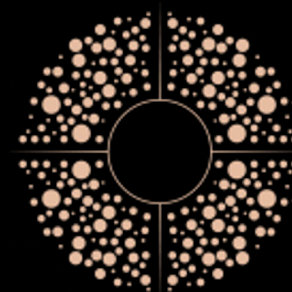
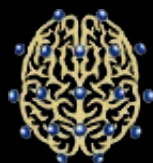




BIDS
BRAIN IMAGING DATA STRUCTURE



open
data



NEMAR

Arnaud Delorme



CerCo
UMR5549

Open Data

Findable

Data and supplementary materials have sufficiently rich metadata and a unique and persistent identifier.

Accessible

Data is deposited in a trusted repository.

Interoperable

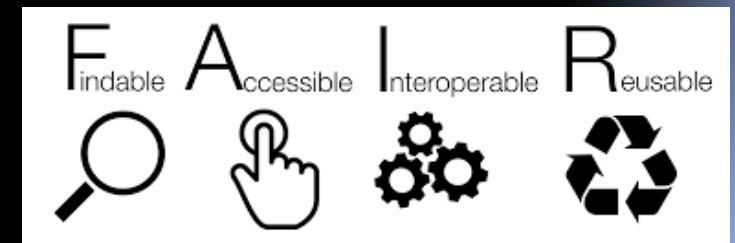
(Meta)data uses a formal, shared, and broadly applicable language or format.

Reusable

Data is described with clear and understandable attributes.

Anonymized.

There should be a clear and acceptable license for re-use.



<https://www.force11.org/group/fairgroup/fairprinciples>

Open Data - BIDS

Findable

Data and supplementary materials have sufficiently rich metadata and a unique and persistent identifier.

Accessible

Data is deposited in a trusted repository.

Interoperable

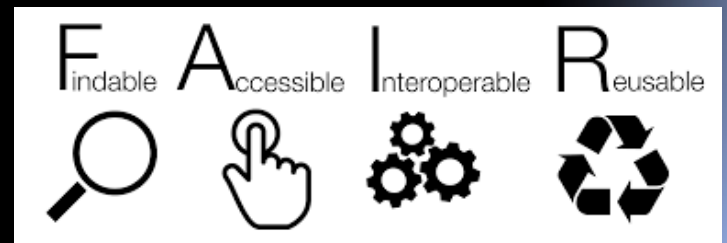
(Meta)data uses a formal, shared, and broadly applicable language or format.

Reusable

Data is described with clear and understandable attributes.

Anonymized.

There should be a clear and acceptable license for re-use.



<https://www.force11.org/group/fairgroup/fairprinciples>

What is BIDS?

BIDS is a way to organize your existing raw data

- ▶ To improve consistent and complete documentation
- ▶ To facilitate re-use by your future self and others

BIDS is not

- ▶ A new file format
- ▶ A search engine
- ▶ A data sharing tool



SCIENTIFIC DATA

OPEN The brain imaging data structure, a format for organizing and describing outputs of neuroimaging experiments

SUBJECT CATEGORIES

- » Data publication and archiving
- » Research data

SCIENTIFIC DATA

OPEN Comment: MEG-BIDS, the brain imaging data structure extended to magnetoencephalography

SCIENTIFIC DATA

OPEN EEG-BIDS, an extension to the brain imaging data structure for electroencephalography

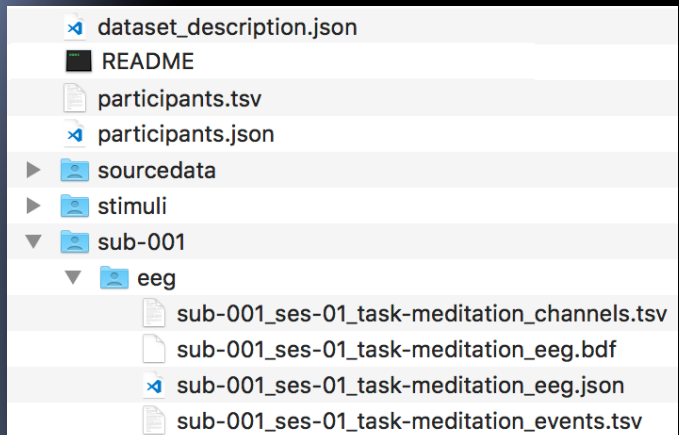
COMMENT

Received: 16 January 2019
Accepted: 7 May 2019
Published online: 25 June 2019

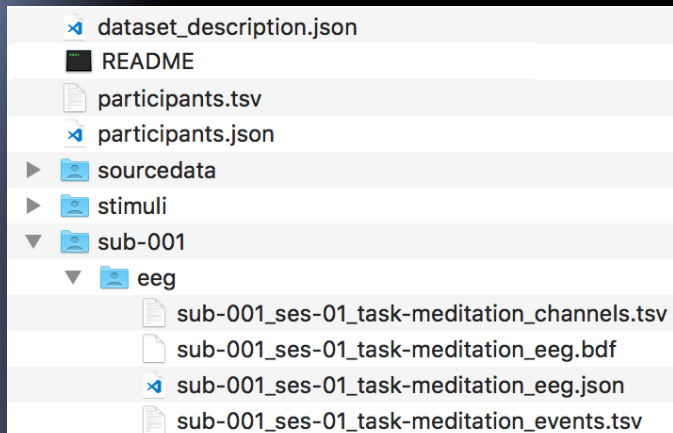
Cyril R. Pernet¹, Stefan Appelhoff², Krzysztof J. Gorgolewski³, Guillaume Flandin⁴, Christophe Phillips⁵, Arnaud Delorme^{6,7} & Robert Oostenveld^{8,9}

The Brain Imaging Data Structure (BIDS) project is a rapidly evolving effort in the human brain imaging research community to create standards allowing researchers to readily

BIDS-EEG



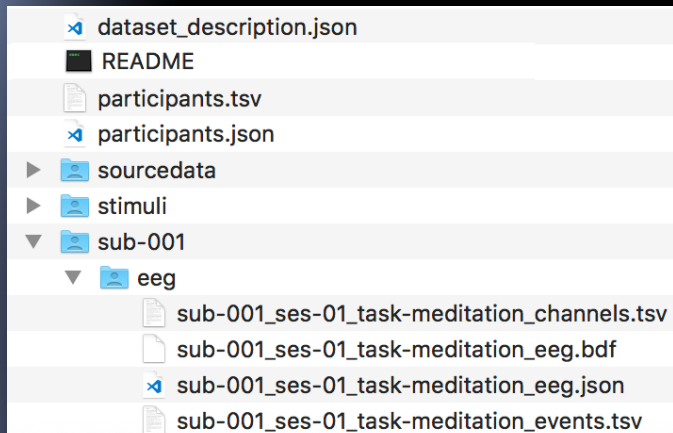
BIDS-EEG



► description of dataset in JSON format

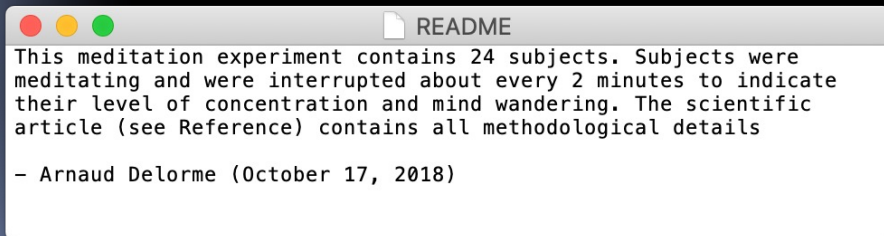
```
dataset_description.json
{
  "Name": "Meditation study",
  "ReferencesAndLinks": ["https://www.ncbi.nlm.nih.gov/pubmed/27815577" ],
  "License": "CC0",
  "BIDSVersion": "1.1.1"
}
```

BIDS-EEG

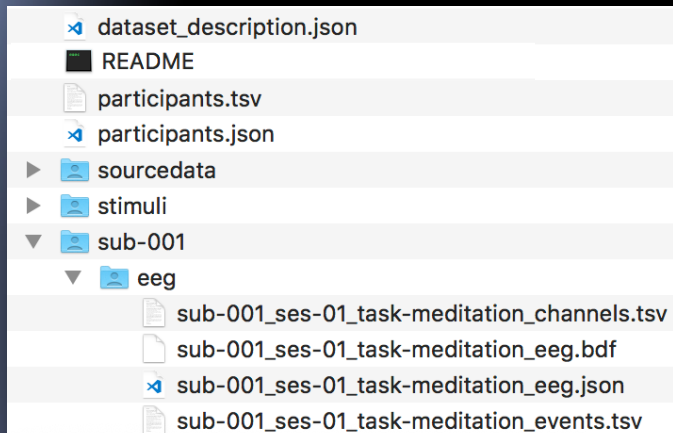


▶ description of dataset in JSON format

▶ readme file for users



BIDS-EEG

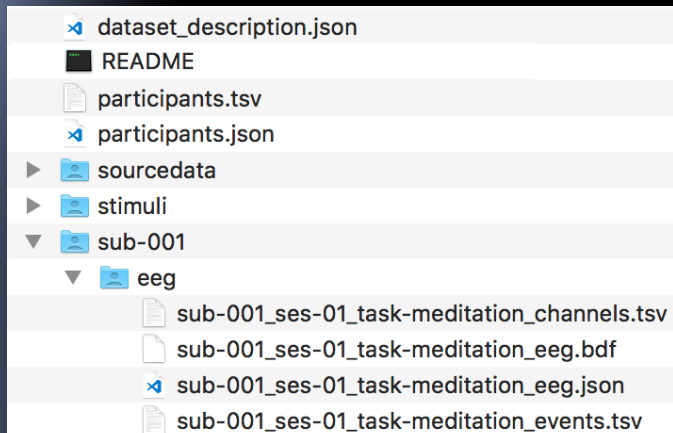


- ▶ description of dataset in JSON format
- ▶ readme file for users
- ▶ participants table files in tab delimited format

A screenshot of a text editor showing the contents of the participants.tsv file. The file is titled "participants.tsv" and contains the following data:

participant_id	gender	age	group
sub-001	M	32	expert
sub-002	M	35	expert
sub-013	F	47	novice
sub-016	M	77	novice

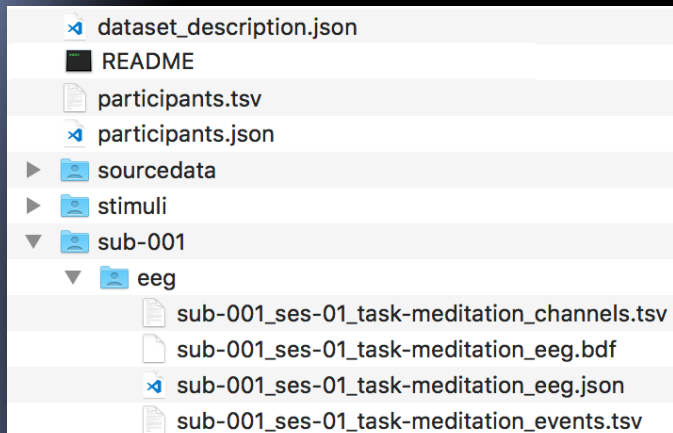
BIDS-EEG



- ▶ description of dataset in JSON format
- ▶ readme file for users
- ▶ participants table files in tab delimited format
- ▶ description of participants table file columns (above)

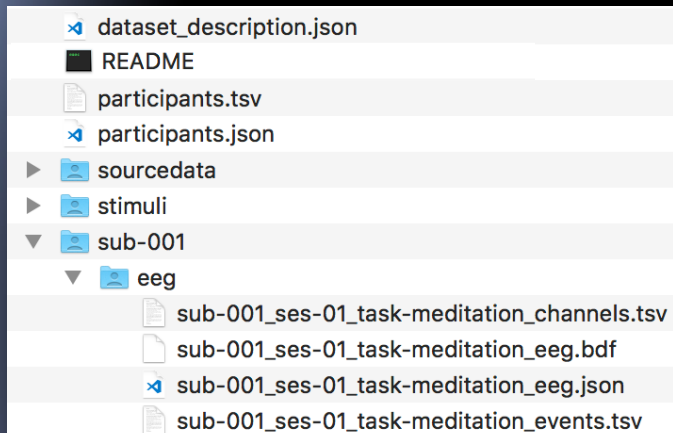
```
participants.json
{
  "participant_id": {
    "Description": "unique participant identifier"
  },
  "gender": {
    "Description": "sex of the participant",
    "Levels": {
      "M": "male",
      "F": "female"
    }
  },
  "age": {
```

BIDS-EEG



- ▶ description of dataset in JSON format
- ▶ readme file for users
- ▶ participants table files in tab delimited format
- ▶ description of participants table file columns (above)
- ▶ original raw data if converted to a supported format
- ▶ original stimuli (sound files and images)
- ▶ anonymized subject 1 folder

BIDS-EEG

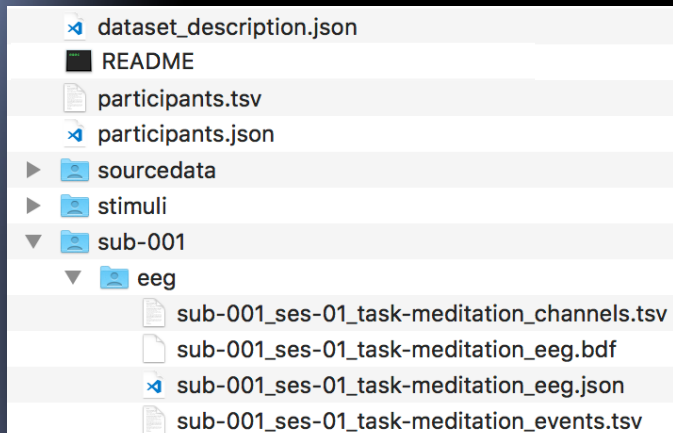


- ▶ description of dataset in JSON format
- ▶ readme file for users
- ▶ participants table files in tab delimited format
- ▶ description of participants table file columns (above)
- ▶ original raw data if converted to a supported format
- ▶ original stimuli (sound files and images)
- ▶ anonymized subject 1 folder
- ▶ file describing channels

A screenshot of a text editor window showing the content of the file sub-001_ses-01_task-meditation_channels.tsv. The content is a table with three columns: name, type, and units.

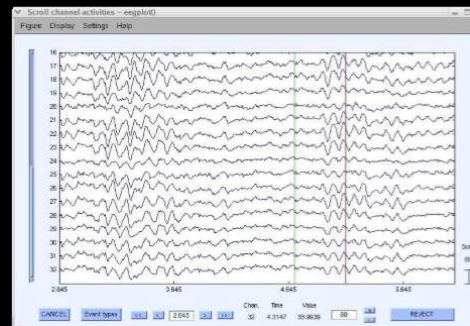
name	type	units
Fp1	EEG	microV
AF7	EEG	microV
AF3	EEG	microV
F1	EEG	microV
F3	EEG	microV
F5	EEG	microV
F7	EEG	microV
FT7	EEG	microV
FC5	EEG	microV
FC3	EEG	microV
FC1	EEG	microV

BIDS-EEG

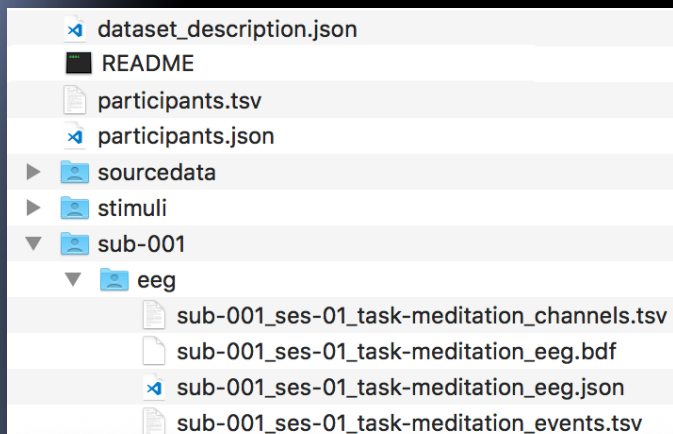


- ▶ description of dataset in JSON format
- ▶ readme file for users
- ▶ participants table files in tab delimited format
- ▶ description of participants table file columns (above)
- ▶ original raw data if converted to a supported format
- ▶ original stimuli (sound files and images)
- ▶ anonymized subject 1 folder
- ▶ file describing channels
- ▶ raw EEG data file (not all raw formats are possible)

```
00000000 00000000 00000000 11101000 10110111 01000010 00000000 00000000
10001011 01010100 00100100 00010000 01101000 01101000 00001101 01000011
00000000 01010010 10001101 01000100 00100100 01110000 01010000 11101000
01110011 01001010 00000000 00000000 11000110 10000100 00100100 11011000
00000000 00000000 00000000 00001001 01010111 10001101 01001100 00100100
01101100 11101000 10010001 01000010 00000000 00000000 10001011 01001100
00100100 00010000 01010001 10001101 01010100 00100100 01001000 01010010
11101000 00110010 01000001 00000000 00000000 11000110 10000100 00100100
11011000 00000000 00000000 00000000 00001010 10111001 00000000 10111011
01000011 00000000 11101000 00110000 01100010 00000000 00000000 01010000
10001101 01001100 00100100 01001000 11101000 10000110 00111010 00000000
00000000 10001011 01000100 00100100 00010000 01101000 10100100 00001101
01000011 00000000 01010000 10001101 01001100 00100100 01011000 01010001
11101000 00100010 01001010 00000000 00000000 11000110 10000100 00100100
11011000 00000000 00000000 00000000 00001011 01010011 10001101 01001100
00100100 01010100 11101000 01000000 01000010 00000000 00000000 10001011
```



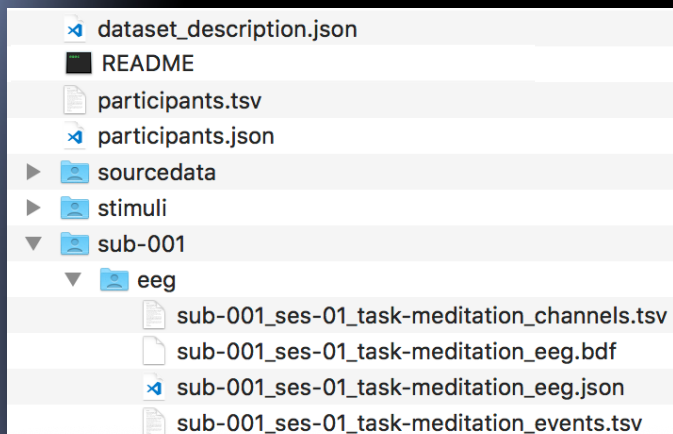
BIDS-EEG



- ▶ description of dataset in JSON format
- ▶ readme file for users
- ▶ participants table files in tab delimited format
- ▶ description of participants table file columns (above)
- ▶ original raw data if converted to a supported format
- ▶ original stimuli (sound files and images)
- ▶ anonymized subject 1 folder
- ▶ file describing channels
- ▶ raw EEG data file (not all raw formats are possible)
- ▶ amplifier and recording information

```
sub-001_ses-01_task-meditation_eeg.json
{
  "InstitutionAddress": "Place du Docteur Baylac, Pavillon Baudot, 31059 Toulouse, France",
  "InstitutionName": "Paul Sabatier University",
  "InstitutionalDepartmentName": "Centre de Recherche Cerveau et Cognition",
  "PowerLineFrequency": 50,
  "ManufacturersModelName": "ActiveTwo",
  "TaskName": "meditation",
  "EEGReference": "CMS/DRL",
  "Manufacturer": "BIOSEMI",
  "EEGChannelCount": 64,
  "MiscChannelCount": 15,
  "RecordingType": "continuous",
  "RecordingDuration": 2721,
  "SamplingFrequency": 256,|
  "EOGChannelCount": 0,
  "ECGChannelCount": 0,
  "EMGChannelCount": 0,
  "SoftwareFilters": "n/a"
}
```

BIDS-EEG



- ▶ description of dataset in JSON format
- ▶ readme file for users
- ▶ participants table files in tab delimited format
- ▶ description of participants table file columns (above)
- ▶ original raw data if converted to a supported format
- ▶ original stimuli (sound files and images)
- ▶ anonymized subject 1 folder
- ▶ file describing channels
- ▶ raw EEG data file (not all raw formats are possible)
- ▶ amplifier and recording information
- ▶ events in tabular format

onset	duration	trial_type	response_time	sample	value
71.3867187500	n/a	stimulus	n/a	18275	128
75.7304687500	n/a	response	n/a	19387	2
79.7734375000	n/a	response	n/a	20422	2
125.6093750000	n/a	stimulus	n/a	32156	128
179.8007812500	n/a	stimulus	n/a	46029	128
183.0976562500	n/a	response	n/a	46873	2
185.6328125000	n/a	response	n/a	47522	4
284.8593750000	n/a	stimulus	n/a	72924	128
287.7578125000	n/a	response	n/a	73666	2
289.9609375000	n/a	response	n/a	74230	2

BIDS specification:

<https://bids-specification.readthedocs.io/en/latest/>

This example:

<https://openneuro.org/datasets/ds001787>

BIDS EEGLAB plugin (export/import)

The image shows three overlapping windows from the BIDS EEGLAB plugin. The top window, titled "BIDS task information -- pop_taskinfo()", contains fields for "Dataset name" (Auditory Oddball Task), "Task name" (AuditoryOddball), and "BIDS EEG acquisition information" (Cap manufacturer: EasyCap, Cap model: 64-channel 10-20 cap, EEG reference location: Active, EEG montage system: 10-20, EEG amplifier maker: Blosemi). It also includes a README and participant task description.

The middle window, titled "Figure 2: Edit BIDS participant info - pop_participantinfo", shows a "Participant information" form with fields for participant_id, Gender, Age, Group, HeadCircumference, Authors, References and links, Task-relevant Cognitive Atlas term, Task-relevant CogPO term, Institution, Department, and Institution location. It also includes a "BIDS metadata for participant fields" table.

The bottom window, titled "Figure 1: Participant information", shows a table of participant information:






participant_id	Gender	Age	Group
s1	M	1	
s2	F	2	

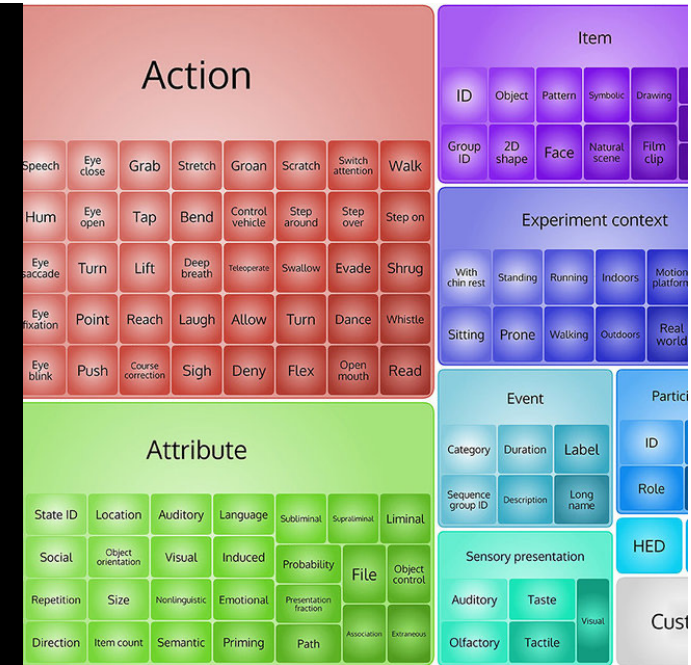
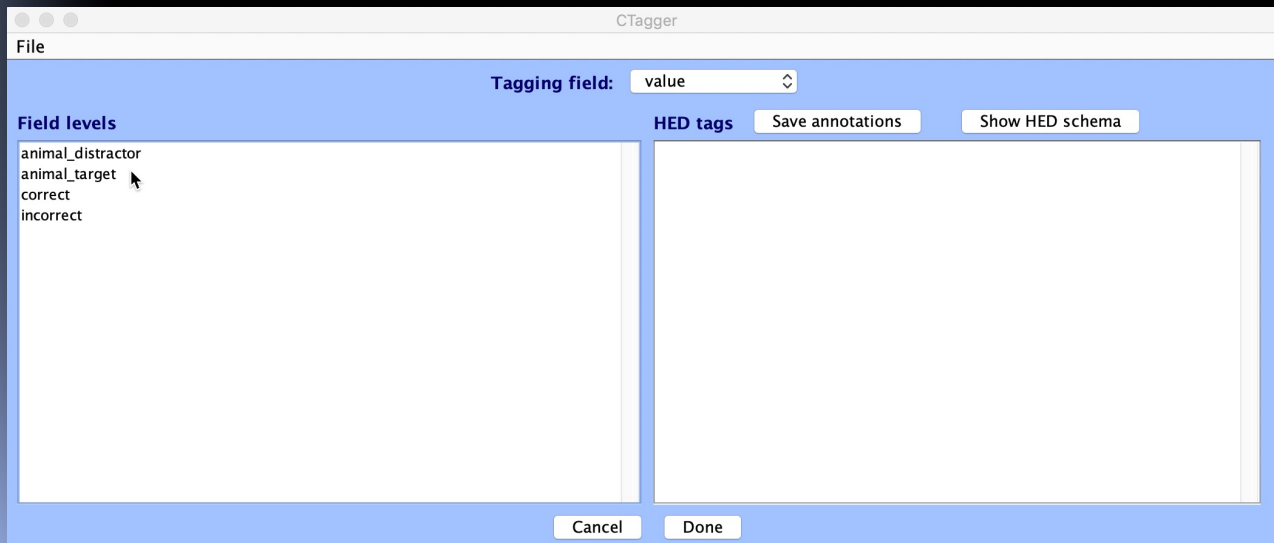
Below this table is a "Choose additional spreadsheet columns" dialog box with a list of columns: ParticipantID, Ethnicity, and Income. The "ParticipantID" column is selected.

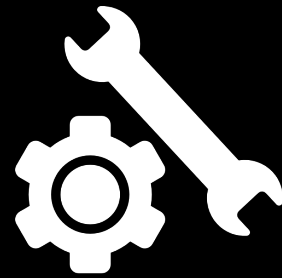
	Status
EEG import	✓
BEH import	✓
iEEG import	✓
MEG import	✓
Montion-cap import	✓
Eye-tracking import	✗
EEG export	✓
Eye-tracking export	✓
BEH export	✓
MRI export	✓
Motion-cap export	✓
iEEG export	✗
MEG export	✗

Task and event details

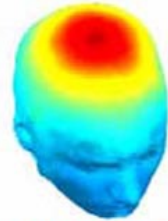
Capturing the nature of events and event context using Hierarchical Event Descriptors (HED)

 Kay Robbins,  Dung Truong,  Stefan Appelhoff,  Arnaud Delorme,  Scott Makeig
doi: <https://doi.org/10.1101/2021.05.06.442841>





TOOLS



EEGLAB



BRAIN PRODUCTS
Solutions for neurophysiological research



BIDScoin



Brainstorm

MNE

MEG + EEG ANALYSIS & VISUALIZATION

<https://bids.neuroimaging.io/benefits.html>



Neuroelectromagnetic Data Archive and Tools Resource




The screenshot shows the OpenNEURO website interface. The main content area displays the 'EEG meditation study' dataset page. Key information includes: 'Files: 180, Size: 5.69GB, Subjects: 24, Sessions: 3', 'Available Tasks: meditation', and 'Available Modalities: /participants, /stimuli/cancel, /stimuli/depth, over, wandering, meditation, mw, tired, /stimuli/resume, /stimuli/resumed, /stimuli/resumemed, /stimuli/self, /stimuli/starting, /stimuli/time, /stimuli/valence, channels, eeg, events'. A 'BIDS Validation' section shows a 'Valid' status with '2 WARNINGS'. A 'Dataset File Tree' is visible on the right, listing files like 'dataset_description.json', 'participants.json', 'participants.tsv', 'README', 'task-meditation_events.json', 'code', 'stimuli', and subfolders 'sub-001' through 'sub-012'. The top navigation bar includes 'MY DASHBOARD', 'PUBLIC DASHBOARD', 'SUPPORT', 'FAQ', and 'UPLOAD DATASET'.

The screenshot shows the NEMAR website home page. The header includes the NEMAR logo and navigation links: 'About', 'Discover', 'Community', 'Support', and 'Login'. The main content area features a large, stylized brain graphic with the text 'Search, Visualize, Analyze, Download' overlaid. Below this, a message states: 'An open access data, tools, and compute resource operating on NeuroElectroMagnetic data made available in OpenNeuro.org.' At the bottom, there is a search bar with the placeholder text 'Search for Datasets and Tools' and a search input field.


The EEGLAB Youtube Channel

Delorme A. EEG is better left alone. Sci Rep. 2023 Feb 9;13(1):2372. doi: 10.1038/s41598-023-27528-0. PMID: 36759667; PMCID: PMC9911389.

 arnodelorme	Update process_eeglab_template.m	d456aa4 on Jul 24	🕒 21 commits
📁 brainstrom	Update process_brainstorm_template.m		7 months ago
📁 data	intial commit		last year
📁 eeglab	Update process_eeglab_template.m		3 months ago
📁 fieldtrip	Update process_fieldtrip_template.m		7 months ago
📁 hatte	updating pipeline to match the final paper		7 months ago
📁 mne	Update process_mne_template.py		7 months ago
📄 .gitignore	updating pipeline to match the final paper		7 months ago
📄 README.md	Update README.md		5 months ago

A simple EEG
automated pipeline
data in 

A simple EEG
automated pipeline
data in 

A simple EEG
automated pipeline
data in 

A simple EEG
automated pipeline
data in 

NEMAR meta-data analysis

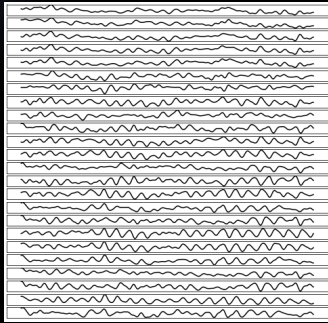
	README	Task Description	Instructions	Event Description	EEG Reference	Power Line Frequency	Channel Types	Electrode Locations	Participants' Age and Gender	Subject Artefact Description	Event Consistency	Channel Consistency	Agregated Score
ds000117													0.18
ds001784													0.73
ds001787													0.67
ds001810													0.75
ds001849													0.45
ds001971													0.83
ds002034													0.58
ds002094													0.45
ds002158													0.33
ds002218													0.67
ds002336													0.36
ds002338													0.36
ds002578													0.75
ds002680													0.67
ds002691													0.75
ds002718													0.83
ds002720													0.5
ds002721													0.5
ds002722													0.42
ds002723													0.42
ds002724													0.42
ds002725													0.42
ds002778													0.45
ds002791													0.45
ds002833													0.64
ds002893													0.83
ds003061													0.83
ds003190													0.55
ds003194													0.45
ds003195													0.55

NEMAR data analysis

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

Delorme, A., Truong, D., Martinez- Cancino, R., Pernet, C., Sivagnanam, S., Yoshimoto, K., Poldrack, R., Majumdar, A., Makeig, S. (2020) Tools for Importing and Evaluating BIDS-EEG Formatted Data. Proceeding of the IEEE transactions on biomedical engineering. 2021 10th International IEEE/EMBS Conference on Neural Engineering (NER) May 4-8, 2021, Rome, Italy.

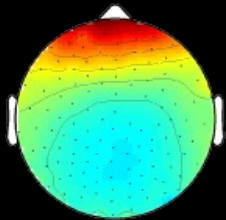
Raw data



Time but no spatial information

vs

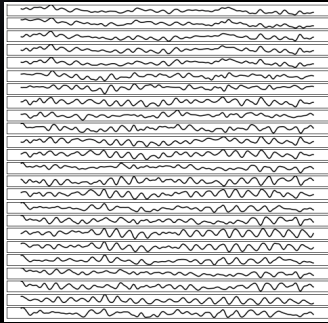
Spectral power



Spatial but no time information

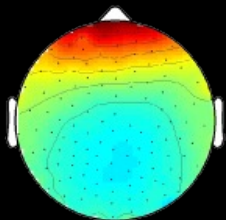


Raw data



vs

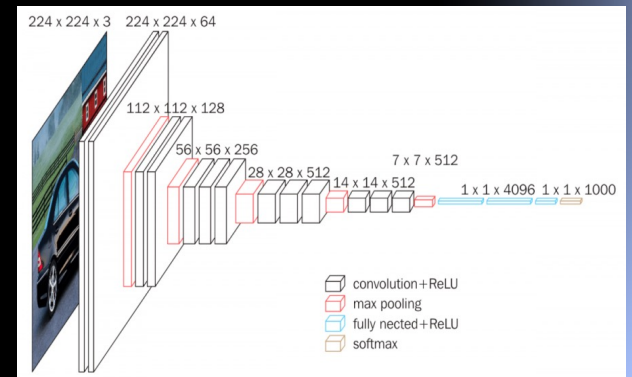
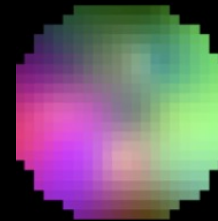
Spectral power



theta

alpha

beta



2021 43rd Annual International Conference of the
IEEE Engineering in Medicine & Biology Society (EMBC)
Oct 31 - Nov 4, 2021. Virtual Conference

Deep Convolutional Neural Network Applied to Electroencephalography: Raw Data vs Spectral Features

Dung Truong, Michael Milham, Scott Makeig, Arnaud Delorme

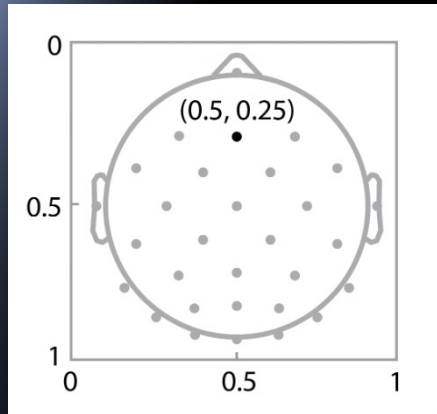
- 1574 participants (50% female) 23-channels (child mind)
- 2-second extracted epochs eyes closed (63k samples)
- predictors: **sex**
- 2 convolutional DL models (SCNN & VGG)
- Raw data or Spectral data given as input

Raw data

Spectral data

Model	Per-sample
R-SCNN	80.6 (79.7 to 81.5)
R-VGG	83.1 (82.7 to 83.4)
S-SCNN	79.0 (78.7 to 79.3)
S-VGG	77.1 (76.8 to 77.4)

<https://github.com/dungscout96/DL-EEG>

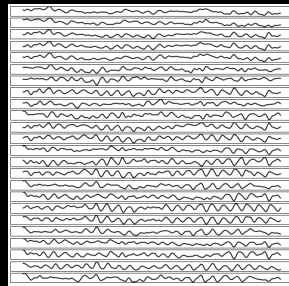


Spatial attention

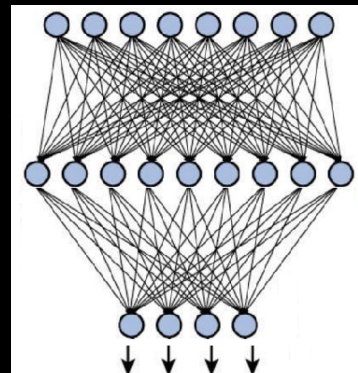
$$a_j(x, y) = \sum_{k=1}^K \sum_{l=1}^K \text{Re}(z_j^{(k,l)}) \cos(2\pi(kx + ly)) + \text{Im}(z_j^{(k,l)}) \sin(2\pi(kx + ly))$$

$$\forall j \in [D1], SA(X)^{(j)} = \frac{1}{\sum_{i=1}^c e^{a_j(x_i, y_i)}} \left(\sum_{i=1}^c e^{a_j(x_i, y_i)} X^{(i)} \right)$$

Raw EEG n channel



Attention layer m channel



Deep learning applied to EEG data with different montages using spatial attention

Dung Truong*
SCCN, INC, UCSD, La Jolla CA, USA
dtruong@ucsd.edu
<https://orcid.org/0000-0003-4540-3551>

Muhammad Abdullah Khalid*
SCCN, INC, UCSD, La Jolla CA, USA
mkhalid.bee18seecs@seecs.edu.pk

Arnaud Delorme
SCCN, INC, UCSD, La Jolla CA, USA |
CerCo CNRS, Paul Sabatier University,
Toulouse, France
arnodelorme@gmail.com
<https://orcid.org/0000-0002-0799-3557>

Model→ Attention↓	128-channel model	23-channel model
No spatial attention	80.4 (0.8)	78.0 (1.8)
Spatial attention	83.7 (1.5)	80.3 (1.4)

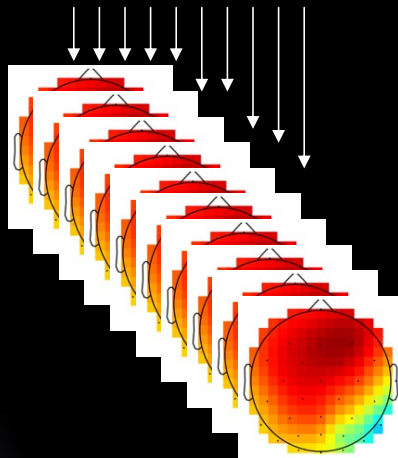
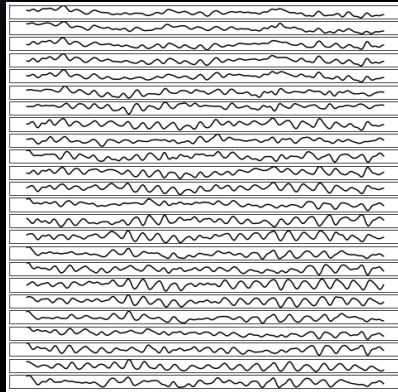
Table 2. Mean classification accuracy (and standard deviation in parenthesis) for the 23-channel and 128-channel models with or without using spatial attention.

Model→ Test data↓	128 channel	23 channel	Mixed channel
128 channel	83.7 (1.5)	76.9 (1.2)	81.4 (2.2)
23 channel	57.2 (3.0)	80.3 (1.4)	79.8 (2.5)
Mixed channel	70.1 (4.2)	78.6 (2.0)	78.9 (1.6)

Table 3. Classification accuracy (standard deviation in parenthesis) of models trained on different channel counts experiments, all while applying spatial attention.

<https://github.com/sccn/deep-channel-harmonization>

Raw data



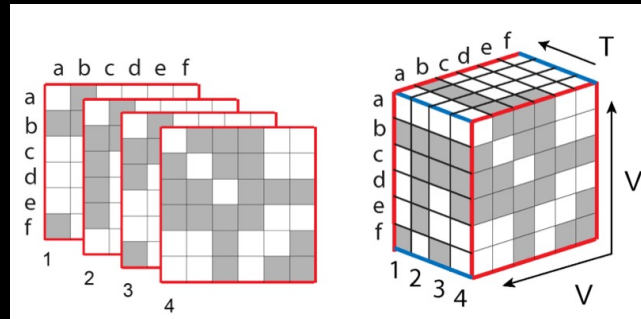
2022 44th Annual International Conference of
the IEEE Engineering in Medicine & Biology Society (EMBC)
Scottish Event Campus, Glasgow, UK, July 11-15, 2022

A streamable large-scale clinical EEG dataset for Deep Learning

Dung Truong¹, Manisha Sinha², Kannan Umadevi Venkataraju², Michael Milham³, Arnaud Delorme^{1,4}

1500 subjects converted
and streamable

3-D array input instead of 2-D



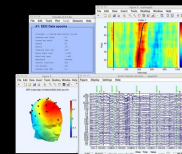
```
import torch
import webdataset as wds
def selectLabel(x, lbl):
    # function to select desired label
    lbl_idx =
    ["id", "sex", "age", "handedness", "index"].index(lbl.lower())
    x = x.decode("utf-8").split(",")[lbl_idx]
    return x if lbl_idx == 0 else float(x)

s3_url = 'https://childmind.s3.us-west-1.amazonaws.com/python/childmind_train.tar' # replace
'train' with 'val' and 'test' accordingly

train_data =
wds.WebDataset(s3_url).decode().map_dict(cls=lambda
x: selectLabel(x, 'ID')).to_tuple("numpy", "cls")

# Check out first sample and its label
sample, label = next(iter(train_data))
print(f'Sample size: {sample.shape}') # (24, 256)
print(f'Label: {label}') # NDARFB908HVX
```

EEGLAB



ML applications



https://github.com/scn/study_ml



