



What is the Brain Imaging Data Structure and why you should know about this!

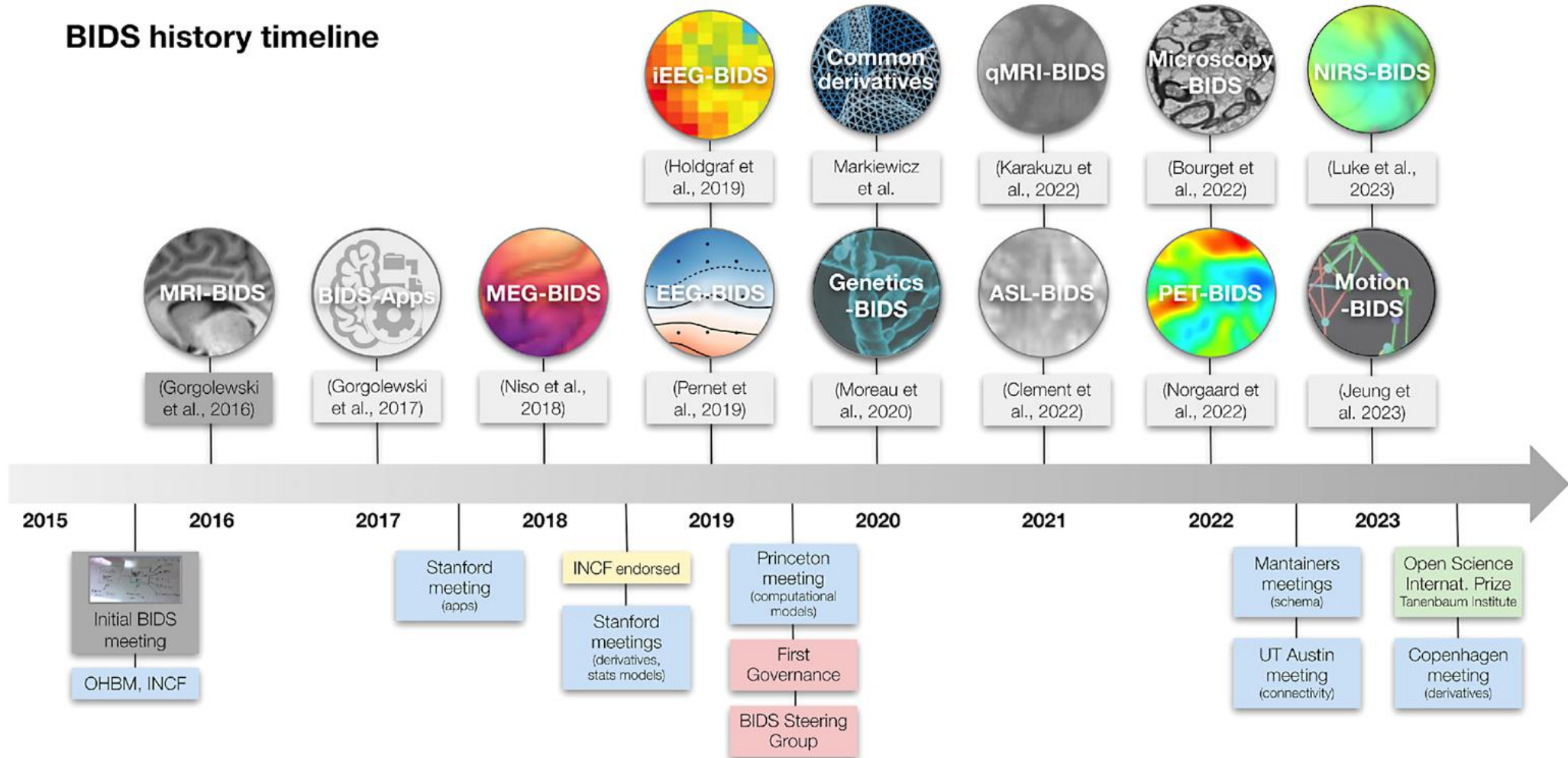
07/11/2024
Julius Welzel



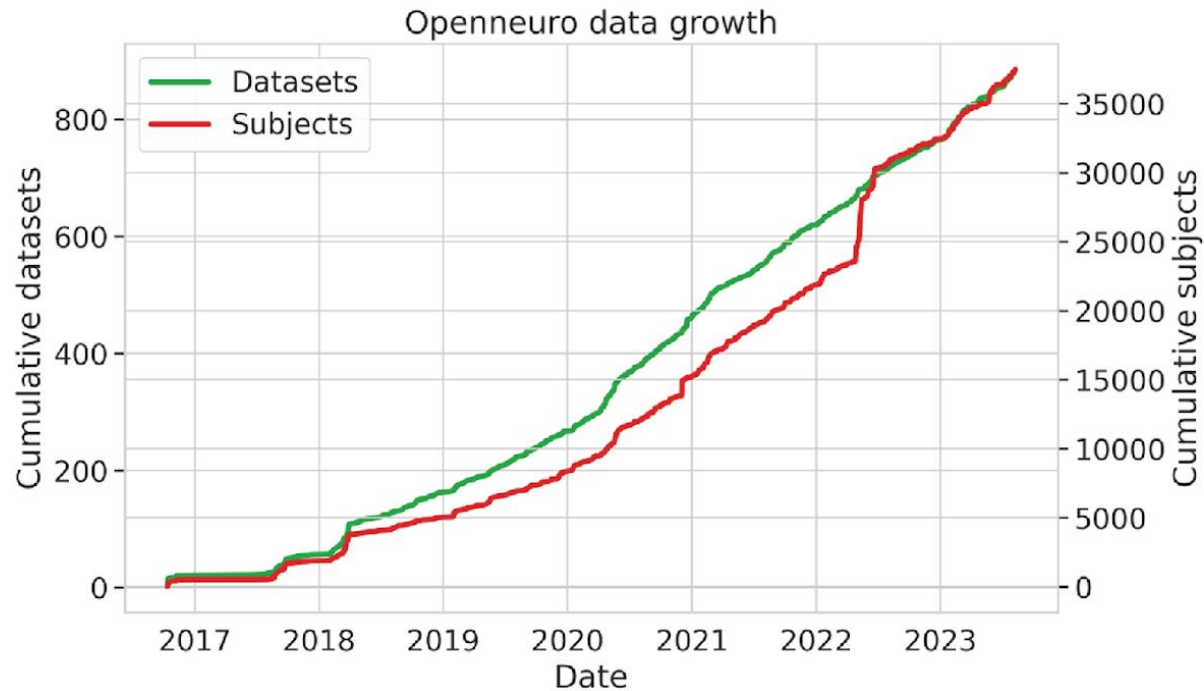
- README.md
- dataset_description.json
- participants.json
- participants.tsv
 - sub-001
 - ses-01
 - eeg/
 - motion/
 - ses-02
 - sub-002


What is the Brain Imaging Data Structure and why you should know about this!

The Past, Present, and Future of BIDS



The Past, Present, and Future of BIDS



 **OpenNEURO**

SEARCH SUPPORT DOCUMENTATION

OpenNeuro EEG

OpenNeuro added support for EEG datasets in 2019 when EEG was incorporated into the BIDS standard.

7.926 Participants **200** Public Datasets

Search EEG Portal

[Search at the participant-level with Neurobagel ?](#)

Keywords ?

These filters return **70** datasets:

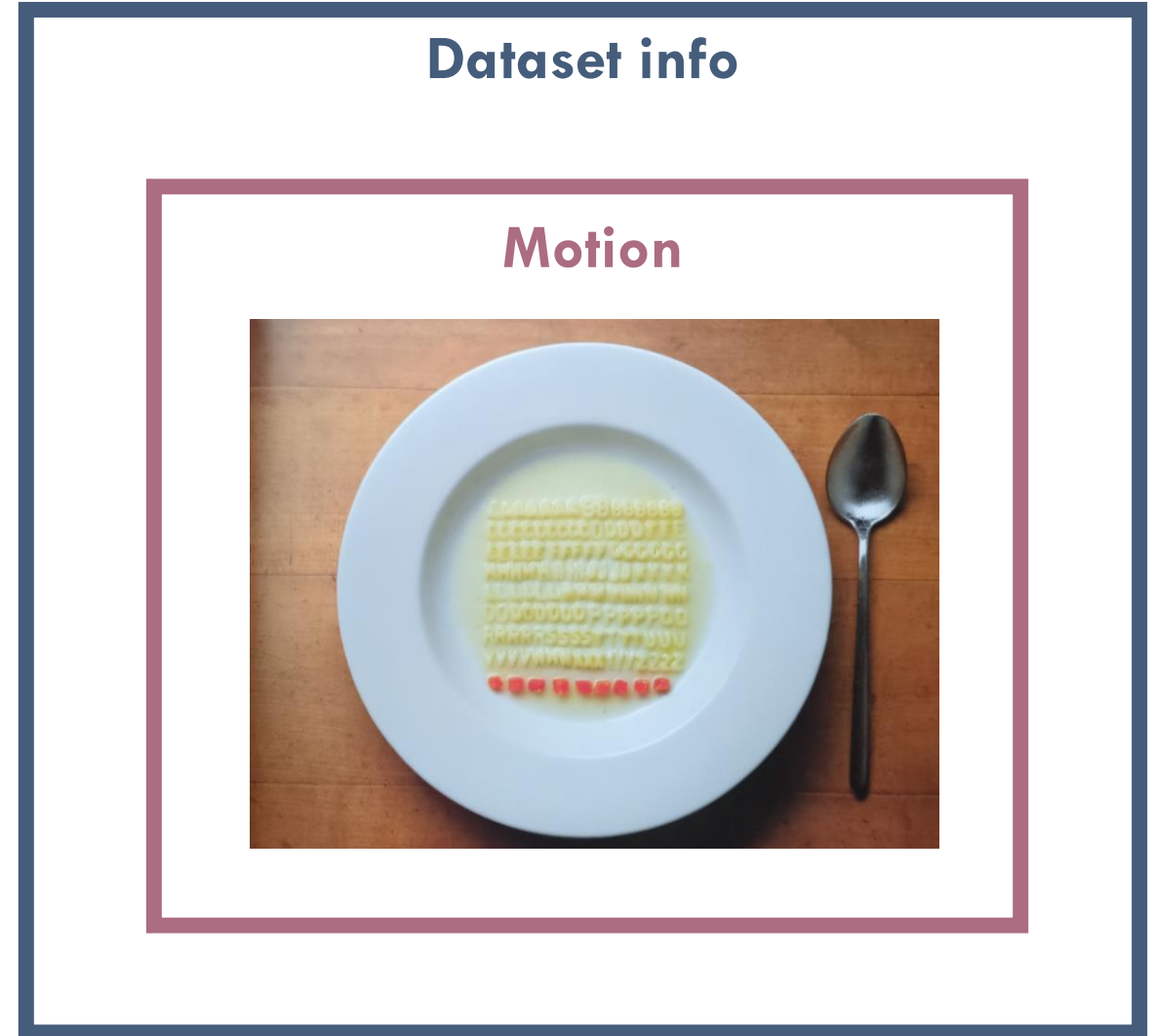
KEYWORD:

MODALITY:

What is BIDS?



About 20+ different EEG systems
record in different file formats



Aims of BIDS

→ Makes (neuroimaging-)data FAIR



DATA STRUCTURE

- BIDS is based on simple file formats and folder structures
- Provides a “easy-to-understand” folder structure
- Provides a naming convention for files and directories
- Specifies file formats to use for modalities

METADATA

- Prevents metadata getting lost
- Some metadata is better than no metadata (80/20 rule)



 Brain Imaging Data Structure v1.9.0  Search

Brain Imaging Data Structure v1.9.0

The BIDS Specification

Introduction

Common principles

Modality agnostic files

Modality specific files >

Derivatives >

Longitudinal and multi-site studies

Glossary

BIDS Extension Proposals

Appendix >

Changelog

The BIDS Starter Kit

Website

Tutorials

GitHub repository

The Brain Imaging Data Structure

The Brain Imaging Data Structure (BIDS) is a simple and intuitive way to organize and describe data.

This document defines the BIDS specification, which provides many details to help implement the standard. It includes the core specification as well as many extensions to specific brain imaging modalities, and increasingly also to other kinds of data.

If BIDS is new to you, and you would like to learn more about how to adapt your own datasets to match the BIDS specification, we recommend exploring the [BIDS Starter Kit](#). Alternatively, to get started please read [the introduction to the specification](#).

For an overview of the BIDS ecosystem, visit the [BIDS homepage](#). The entire specification can also be [downloaded as PDF](#).



Remi Gau

Remi-Gau

Follow

I wanted to understand the brain, now I tell people how to name files, folders and variables.

**My favourite two
BIDS principles**

Requirement levels

Level of requirement **for**
files and **within** files:

1. **REQUIRED**
2. **RECOMMENDED**
3. **OPTIONAL**

```
├── README.md
├── dataset_description.json
├── participants.tsv
├── sub-001
│   ├── ses-01
│   │   ├── eeg/
│   │   └── motion/
│   └── ses-02
└── sub-002
```

| participant_id | age | group |
|----------------|-----|-------|
| sub-001 | 34 | C |
| sub-002 | 12 | P |
| sub-003 | 33 | M |

BIDS definitions

Dataset - A set of neuroimaging and behavioral data acquired for a purpose of a particular study. A dataset consists of data acquired from one or more subjects, possibly from multiple sessions.

Subject - A person or animal participating in the study. Used interchangeably with term **Participant**.

Session - A logical grouping of neuroimaging and behavioral data consistent across subjects.

Task - A set of structured activities performed by the participant.

Overview files:

- participant.tsv
- scans.tsv
- ...

Subject, session and task specific files

- sub-001_ses-01_task-GoNoGo.eeg
- sub-001_task-Oddball_channels.tsv

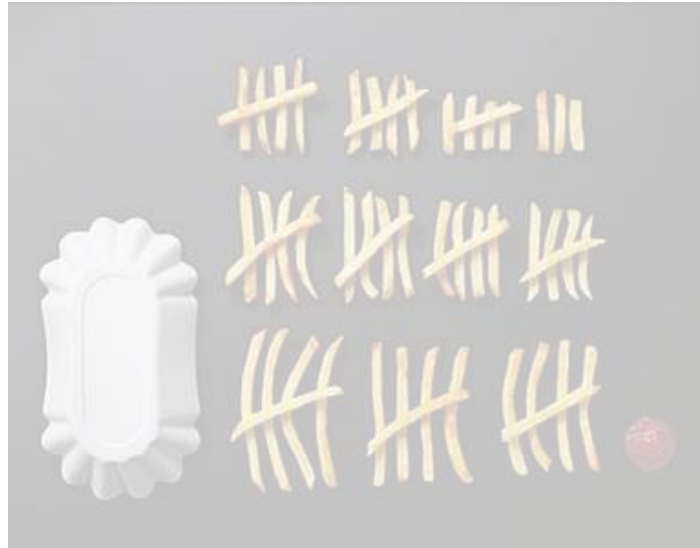
Directories and file naming structure

- Data for each subject are organized in subdirectories labeled "**sub-label**", with "<label>" representing the unique identification label of each subject
- If subjects participated in multiple sessions, subdirectories labeled "**ses-label**" contain session-specific data within the subject directory
- Each session subdirectory (or the subject subdirectory if no session data exists) contains subdirectories for different data types, only defined if files are present for that type
- For a data file that was collected in a given **session** from a given **subject**, the file name **MUST** begin with the string **sub-label****[_ses-label]****_task-label**

```
├── README.md
├── dataset_description.json
├── participants.json
├── participants.tsv
└── sub-001
    ├── ses-01
    │   └── eeg/
    │       └── sub-001_ses-01_task-GoNoGo.eeg
```

Dataset info

EEG



- README.md
- dataset_description.json
- participants.json
- participants.tsv
- └─ sub-001
 - ses-01
 - eeg/
 - motion/
 - ses-02
- └─ sub-002

Modality agnostic data

dataset_description.json

├── README.md
├──
dataset_description.json
├── participants.json
├── participants.tsv
├── sub-001
│ ├── ses-01
│ └── ses-02
└── sub-002



```
{  
  "Name": "The mother of all experiments",  
  "BIDSVersion": "1.4.0",  
  "DatasetType": "raw",  
  "License": "CC0",  
  "Authors": [ "Paul Broca", "Carl Wernicke" ],  
  "Funding": [ „NIH F37823MFH1" ],  
  "EthicsApprovals": [ „HRPO (Protocol AR0928" ],  
  "DatasetDOI": "10.0.2.3/dfjj.10"  
}
```

participants

- README.md
- dataset_description.json
- **participants.tsv**
- **participants.json**
- sub-001
 - ses-01
 - ses-02
- sub-002

participants.tsv

| participant_id | age | group |
|----------------|-----|-------|
| sub-001 | 34 | C |
| sub-002 | 12 | P |
| sub-003 | 33 | M |

participants.json

```
"age": {  
  "Description": "age of the  
  participant", "Units": "years"  
},  
„group“: {  
  "Description": „assigned group“,  
  "Levels": {  
    „C“: „control“,  
    „M“: „medication“,  
    „P“: „placebo“,  
  }  
}
```


Dataset info

EEG

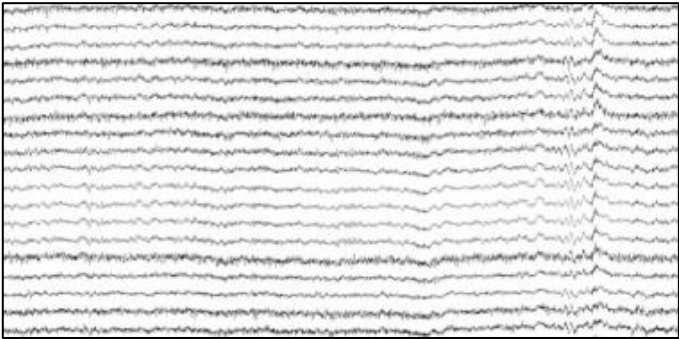


- README.md
- dataset_description.json
- participants.json
- participants.tsv
 - sub-001
 - ses-01
 - eeg/
 - motion/
 - ses-02
 - sub-002

Modality specific data

Files | eeg

└─ sub-001
└─ ses-01
└─ eeg/



```
{
  "TaskName": "TASKNAME",
  "SamplingFrequency": 1000,
  "SoftwareFilters": "n/a",
  "EEGChannelCount": 4,
  "EOGChannelCount": 1,
  "EEGReference": "placed on Cz",
  "PowerLineFrequency": 50
}
```

- └─ sub-001_ses-01_task-<label>_eeg.<extension>
- └─ sub-001_ses-01_task-<label>_eeg.json
- └─ sub-001_ses-01_task-<label>_channels.tsv
- └─ sub-001_ses-01_task-<label>_electrodes.tsv
- └─ sub-001_ses-01_task-<label>_electrodes.tsv

| name | type | units | status | status_description |
|------|------|--------|--------|--------------------|
| CP5 | EEG | microV | good | n/a |
| FC5 | EEG | microV | bad | high freq noise |
| FC1 | EEG | microV | good | n/a |
| C3 | EEG | microV | good | n/a |
| VEOG | EOG | microV | good | n/a |

| name | x | y | z | impedance |
|------|-------|-------|------|-----------|
| CP5 | -0.77 | -0.30 | 0.57 | 8 |
| FC5 | -0.77 | 0.30 | 0.57 | 12 |
| FC1 | -0.29 | 0.31 | 0.91 | 2 |
| C3 | -0.59 | 0.00 | 0.81 | 5 |
| VEOG | n/a | n/a | n/a | n/a |

```
{
  "EEGCoordinateSystem": "T1w",
  "EEGCoordinateUnits": "mm",
  "AnatomicalLandmarkCoordinates": {
    "LPA": [-0.067, 1.736e-09, -3.844e-09],
    "NAS": [-4.11e-09, 0.091, -4.541e-10],
    "RPA": [0.064, -6.435e-09, -4.566e-09]
  },
  "AnatomicalLandmarkCoordinateSystem": "T1w",
  "AnatomicalLandmarkCoordinateUnits": "mm",
  "IntendedFor": "sub-01_T1w.nii.gz"
}
```

Metadata | eeg.json

REQUIRED

TaskName, SamplingFrequency,
PowerlineFrequency, EEGReference,
SoftwareFilters

RECOMMENDED

TaskDescription, Manufacturer,
EEGChannelCount,
EOGChannelCount, ...

OPTIONAL

-

```
{  
  TaskName: Seeing stuff,  
  SamplingFrequency: 2400,  
  Manufacturer: Brain Products,  
  ManufacturersModelName: BrainAmp DC,  
  EEGChannelCount: 87,  
  EOGChannelCount: 2,  
  TriggerChannelCount: 1,  
  PowerLineFrequency: 50,  
  EEGPlacementScheme: 10 percent system,  
  EEGReference: single electrode placed on FCz,  
  SoftwareFilters: {  
    Anti-aliasing filter: {  
      half-amplitude cutoff (Hz):  
        500,  
      Roll-off: 6dB/Octave  
    }  
  }  
}
```

Metadata | channels.tsv

REQUIRED

name, type, units

RECOMMENDED

-

OPTIONAL

description, sampling_frequency

| name | type | units | description | reference | status | status_description |
|---------|------|-------|--------------------------------|--------------|--------|----------------------|
| VEOG | VEOG | uV | left eye | VEOG-, VEOG+ | good | n/a |
| FDI | EMG | uV | left first dorsal interosseous | FDI-, FDI+ | good | n/a |
| Cz | EEG | uV | n/a | REF | bad | high frequency noise |
| UADC001 | MISC | n/a | envelope of audio signal | n/a | good | n/a |

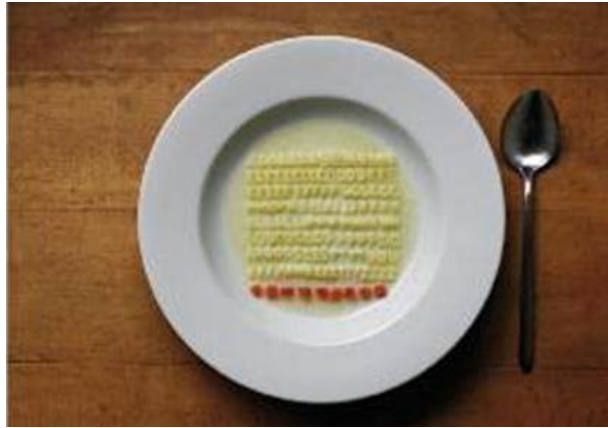
Metadata | electrodes.tsv

| name | x | y | z | type | material | color |
|-------|-------|-------|--------|---------|----------|--------|
| VEOG+ | n/a | n/a | n/a | cup | Ag/AgCl | blue |
| VEOG- | n/a | n/a | n/a | cup | Ag/AgCl | white |
| FDI+ | n/a | n/a | n/a | cup | Ag/AgCl | red |
| FDI- | n/a | n/a | n/a | cup | Ag/AgCl | red |
| GND | -0.07 | 0.00 | -0.070 | clip-on | Ag/AgCl | pink |
| Cz | 0.00 | 0.07 | 0.06 | cup | Ag/AgCl | yellow |
| REF | -0.07 | -0.02 | -0.01 | cup | Ag/AgCl | grey |

| Column name | Requirement Level | Data type | Description |
|--------------------|-------------------|-----------------|--|
| name | REQUIRED | string | Name of the electrode contact point. Values in <code>name</code> MUST be unique. This column must appear first in the file. |
| x | REQUIRED | number | Recorded position along the x-axis. This column must appear second in the file. |
| y | REQUIRED | number | Recorded position along the y-axis. This column must appear third in the file. |
| z | REQUIRED | number or "n/a" | Recorded position along the z-axis. This column must appear fourth in the file. |
| type | RECOMMENDED | string | Type of the electrode (for example, cup, ring, clip-on, wire, needle). This column may appear anywhere in the file. |
| material | RECOMMENDED | string | Material of the electrode (for example, <code>Tin</code> , <code>Ag/AgCl</code> , <code>Gold</code>). This column may appear anywhere in the file. |
| impedance | RECOMMENDED | number | Impedance of the electrode, units MUST be in <code>kOhm</code> . This column may appear anywhere in the file. |
| Additional Columns | OPTIONAL | n/a | Additional columns are allowed if they are defined in the associated metadata file. |

Dataset info

Motion



- └─ README.md
- └─ dataset_description.json
- └─ participants.json
- └─ participants.tsv
- └─ sub-001
 - └─ ses-01
 - └─ eeg/
 - └─ **motion/**
 - └─ ses-02
- └─ sub-002

Modality specific data

Files | motion

```
└─ sub-001/  
  └─ ses-01/  
    └─ eeg/  
    ...  
    └─ motion/
```

```
└─ sub-001_ses-01_task-<label>_tracksys-<label>_motion.tsv  
└─ sub-001_ses-01_task-<label>_tracksys-<label>_motion.json  
└─ sub-001_ses-01_task-<label>_channels.tsv
```

| | | | | | |
|-----------|-----------|-----------|-----------|-----------|-----------|
| 0,2634511 | 0,092295 | 0,0086682 | 0,9305117 | 0,690106 | 0,8098815 |
| 0,694520 | 0,1918243 | 0,8437273 | 0,3975710 | 0,885496 | 0,8952724 |
| 0,0766395 | 0,2587211 | 0,5434792 | 0,2822837 | 0,2789791 | 0,2326254 |
| 0,5779993 | 0,0456141 | 0,0490745 | 0,9408899 | 0,1533421 | 0,6683652 |
| 0,054556 | 0,7915927 | 0,5871733 | 0,4669577 | 0,9754468 | 0,0480541 |
| 0,966024 | 0,1962834 | 0,7114406 | 0,3389448 | 0,7194495 | 0,4384892 |
| 0,984172 | 0,5079461 | 0,1180168 | 0,7966978 | 0,1753768 | 0,4886533 |
| 0,9883907 | 0,1557346 | 0,8002013 | 0,6334882 | 0,7526906 | 0,8529441 |
| ... | ... | ... | ... | ... | ... |

```
{  
  "SamplingFrequency": 60,  
  "SamplingFrequencyEffective": 60.19,  
  "TaskName": "BIDS Motion fictive example",  
  "TrackingSystemName": "imu1",  
  "TaskDescription": "walking and talking",  
  "MotionChannelCount": 6,  
  "SubjectArtefactDescription": "n/a",  
  "TrackedPointsCount": 2,  
  "ACCELChannelCount": 3,  
  "GYROChannelCount": 3,  
  "Manufacturer": "BWSensing",  
  "ManufacturersModelName": "BW-imu600",  
}
```

| name | component | type | tracked_point | units | placement |
|----------------|-----------|-------|---------------|------------------|------------|
| imu1_rf_acc_x | x | ACCEL | rf | m/s ² | right_foot |
| imu1_rf_acc_y | y | ACCEL | rf | m/s ² | right_foot |
| imu1_rf_acc_z | z | ACCEL | rf | m/s ² | right_foot |
| imu1_rf_gyro_x | x | GYRO | rf | rad/s | right_foot |
| imu1_rf_gyro_y | y | GYRO | rf | rad/s | right_foot |
| imu1_rf_gyro_z | z | GYRO | rf | rad/s | right_foot |

Metadata | motion.json

REQUIRED

TaskName, SamplingFrequency

RECOMMENDED

TaskDescription, RecordingDuration,
MotionChannelCount, <type>ChannelCount,
SoftwareFilters, ...

OPTIONAL

Manufacturer, RecordingSoftware, ...

```
{  
  "SamplingFrequency": 60,  
  "TaskName": "BIDS Motion fictive example",  
  "TrackingSystemName": "IMU Right Hand",  
  "TaskDescription": "walking and talking",  
  "InstitutionAddress": "Fictive address",  
  "InstitutionName": "Fictive Institution",  
  "MotionChannelCount": 18,  
  "RecordingDuration": 4667.641106,  
  "SubjectArtefactDescription": "n/a",  
  "TrackedPointsCount": 2,  
  "ACCELChannelCount": 6,  
  "GYROChannelCount": 6,  
  "MAGNChannelCount": 6,  
  "Manufacturer": "BWSensing",  
  "ManufacturersModelName": "BW-IMU600",  
}
```

Metadata | channels.tsv

REQUIRED

name, component, type, tracked_point, units

RECOMMENDED

placement, reference_frame

OPTIONAL

status, status_description, sampling_frequency

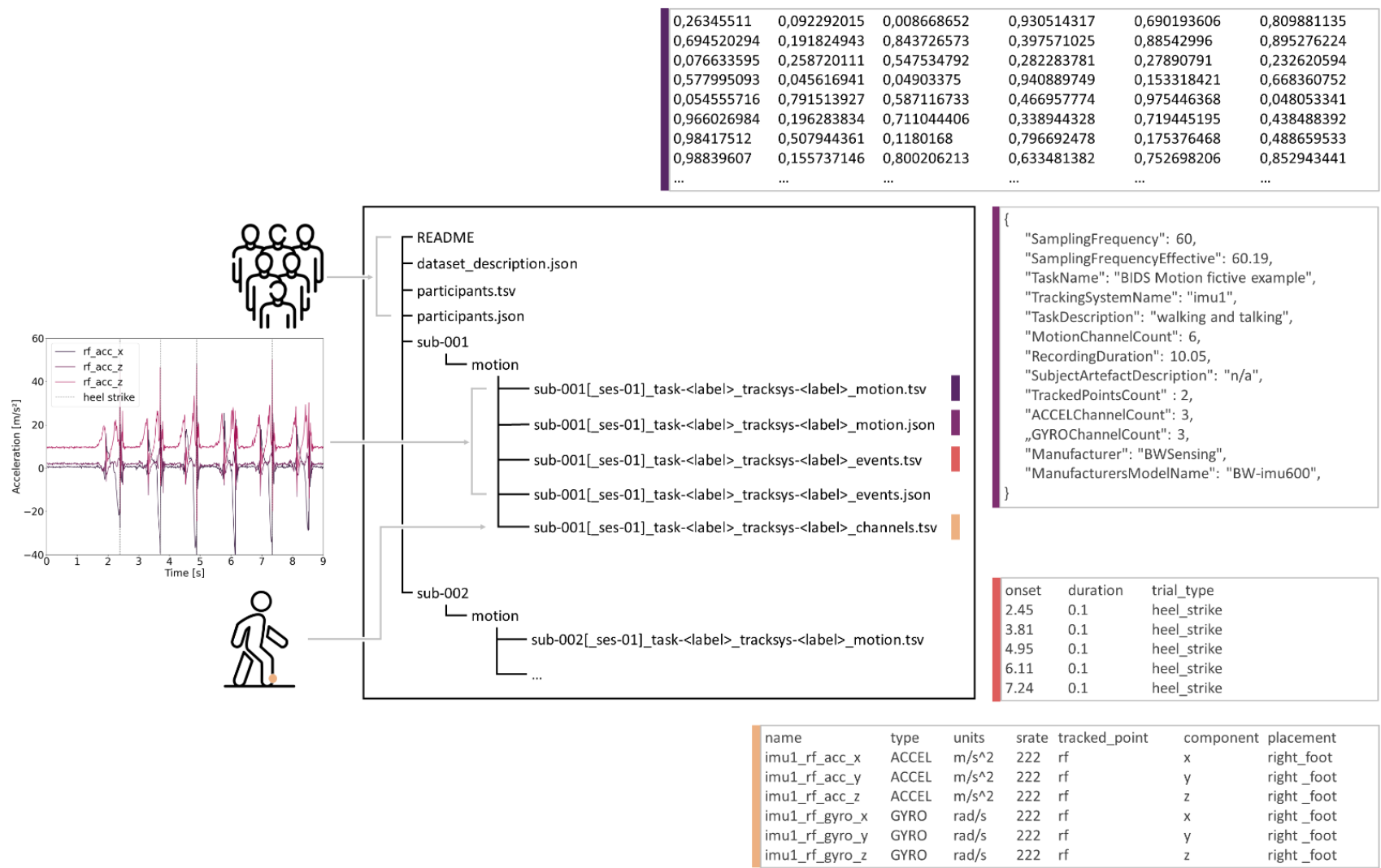
Restricted keyword list for channel type

Restricted keyword list for column `type` in alphabetic order. Note that upper-case is REQUIRED:

| Keyword | Description |
|----------|---|
| ACCEL | Accelerometer channel, one channel for each spatial axis. Column component for the axis MUST be added to the *_channels.tsv file (x, y, or z). |
| ANGACCEL | Angular acceleration channel, one channel for each spatial axis. Column component for the axis MUST be added to the *_channels.tsv file (x, y, or z). |
| GYRO | Gyrometer channel, one channel for each spatial axis. Column component for the axis MUST be added to the *_channels.tsv file (x, y, or z). |
| JNTANG | Joint angle channel between two fixed axis belonging to two bodyparts. Angle SHOULD be defined between proximal and distal bodypart in deg. |

| name | component | type | tracked_point | units | reference_frame |
|-----------|-----------|-------|---------------|-------|-----------------|
| t1_acc_x | x | ACCEL | LeftFoot | m/s^2 | global |
| t1_acc_y | y | ACCEL | LeftFoot | m/s^2 | global |
| t1_acc_z | z | ACCEL | LeftFoot | m/s^2 | global |
| t1_gyro_x | x | GYRO | LeftFoot | rad/s | global |
| t1_gyro_y | y | GYRO | LeftFoot | rad/s | global |
| t1_gyro_z | z | GYRO | LeftFoot | rad/s | global |
| ... | | | | | |

Overview | motion



Dataset info

Events



- README.md
- dataset_description.json
- participants.json
- participants.tsv
 - └ sub-001
 - ses-01
 - eeg/
 - motion/
 - ses-02
 - └ sub-002

Modality specific data

Task events

sub-<label>/
<data_type>/
 <matches>_events.tsv
 <matches>_events.json

```
{
  "channel": {
    "Description": "Channel(s) associated with the event"
  },
  "annots": {
    "LongName": "Annotations",
    "Description": "Annotations associated with channels indicated in the channel column.",
    "Levels": {
      "musc": "Muscle artifact. A very common, high frequency, sharp artifact that corresponds with agitation/nervousness in a patient."
    }
  }
}
```

| onset | duration | trial_type | response_time | stim_file | channel | annots |
|-------|----------|------------|---------------|-----------------|---------|--------|
| 1.23 | 0.65 | start | 1.435 | red_square.jpg | n/a | n/a |
| 5.65 | 0.65 | stop | 1.739 | blue_square.jpg | n/a | n/a |
| 12.1 | 2.35 | n/a | n/a | n/a | Cz | musc |

scans.tsv

```
sub-<label>/  
[ses-<label>/]  
sub-<label>[_ses-<label>]_scans.tsv
```

filename

```
eeg/sub-control01_task-nback.eeg  
eeg/sub-control01_task-walking.eeg  
motion/sub-control01_task-walking.tsv
```

acq_time

```
1877-06-15T13:45:30  
1877-06-15T13:55:33  
1877-06-15T13:56:27
```

*nback.eeg

*walking.eeg

*walking.tsv

| Column name | Requirement Level | Data type | Description |
|--------------------|-------------------|-----------|--|
| filename | REQUIRED | string | <p>Relative paths to files. There MUST be exactly one row for each file.</p> <p>Values in <code>filename</code> MUST be unique.</p> <p>This column must appear first in the file.</p> |
| acq_time | OPTIONAL | string | <p>Acquisition time refers to when the first data point in each run was acquired. Furthermore, if this header is provided, the acquisition times of all files from the same recording MUST be identical. Datetime format and their anonymization are described in Units.</p> <p>This column may appear anywhere in the file.</p> |
| Additional Columns | OPTIONAL | n/a | Additional columns are allowed. |

What is the Brain Imaging
Data Structure and **why you
should know about this!**

For research

- Validation and optimizing pipelines

- Building a sustainable research environment




Article | [Open access](#) | Published: 09 February 2023

EEG is better left alone

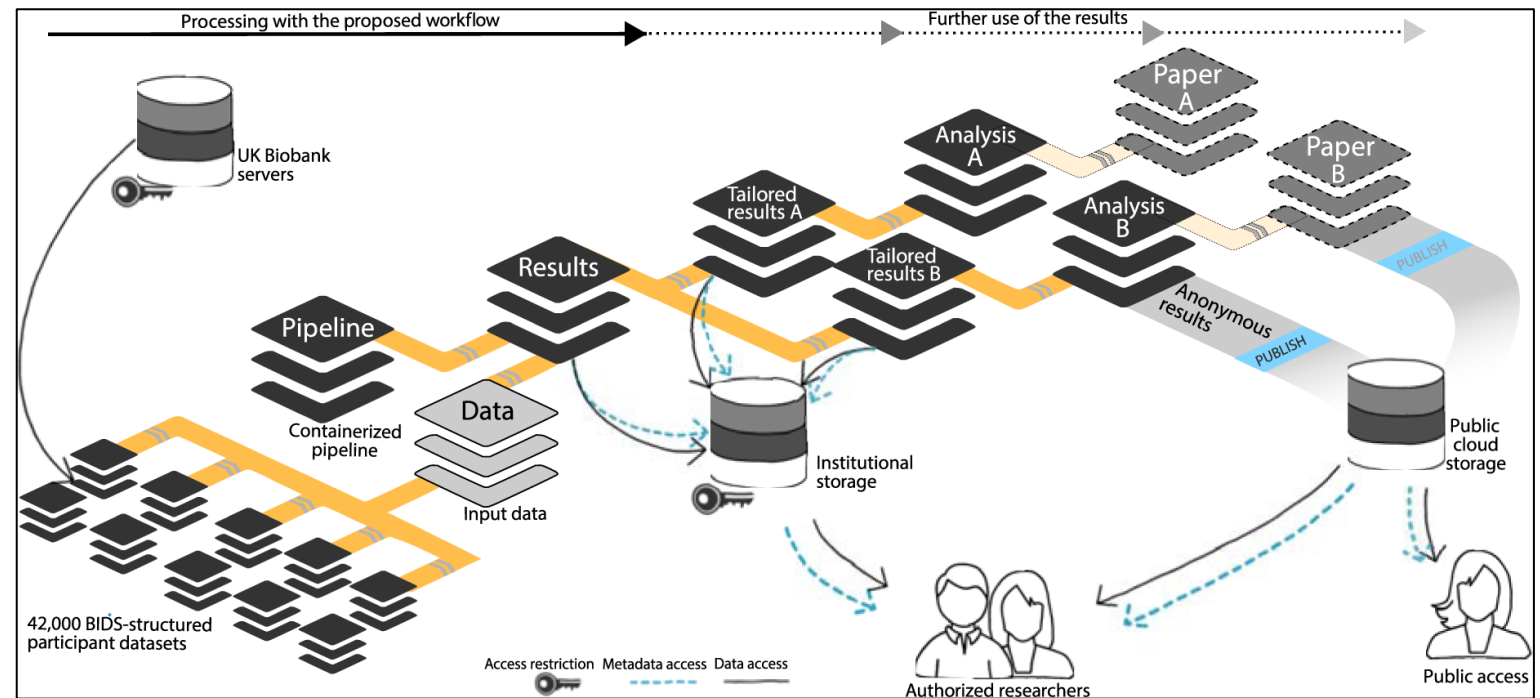
[Arnaud Delorme](#) 

[Scientific Reports](#) **13**, Article number

Preprocessing Choices for P3 Analyses with Mobile EEG: A Systematic Literature Review and Interactive Exploration

 Nadine S. J. Jacobsen,  Daniel Kristanto, Suong Welp, Yusuf Cosku Inceler,  Stefan Debener

doi: <https://doi.org/10.1101/2024.04.30.591874>

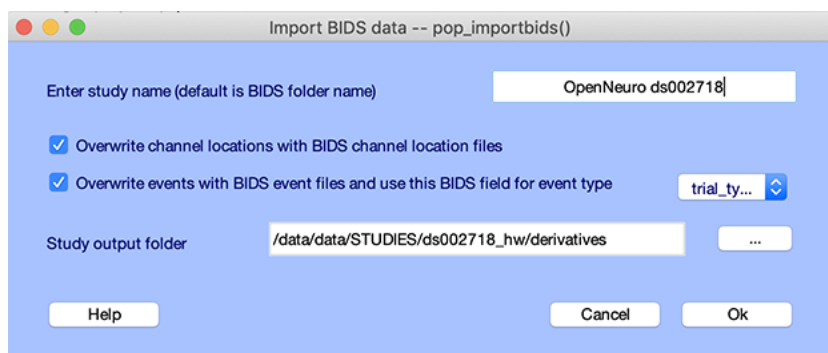


For you



Load public data from the BIDS format

- EEGLab, FieldTrip and MNE provide importers



Converting your data to BIDS

- Reuse data from your Lab and others
- Foster collaborations

```
sub = {'01', '02', '03', '04', '05', '06', '07', '08', '09', '10'};
```

```
% for subject 3 the age is unknown, for subject 2 the sex is not specified
```

```
age = [11 96 nan 77 82 87 18 40 26 80];
```

```
sex = {'f' [] 'f' 'f' 'f' 'm' 'm' 'm' 'm' 'm'};
```

```
for subindx=1:numel(sub)
```

```
    cfg = [];
```

```
    cfg.datatype = 'eeg';
```

```
    % Load your data
```

```
    cfg.dataset = sub{subindx} + '.edf';
```

```
    % specify the output directory
```

```
    cfg.bidsroot = 'bids';
```

```
    cfg.sub = sub{subindx};
```

```
    % specify the information for the participants.tsv file
```

```
    % this is optional, you can also pass other pieces of info
```

```
    cfg.participants.age = age(subindx);
```

```
    cfg.participants.sex = sex{subindx};
```

```
    % provide the mnemonic and long description of the task
```

```
    cfg.TaskName = 'changedetection';
```

```
    cfg.TaskDescription = 'Subjects were responding as fast as possible upon a  
                           change in a visually presented stimulus.';
```

```
    % these are EEG specific
```

```
    cfg.eeg.PowerLineFrequency = 50; % since recorded in the EU
```

```
    cfg.eeg.EEGReference = 'M1'; % left mastoid
```

```
    data2bids(cfg);
```

```
end
```

BIDS Starterpack

Website

[<https://bids-specification.readthedocs.io/en/stable/>]

Validator

[<https://bids-standard.github.io/bids-validator/>]

BIDS Validator v1.14.6

Select a [BIDS dataset](#) to validate

Brain Imaging Data Structure v1.9.0

Search

GitHub v1.9.0 255 153

Brain Imaging Data Structure v1.9.0

The BIDS Specification

Introduction

Common principles

Modality agnostic files

Modality specific files >

Derivatives >

Longitudinal and multi-site studies

Glossary

BIDS Extension Proposals

Appendix >

Changelog

The BIDS Starter Kit

Website

Tutorials

GitHub repository

The Brain Imaging Data Structure

The Brain Imaging Data Structure (BIDS) is a simple and intuitive way to organize and describe data.

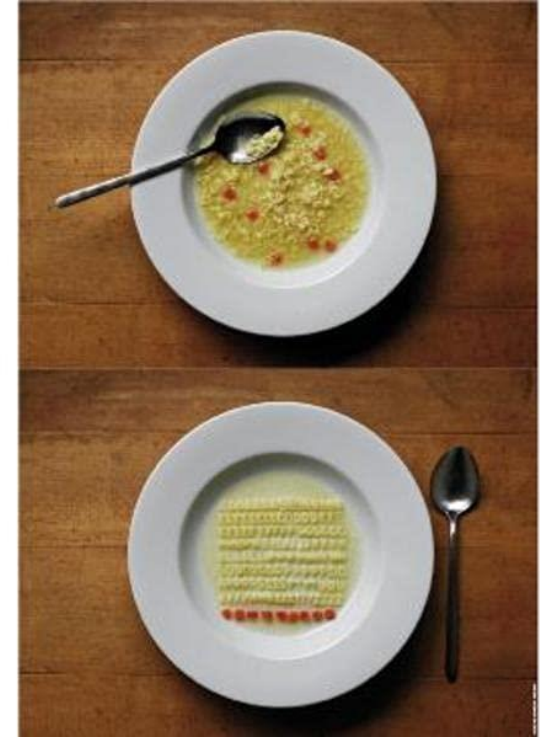
This document defines the BIDS specification, which provides many details to help implement the standard. It includes the core specification as well as many extensions to specific brain imaging modalities, and increasingly also to other kinds of data.

If BIDS is new to you, and you would like to learn more about how to adapt your own datasets to match the BIDS specification, we recommend exploring the [BIDS Starter Kit](#). Alternatively, to get started please read [the introduction to the specification](#).

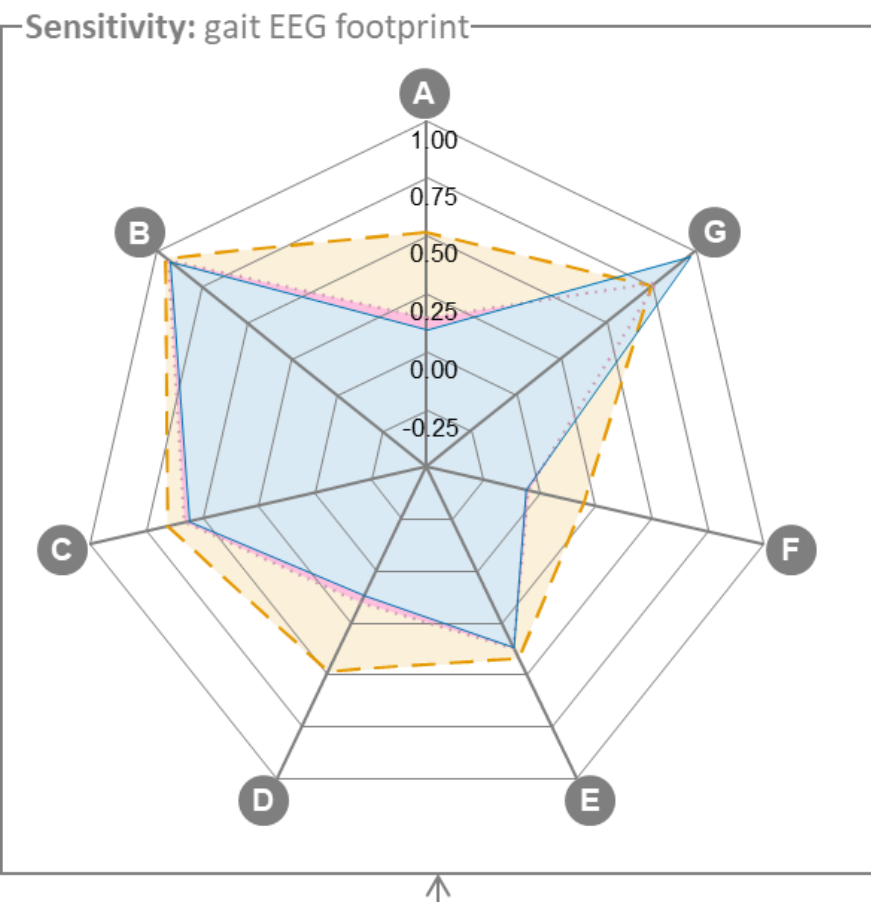
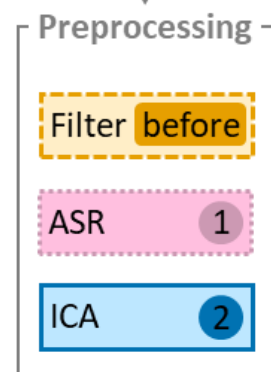
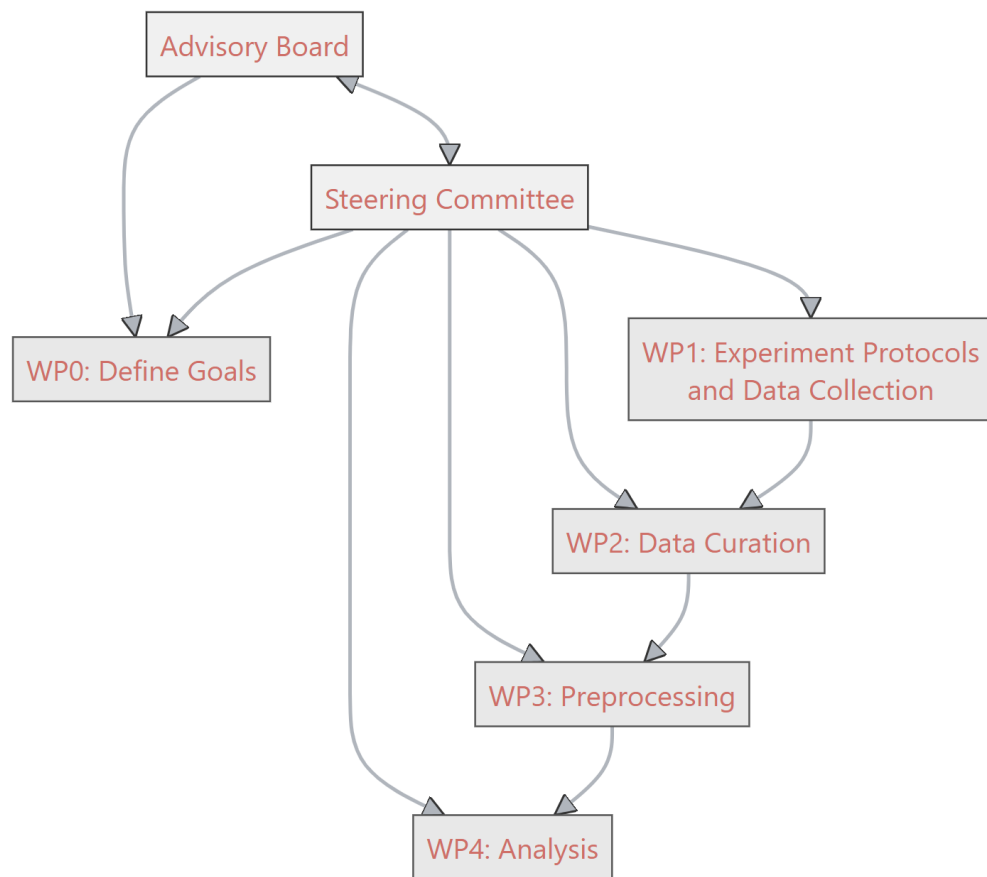
For an overview of the BIDS ecosystem, visit the [BIDS homepage](#). The entire specification can also be [downloaded as PDF](#).

Copyright © BIDS Contributors - CC BY 4.0

Conclusions



Outlook: EEGManySteps



Thank you for listening carefully



Thanks to **Sein Jeung** for pushing this to completion over the past three years
Thanks to all of the **BIDS Maintainers and Devs** who help to implement this
Thanks to my **working group**, who have given me the freedom to work in this
Thanks to all **participants** who ever provided data which is now in BIDS