar/openneuro/ds002691');

% duration and number of events for EEG dataset
duration = round([EEG.pnts]'/EEG(1).srate);
nevents  = cellfun(@length, {EEG.event}');

% display some statistics about the RAW EEG data
fid = fopen('results.txt', 'w'); % open result file
fprintf(fid,'BIDS dataset %s\n', bidsName);
fprintf(fid,'%d datasets\n', length(EEG));
fprintf(fid,'%d to %d channels\n', min([EEG.nbchan]), max([EEG.nbchan]));
fprintf(fid,'%d to %d seconds\n', min(duration), max(duration));
fprintf(fid,'%d to %d events\n', min(nevents), max(nevents));
fclose(fid); % close result file
```

Sample Python script

```
import os
import json
import pprint
import shutil

bidsName = 'ds002691'                    # Specify the NEMAR dataset control ID
DATASETDIR = os.environ['NEMARPATH'] + f'{bidsName}'
copiedDir = '/BIDScopied'
shutil.copytree(DATASETDIR, copiedDir)          # Copy dataset to current working directory before processing
with open(os.path.join(copiedDir, 'dataset_description.json')) as f:
    dataset_description = json.load(f.read()) # Load the dataset description from json file
    pprint.pprint(dataset_description)        # Pretty-print the json file contents
subprocess.run(["chmod", "-R", "u+w", copiedDir]) # Make sure permissions are correct for successful deletion.
shutil.rmtree(copiedDir)                        # Remove the copied dataset files to avoid a large download.
```

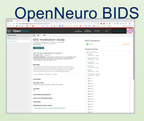
**NEMAR Dataset** metadata



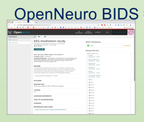

Current **NEMAR Dataset** statistics

| Dataset | n | Flat | Chan. | Good chan. | Good data | Brain ICs |
|---|---|---|---|---|---|---|
| ds000117 | 84 | - | 74 | 89 - 90 | 68 - 78 | 28 - 32 |
| ds001784 | 30 | - | 60 | 85 - 92 | 63 - 77 | 5 - 11 |
| ds001787 | 40 | - | 64 | 93 - 95 | 85 - 90 | 24 - 30 |
| ds001810 | 263 | - | 64 | 78 - 80 | 71 - 74 | 35 - 37 |
| ds001849 | 120 | - | 30 | 86 - 88 | 66 - 70 | 41 - 45 |
| ds001971 | 273 | - | 112* | 91 - 95 | 81 - 84 | 8 - 10 |
| ds002034 | 167 | 21 | 62 | 95 - 98 | 47 - 52 | 29 - 33 |
| ds002094 | 43 | - | 30 | 88 - 93 | 66 - 77 | 37 - 45 |
| ds002158 | 8 | - | 63 | 80 - 86 | 19 - 66 | 11 - 23 |
| ds002218 | 18 | - | 32 | 94 - 97 | 66 - 78 | 35 - 46 |
| ds002336 | 54 | 1 | 63 | 81 - 83 | 74 - 89 | 33 - 39 |
| ds002338 | 85 | - | 63 | 74 - 77 | 87 - 95 | 24 - 28 |
| ds002578 | 2 | - | 18 | 78 - 89 | 96 - 97 | 46 - 60 |
| ds002680 | 350 | - | 19 | 92 - 93 | 81 - 84 | 53 - 56 |
| ds002691 | 20 | - | 32 | 93 - 97 | 82 - 87 | 30 - 40 |
| ds002718 | 18 | - | 74 | 95 - 96 | 58 - 76 | 20 - 28 |
| ds002720 | 165 | - | 19 | 93 - 95 | 68 - 72 | 52 - 56 |
| ds002721 | 185 | - | 19 | 95 - 96 | 67 - 72 | 45 - 49 |
| ds002722 | 94 | - | 32 | 92 - 95 | 63 - 70 | 36 - 39 |
| ds002723 | 44 | - | 32 | 95 - 97 | 66 - 76 | 36 - 40 |
| ds002724 | 96 | - | 32 | 94 - 96 | 72 - 77 | 35 - 39 |
| ds002725 | 105 | - | 31 | 90 - 94 | 80 - 85 | 53 - 58 |
| ds002778 | 46 | - | 32 | 90 - 95 | 68 - 78 | 41 - 48 |
| ds002791 | 92 | - | 256 | 90 - 92 | 62 - 70 | 6 - 7 |
| ds002833 | 80 | - | 256 | 91 - 94 | 82 - 88 | 6 - 8 |
| ds002893 | 55 | 12 | 36 | 89 - 93 | 76 - 83 | 29 - 37 |
| ds003061 | 39 | - | 64 | 84 - 89 | 86 - 92 | 22 - 28 |
| ds003190 | 384 | 2 | 8 | 81 - 83 | 89 - 91 | 74 - 78 |
| ds003194 | 29 | - | 19 | 93 - 97 | 90 - 95 | 72 - 79 |
| ds003195 | 20 | - | 19 | 90 - 97 | 89 - 94 | 65 - 76 |

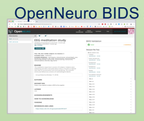
OpenNeuro BIDS #1



OpenNeuro BIDS #2



OpenNeuro BIDS #3

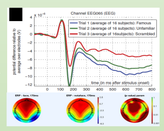


## NEM data meta-analysis!

HED event selection

EEGLAB automated pipeline

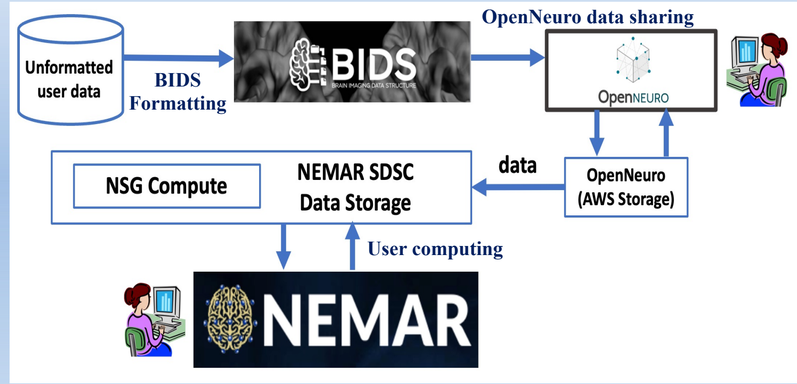
Output



Example using the integrated EEGLAB / NSG / NEMAR compute environment

## Data analysis pipeline for BIDS-format data (*bids.neuroimaging.io*)

Users first format their NEM data to BIDS specifications and then upload their data to the ***OpenNeuro.org*** web interface, which stores it on its AWS back end. The NEMAR SDSC Data Storage back-end then periodically copies and updates the data from OpenNeuro. Data statistics and visualizations are precomputed for display using NSG computing. The ***NEMAR.org*** web interface (bottom box) serves the data to NEMAR users. Dataset identifiers discovered using NEMAR or OpenNeuro can then be used in any custom analysis script the user uploads to the Neuroscience Gateway (***www.nsgportal.org***), which retrieves and processes the selected data without need for data download and subsequent re-upload.





Funding: R24-MH120037, R01-EB023297, RF1-MH126700, R01-NS04293