Research Data Management in Neuroscience

An introduction to BIDS

```
The Brain Imaging Data Structure

Michael Sonntag

Department Biologie II
Ludwig-Maximilians-Universität München

Friday, 12 June, 2020
```


Take home message

BIDS

- can help you organise (imaging) data
- exposes you to a community standard of data organisation
- exposes you to a standard of project and metadata organisation

BIDS background

- inspired by addressing problems at openNeuro.org
- developed in Stanford at the Poldrack lab in 2016

> Gorgolewski, K.J., Auer, T., Calhoun, V.D., Craddock, R.C., Das, S., Duff, E.P., Flandin, G., Ghosh, S.S., Glatard, T., Halchenko, Y.O., Handwerker, D.A., Hanke, M., Keator, D., Li, X., Michael, Z., Maumet, C., Nichols, B.N., Nichols, T.E., Pellman, J., Poline, J.-B., Rokem, A., Schaefer, G., Sochat, V., Triplett, W., Turner, J.A., Varoquaux, G., Poldrack, R.A., 2016. The brain imaging data structure, a format for organizing and describing outputs of neuroimaging experiments. Sci Data 3, 160044.

Research Resource Identifier RRID:SCR_016124

- orignally aimed to document MRI and fMRI data
- data structure specification to consistently organize and document neuroimaging and connected behavioral data
- out of this developed the larger BIDS project: https://bids.neuroimaging.io (https://bids.neuroimaging.io)

The BIDS specification

- BIDS provides a specification but is not a standard
- it should be viewed as a best practice in project structure and documentation

To this end the BIDS standard specifies

- the naming convention for files and directories
- which file formats are to be used for a use case (i.e. Nifti, json, tsv)
- core metadata and how they are to be stored e.g. about participants, stimuli and key recording settings

Besides the imaging aspect it tries to cover

- behavior
- physiology

The BIDS specification

So far full BIDS specifications exist for

- MRI (2016)
- fMRI (2016)
- MEG (2018)
- EEG (2019)
- iEEG (2019)

Specification extensions e.g. for PET or CT are currently being developed by the community.

Introduction to the standard specification

The BIDS structure

BIDS specifies

- folder structures
- supported file types for different types of (neuroimaging) data
- file naming
- partially file content

The specification can be found at https://bids-specification.readthedocs.io/en/stable (https://bids-specification.readthedocs.io/en/stable).

BIDS file type support

- `.json` files to document metadata
- `.tsv` files containing tab separated tabular metadata - no CSV, no excel, only true tabs, no spacing
- raw data files specific to the modality that the project contains
  - e.g. nii.gz files for an anatomical MRI project
 - only NIFTI files are supported

BIDS general folder structure

```
project                    exampleProjectName
└── subject                └── sub-subject_id_01
    └── session                └── ses-session_number_01
        └── datatype               └── anat
        └── datatype               └── func
```

BIDS folder names and contrstraints

- `project` ... can have any name but should be descriptive
- `subject` ... sub-<participant label>
  - Label has to be specific for each subject
  - Only one folder per subject per dataset
- `session` ... sub-<session label>
  - Each folder represents a recording session
  - If required use multiple sessions per subject
  - The session label has to be unique per subject
- `datatype` ... `func`, `dwi`, `fmap`, `anat`, `meg`, `eeg`, `ieeg`, `beh`
  - defines the types of data contained in this dataset

BIDS general folder structure

```
project                    exampleProjectName
└── subject                └── sub-subject_id_01
    └── session                └── ses-session_number_01
        └── datatype               └── anat
        └── datatype               └── func
```
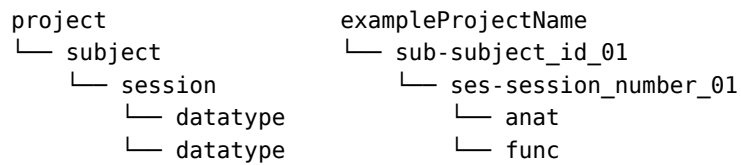
BIDS datatypes

- `func` ... functional MRI data
- `dwi` ... diffusion Imaging Data
- `fmap` ... fieldmap MRI data
- `anat` ... anatomical MRI data
- `meg` ... MEG data
- `eeg` ... EEG Data
- `ieeg` ... intracranial EEG data
- `beh` ... behavior

Folders of these datatypes allow only specific files that are named according to the BIDS specificiation.

BIDS general folder structure

File naming contstraints

Metadata and data file names depend on the project type and the folder names!

Anatomical MRI data example: `anat`

Folder structure and naming constraints `./myProject/sub-01/ses-01/anat/`

Data file naming constraints in an anatomical MRI data project:

- `sub-<>_ses-<>_T1w.nii.gz`

Metadata file naming constraints:

- `sub-<>_ses-<>_T1w.json`

Other files are not allowed in an `anat` folder.

BIDS general folder structure - full example

```
ds001
├── dataset_description.json
├── participants.tsv
├── sub-01
│   ├── anat
│   │   ├── sub-01_inplaneT2.nii.gz
│   │   └── sub-01_T1w.nii.gz
│   └── func
│       ├── sub-01_task-balloonanalogrisktask_run-01_bold.nii.gz
│       ├── sub-01_task-balloonanalogrisktask_run-01_events.tsv
│       ├── sub-01_task-balloonanalogrisktask_run-02_bold.nii.gz
│       ├── sub-01_task-balloonanalogrisktask_run-02_events.tsv
├── sub-02
│   ├── anat
│   │   ├── sub-02_inplaneT2.nii.gz
│   │   └── sub-02_T1w.nii.gz
│   └── func
│       ├── sub-02_task-balloonanalogrisktask_run-01_bold.nii.gz
│       ├── sub-02_task-balloonanalogrisktask_run-01_events.tsv
│       ├── sub-02_task-balloonanalogrisktask_run-02_bold.nii.gz
│       ├── sub-02_task-balloonanalogrisktask_run-02_events.tsv
...
└── task-balloonanalogrisktask_bold.json
```

BIDS general folder structure - BIDS is rigid

- folder structure and naming scheme has to be followed
- empty or additional files as well as unsupported file types are not allowed

Adding additional files

The root of the project folder may (and should) contain the following files

- README
- dataset_description.json
- participants.tsv

Dealing with unsupported files

- keep unsupported files one level above the project root
- add non-BIDS files to `.bidsignore` at the project root; works like `.gitignore`.

```
*_not_bids.txt
extra_data/
```

BIDS validation

Core of BIDS is a validation service

- needs to be run on a regular basis to ensure adherence to specification
- online service at https://bids-standard.github.io/bids-validator (https://bids-standard.github.io/bids-validator)

- local service installation at https://github.com/bids-standard/bids-validator (https://github.com/bids-standard/bids-validator)
  - nodejs (full functionality, commandline tool)
  - Python (reduced functionality)
  - docker

BIDS usage example

We will now use the online validator to build and troubleshoot a BIDS project from scratch.

https://bids-standard.github.io/bids-validator (https://bids-standard.github.io/bids-validator)

Find step by step example directories in the RDM course folder on gin: https://gin.g-node.org/RDMcourse2020/Lectures/Lecture04/BIDS_validation_examples (https://gin.g-node.org/RDMcourse2020/Lectures/Lecture04/BIDS_validation_examples)

The folder contains four BIDS projects with various validation issues:

```
01_empty_example/
02_invalid_structure/
03_invalid_file_annotation/
04_invalid_additional_file/
```

The folders can be uploaded to the validator and will return the individual issues.

BIDS troubleshooting: use the specification

- BIDS is available for fMRI, MRI, EEG, iEEG and other data sources
- the exact specification, allowed structure, naming and file format varies
- use the specification for all details to get to a valid BIDS structure

https://bids-specification.readthedocs.io/en/stable (https://bids-specification.readthedocs.io/en/stable)

- use the specification to collect and document metadata

BIDS specifications in the making - get involved

Besides the published supported data many more are currently in development

e.g. BIDS for PET is close to finishing: https://docs.google.com/document /d/1mqMLnxVdLwZjDd4ZiWFqjEAmOmfcModA_R535v3eQs0/edit (https://docs.google.com/document /d/1mqMLnxVdLwZjDd4ZiWFqjEAmOmfcModA_R535v3eQs0/edit)

Everyone can look up the status of a project and also contribute: https://bids.neuroimaging.io/get_involved.html (https://bids.neuroimaging.io/get_involved.html)

BIDS converters, tools and apps

List of tools

- https://bids.neuroimaging.io/benefits.html (https://bids.neuroimaging.io/benefits.html)

Example tool: Raw data to BIDS converter

- https://github.com/Donders-Institute/bidscoin (https://github.com/Donders-Institute/bidscoin)

BIDS apps - applications that work with BIDS datasets

- https://bids-apps.neuroimaging.io/about (https://bids-apps.neuroimaging.io/about)

Example app

- https://github.com/poldracklab/fmriprep (https://github.com/poldracklab/fmriprep)

Linklist

BIDS home page

- https://bids.neuroimaging.io (https://bids.neuroimaging.io)

BIDS specification

- https://bids-specification.readthedocs.io/en/stable/ (https://bids-specification.readthedocs.io/en/stable/)

BIDS validator

- https://bids-standard.github.io/bids-validator/ (https://bids-standard.github.io/bids-validator/)
- https://github.com/bids-standard/bids-validator (https://github.com/bids-standard/bids-validator)

Introductions and examples

- https://github.com/bids-standard/bids-starter-kit (https://github.com/bids-standard/bids-starter-kit)
- https://github.com/bids-standard/bids-starter-kit/wiki/Tutorials (https://github.com/bids-standard/bids-starter-kit/wiki /Tutorials)
- https://github.com/bids-standard/bids-examples (https://github.com/bids-standard/bids-examples)

BIDS papers

BIDS https://doi.org/10.1038/sdata.2016.44 (https://doi.org/10.1038/sdata.2016.44)

EEG-BIDS https://doi.org/10.1038/s41597-019-0104-8 (https://doi.org/10.1038/s41597-019-0104-8)

iEEG BIDS https://doi.org/10.1038/s41597-019-0105-7 (https://doi.org/10.1038/s41597-019-0105-7)

MEG BIDS https://doi.org/10.1038/sdata.2018.110 (https://doi.org/10.1038/sdata.2018.110)

BIDS apps https://doi.org/10.1371/journal.pcbi.1005209 (https://doi.org/10.1371/journal.pcbi.1005209)

Assignment

- read through https://github.com/bids-standard/bids-starter-kit (https://github.com/bids-standard/bids-starter-kit)
- try to map your data to the BIDS structure
- if you have problems check
  - the specification: https://bids-specification.readthedocs.io/en/stable/ (https://bids-specification.readthedocs.io/en/stable/)
  - the examples page: https://github.com/bids-standard/bids-examples (https://github.com/bids-standard/bids-examples)
- make sure your example is valid using the online validator
- read through the specification for your dataset and try to find some metadata, validate again