



STANFORD  
UNIVERSITY

Stanford Center For  
Reproducible Neuroscience

# Brain Imaging Data Structure (BIDS)ify your data: help/info session

June 29, 2018  
Blg. 420 Rm. 419  
Franklin Feingold  
ffein@stanford.edu

# What is BIDS?

- No organizational standard in Neuroimaging community
- Simple and intuitive way to organize and describe neuroimaging and behavioral data
- Developed with input from experts and the broader neuroimaging community
- Covers many common experimental methodologies
- Flexible and easy to adopt
- BIDS Validator to confirm the data structure is correct

# BIDS Benefits

## Data reusability

- Within the same lab, PI asks a new postdoc to ask a different question from a dataset that a previous graduate student analyzed. With BIDS you would know how to interact with the data and filenames are human-readable.

## Data sharing

- Standardization with documentation so other researchers will understand the data organization and more easily use the data

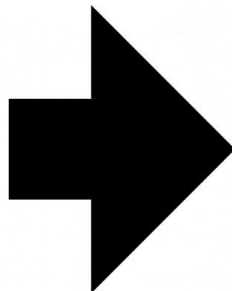
## Study reproducibility

- Standard organization will streamline the process to replicate the results from studies and write analysis pipelines that can be used across different datasets

# What do we cover?

- Specification covers the raw data organization for anatomical, functional, diffusion MRI scans as well as MEG scans
- Common neuroimaging file types and general formats supported (i.e. Nifti, json, tsv, bvec, bval)
- Full current specification on our website ([bids.neuroimaging.io](https://bids.neuroimaging.io))
- Current Specification version 1.1.1: [http://bids.neuroimaging.io/bids\\_spec.pdf](http://bids.neuroimaging.io/bids_spec.pdf)
- BIDS is expanding to more modalities, processed data, and describing models

dicomdir/  
1208200617178\_22/  
1208200617178\_22\_8973.dcm  
1208200617178\_22\_8943.dcm  
1208200617178\_22\_2973.dcm  
1208200617178\_22\_8923.dcm  
1208200617178\_22\_4473.dcm  
1208200617178\_22\_8783.dcm  
1208200617178\_22\_7328.dcm  
1208200617178\_22\_9264.dcm  
1208200617178\_22\_9967.dcm  
1208200617178\_22\_3894.dcm  
1208200617178\_22\_3899.dcm  
1208200617178\_23/  
1208200617178\_24/  
1208200617178\_25/



my\_dataset/  
participants.tsv  
sub-01/  
anat/  
sub-01\_T1w.nii.gz  
func/  
sub-01\_task-rest\_bold.nii.gz  
sub-01\_task-rest\_bold.json  
dwi/  
sub-01\_dwi.nii.gz  
sub-01\_dwi.json  
sub-01\_dwi.bval  
sub-01\_dwi.bvec  
sub-02/  
sub-03/  
sub-04/

participant_id	age	sex	group
sub-01	34	M	control
sub-02	12	F	control
sub-03	33	F	patient
sub-04	23	M	patient

```
my_dataset/  
  participants.tsv  
  sub-01/  
    anat/  
      sub-01_T1w.nii.gz
```

# Anatomical Template

Template:

sub-<participant\_label>/[ses-<session\_label>/]

**anat/**

**sub**-<participant\_label>[\_ses-<session\_label>][\_acq-<label>][\_ce-<label>][\_rec-<label>][\_run-<index>]\_<modality\_label>.nii[.gz]









Example: sub-01/ses-02/anat/sub-01\_ses-02\_run-2\_T1w.nii.gz

Page 20 of the BIDS Specification 1.1.1

The bracketed pieces are optional

```
my_dataset/  
  participants.tsv  
  sub-01/  
    anat/  
      sub-01_T1w.nii.gz
```



 my\_dataset/  
     participants.tsv  
     sub-01/  
         anat/  
             sub-01\_T1w.nii.gz  
         func/  
             sub-01\_task-rest\_bold.nii.gz  
             sub-01\_task-rest\_bold.json

# Functional Template

Template:

sub-<participant\_label>/[ses-<session\_label>/]

**func/**

**sub**-<participant\_label>[\_ses-<session\_label>]\_**task**-<task\_label>[\_acq-<label>][\_rec-<label>][\_run-<index>][\_echo-<index>]\_**bold.nii**[.gz]

Example: sub-control01/func/sub-control01\_task-nback\_bold.nii.gz

Page 21 of the BIDS Specification 1.1.1

Covers both rest and task. Stored in the 'task\_label'

Example: sub-control01/func/sub-control01\_task-nback\_sbref.nii.gz









# Json file example

```
sub-control01/  
  func/  
    sub-control01_task-nback_bold.json
```

This example can be found on page 24  
of the BIDS Specification 1.1.1

```
{  
  ➡ "TaskName": "N Back",  
  ➡ "RepetitionTime": 0.8,  
    "EchoTime": 0.03,  
    "FlipAngle": 78,  
    "SliceTiming": [0.0, 0.2, 0.4, 0.6, 0.0, 0.2, 0.4, 0.6, 0.0, 0.2, 0.4, 0.6,  
0.0, 0.2, 0.4, 0.6],  
    "MultibandAccelerationFactor": 4,  
    "ParallelReductionFactorInPlane": 2,  
    "PhaseEncodingDirection": "j",  
    "InstitutionName": "Stanford University",  
    "InstitutionAddress": "450 Serra Mall, Stanford, CA 94305-2004, USA",  
    "DeviceSerialNumber": "11035"  
}
```

Common metadata fields on page 16 of the BIDS  
Specification 1.1.1

 my\_dataset/  
     participants.tsv  
     sub-01/  
         anat/  
             sub-01\_T1w.nii.gz  
         func/  
             sub-01\_task-rest\_bold.nii.gz  
             sub-01\_task-rest\_bold.json

- my\_dataset/
  - participants.tsv
  - sub-01/
    - anat/
      - sub-01\_T1w.nii.gz
    - func/
      - sub-01\_task-rest\_bold.nii.gz
      - sub-01\_task-rest\_bold.json
    - dwi/
      - sub-01\_dwi.nii.gz
      - sub-01\_dwi.json
      - sub-01\_dwi.bval
      - sub-01\_dwi.bvec

# Diffusion Template

Template:

sub-<participant\_label>/[ses-<session\_label>/]

**dwi/**

**sub-<participant\_label>[\_ses-<session\_label>][\_acq-<label>][\_run-<index>]\_dwi.nii[.gz]**

**sub-<participant\_label>[\_ses-<session\_label>][\_acq-<label>][\_run-<index>]\_dwi.bval**

**sub-<participant\_label>[\_ses-<session\_label>][\_acq-<label>][\_run-<index>]\_dwi.bvec**

**sub-<participant\_label>[\_ses-<session\_label>][\_acq-<label>][\_run-<index>]\_dwi.json**

Example: sub-003/ses-005/dwi/sub-003\_ses-005\_run-009\_dwi.nii.gz

Page 24 of the BIDS Specification 1.1.1

# Field Maps Template

Phase and Magnitude maps are accepted modality types

Template:

sub-<participant\_label>/[ses-<session\_label>/]

**fmap/**

**sub-<label>[\_ses-<session\_label>][\_acq-<label>][\_run-<run\_index>]\_phasediff.nii.gz]**

**sub-<label>[\_ses-<session\_label>][\_acq-<label>][\_run-<run\_index>]\_phasediff.json]**

**sub-<label>[\_ses-<session\_label>][\_acq-<label>][\_run-<run\_index>]\_magnitude1.nii.gz]**

Example: sub-9/fmap/sub-9\_phasediff.nii.gz

More examples can found on page 26 of the BIDS Specification 1.1.1

# Magnetoencephalography (MEG) Template

Template:

sub-<participant\_label>/

[ses-<label>]/

**meg/**

**sub-**<participant\_label>[\_ses-<label>]**\_task-**<task\_label>[\_acq-<label>][\_run-<index>][\_proc-<label>]**\_meg.**<manufacturer\_specific\_extension>

[**sub-**<participant\_label>[\_ses-<label>]**\_task-**<task\_label>[\_acq-<label>][\_run-<index>][\_proc-<label>]**\_meg.json**]

Example: sub-01/meg/sub-01\_task-rest\_run-02\_meg.fif

Page 28 of the BIDS Specification 1.1.1

Required fields within the json file: TaskName, SamplingFrequency, PowerLineFrequency, DewarPosition, SoftwareFilters, DigitizedLandmarks, DigitizedHeadPoints



# Additional files

dataset\_description.json - Required information includes: Name and BIDSVersion

Further information - page 14 of BIDS Specification 1.1.1. Examples on page 15.

events.tsv - describes the timing and other relevant information of the events recorded during the scan. Required columns include: onset, duration

Further information - page 41 of BIDS Specification 1.1.1. Examples on page 42.

These files will be placed at the same level as the subject folders

# Tutorials

A detailed introductory step-by-step walkthrough -

<https://stanford.io/2tBmexW>

Automating the introductory walkthrough -

<https://stanford.io/2KxIBOo>

Using an automated BIDS converter (HeuDiConv) -

<https://stanford.io/2yRjwcC>

# Examples

Stripped down datasets - <https://github.com/INCF/BIDS-examples>

OpenNeuro - <https://openneuro.org/>

# Questions about your particular dataset?

Please post your questions on neurostars.org with the BIDS tag

To find these questions - <https://neurostars.org/tags/bids>

# How do you know you if you converted correctly?

We have a web-based validator to verify that the specification is being followed

<http://incf.github.io/bids-validator/>

Privacy point: We do not have access to your dataset. The validator is checking the files you pointed to and confirming the data structure are within the specification on your local computer. There is **no** data sharing or uploading taking place. This is simply a tool to perform validation.



# How can I join the BIDS community?

We have a google group that holds all of our discussions related to BIDS specification development

[goo.gl/wR9Qu9](https://goo.gl/wR9Qu9)

BIDS events files for different modalities.

4 posts by 3 authors



Horea Christian

May 6



I notice that the eventfile `_events` suffix is treated like a hierarchical equivalent of contrast specifications e.g. `_cbv` or `_bold`.

Sadly, this is not always representative of the field hierarchy, e.g.:

In one session a `_cbv` and `_bold` scan may be acquired, and they might have varying events files.

Any suggestions on how this might best be handled? Is the concatenation `_cbv_events` permitted?

I am thinking this might indicate a larger issue with the implicit-field suffixes (`_bold`, `_bval`, `_events`) - maybe some of them, such as the contrasts, should be migrated to their own tag, e.g. `_con-cbv`.

[Click here to Reply](#)



Tal Yarkoni

May 6



Can you provide a more detailed example of a case where you'd run into problems? In the case you describe, couldn't you resolve the problem by, say, using the "task" keyword to specify different tasks for `cbv` and `bold`? E.g., `task-mytaskcbv_events.tsv` and `task-mytaskbold_events.tsv`? I don't see why the structure of `events.tsv` files has to mirror the rest of the hierarchy; minimally, there only needs to be *some* distinguishing feature somewhere that allows you to determine which file you want.

- show quoted text -  
- show quoted text -

---

You received this message because you are subscribed to the Google Groups "bids-discussion" group.

To unsubscribe from this group and stop receiving emails from it, send an email to [bids-discussi...@googlegroups.com](mailto:bids-discussi...@googlegroups.com).

To post to this group, send email to [bids-di...@googlegroups.com](mailto:bids-di...@googlegroups.com).

To view this discussion on the web visit <https://groups.google.com/d/msgid/bids-discussion/9bdb2079-a393-49a7-837e-2c9c310cebc1%40googlegroups.com>.

For more options, visit <https://groups.google.com/d/optout>.



Tal Yarkoni

May 6



On re-read, my response may have sounded a bit flippant. To clarify, I'm not saying it would be a bad idea in principle to have `_events.tsv` respect the rest of the naming conventions; I just think this is a pretty clear case where the 80/20 rule favors keeping things as-is. The proportion of datasets that need to use `_events.tsv` to encode task events from multiple kinds of scans is probably well under 10%, and in those cases, I think it makes sense to require a very small amount of extra work from users to encode salient task differences in another identifier (most likely "task"), rather than adding additional complexity to the spec (and more cognitive overhead for users).

- show quoted text -

# Want to contribute?

We are always welcome to new ideas and further expansion of the specification

This google doc will explain the process for contributing to BIDS -

<https://goo.gl/Pfc69o>

Page 9 of BIDS Specification 1.1.1. lists the extension projects currently underway



# Want to contribute?

## Minimally processed and resampled volumes

Template:

```
<pipeline_name>/
sub-<participant_label>/
  func|anat|dwi/
    <source_file>_space-<space>[_res-<XxYxZ>][_variant-<label>]_preproc.nii.gz]
  c.nii.gz]

for example:

sub-001/
```

```
func/
sub-001_task-rest_run-1_bold_space-MNI305_preproc.nii.gz
```

Minimally preprocessed, coregistered and resampled (interpolated) volumes. "space" denotes the last template a given volume was coregistered to. It can be one of the values denoted in Table targets/spaces. "res" denotes the resolution to which the volumes have been resliced to (in mm) for example: "1x1x1". If slice time correction changes the scan onset to the middle of the scan, adjusted event timings must be provided.

Multiple different versions of preprocessing can be stored for the same source data. To distinguish them from each other the "variant" keyword can be used. Details of preprocessing performed for each variant should be included in the pipeline documentation.

**The standard does not define what "minimally preprocessed". This will depend on each preprocessing pipeline.**

Label name	Description
orig	A (potentially unique) per-image space. Useful for describing the source of transforms from an input



Chris Markiewicz

5:43 AM May 11

Resolve

I've asked at Singularity, but I don't think we should assume the hash will always be MD5. SingularityHub has a field called "version", which contains a hash digest, so perhaps we should make this SingularityContainerVersion?



Cyril Pernet

4:10 AM Mar 15

Resolve

or SourceDatasetsDOIs



Chris Markiewicz

4:58 AM May 11

What about moving to SourceDatasets, where the values must be URLs or DOIs?



Jonathan R. Wi...

Aug 4, 2017

Resolve

What if the space is native (original for the underlying data) space? For example, if the volume was only resampled to 1x1x1?



Chris Markiewicz

Sep 21, 2017

This is a good point. I think space should be optional. This is consistent with FMRIPREP's anat/sub-<participant\_label>\_T1w\_preproc.nii.gz. There the space is implicitly T1w.



Martin Felipe P...

Aug 4, 2018

Resolve

One question I still have is about the file parsing convention. Does keeping

# BIDS Tools

BIDS Apps - <http://bids-apps.neuroimaging.io/>

- Collection of application developed by the community and easy to use once data is in the BIDS format

## Available BIDS Apps

BIDS-Apps/example	version 0.0.7	open bug issues 0
BIDS-Apps/freesurfer	version v6.0.1-4	open bug issues 0
BIDS-Apps/ndmg	version v0.1.0	open bug issues 0
BIDS-Apps/BROCCOLI	version v1.0.1	open bug issues 1
BIDS-Apps/FibreDensityAndCrosssection	version v0.0.1	open bug issues 0
BIDS-Apps/SPM	version v0.0.14	open bug issues 1
poldracklab/mriqc	version 0.11.0	open bug issues 25
BIDS-Apps/QAP	Image not found	open bug issues 0
BIDS-Apps/CPAC	version v1.0.1a_22	open bug issues 0
BIDS-Apps/hyperalignment	Image not found	open bug issues 0
BIDS-Apps/mindboggle	version 0.0.4-1	open bug issues 2
BIDS-Apps/MRtrix3_connectome	version 0.3.0	open bug issues 0
BIDS-Apps/rs_signal_extract	version 0.1	open bug issues 0
BIDS-Apps/aa	version enh_vario...	open bug issues 0
BIDS-Apps/niak	version latest	open bug issues 1
BIDS-Apps/oppni	version v0.7.0-1	open bug issues 1
poldracklab/fmriprip	version 1.1.1	open bug issues 11

# BIDS Tools

BIDS Apps - <http://bids-apps.neuroimaging.io/>

- Collection of application developed by the community and easy to use once data is in the BIDS format

MRQC - <http://mrqc.org/>

- MRQC provides reports to more systemically evaluate the scan quality

fMRIPrep - <http://fmripred.org/>

- fMRIPrep is a preprocessing pipeline that outputs reports and data prepared for group level analysis

# Acknowledgements



Russ Poldrack

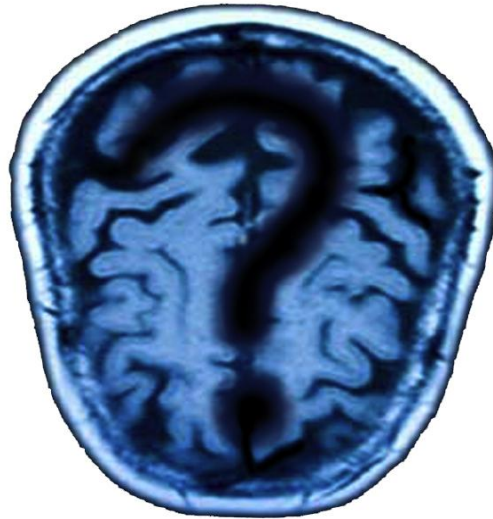


Chris Gorgolewski



Poldrack lab

NIH for their funding support



# Let's start BIDSifying!

[ffein@stanford.edu](mailto:ffein@stanford.edu)