Introduction to BIDS

Shawn Rhoads
Georgetown University
Brainhack DC 2020

“A simple and intuitive way to organize and describe your neuroimaging and behavioral data.”
https://bids.neuroimaging.io/index.html

Gorgolewski, et al. (2016)

Why use BIDS?

- A standardized, community-driven way to organize and share data!

Why use BIDS?

- A standardized, community-driven way to organize and share data!
- Saves time!

Why use BIDS?

- A standardized, community-driven way to organize and share data!
- Saves time!
  - Multiple people on one project?
  - Longitudinal study?
  - Want to reuse data for a new project or grant?

Why use BIDS?

- A standardized, community-driven way to organize and share data!
- Saves time!
  - Multiple people on one project?
  - Longitudinal study?
  - Want to reuse data for a new project or grant?
- Easy to understand!

Why use BIDS?

- A standardized, community-driven way to organize and share data!
- Saves time!
  - Multiple people on one project?
  - Longitudinal study?
  - Want to reuse data for a new project or grant?
- Easy to understand!
- Growing number of applications built around BIDS!

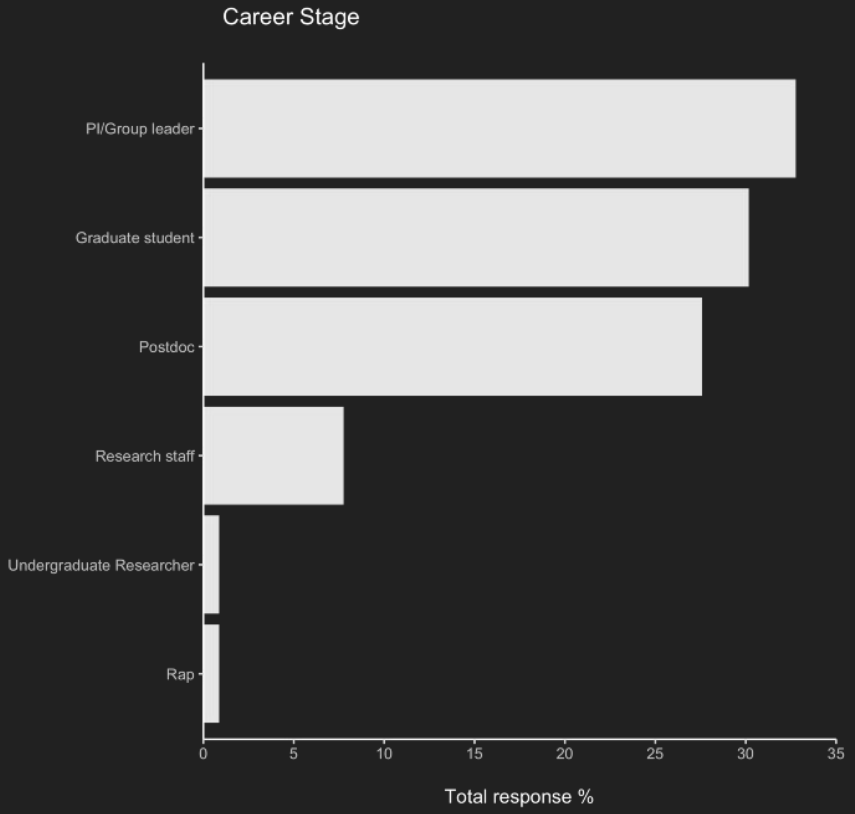
Why use BIDS?

- A standardized, community-driven way to organize and share data!
- Saves time!
  - Multiple people on one project?
  - Longitudinal study?
  - Want to reuse data for a new project or grant?
- Easy to understand!
- Growing number of applications built around BIDS!
- Contribute to public data repositories (e.g., OpenNeuro, LORIS)

Why use BIDS?

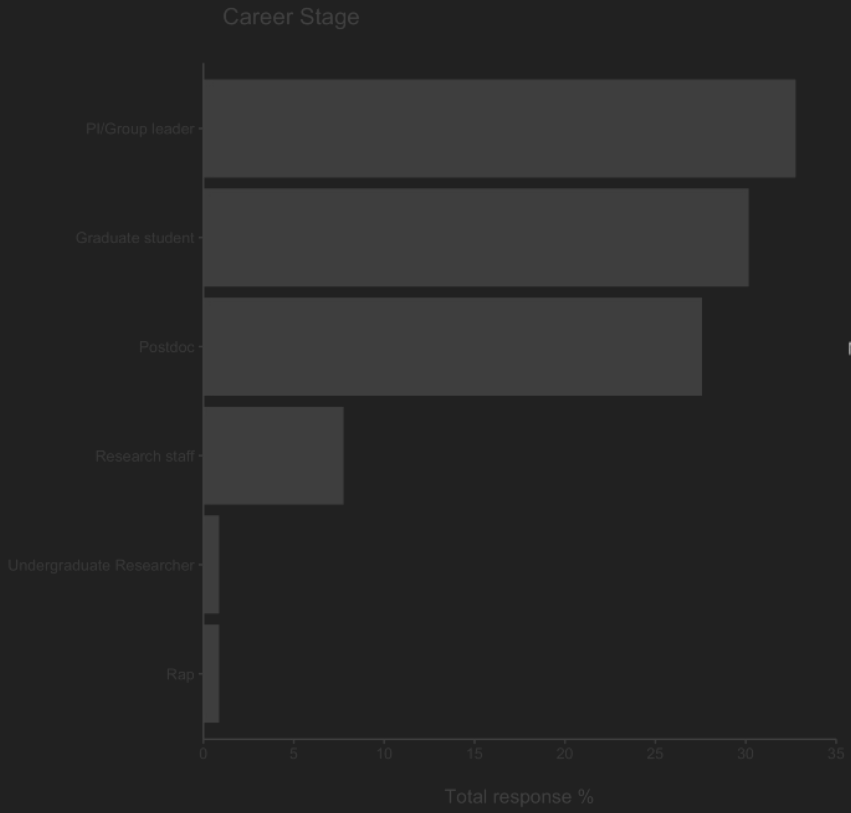
- A standardized, community-driven way to organize and share data!
- Saves time!
  - Multiple people on one project?
  - Longitudinal study?
  - Want to reuse data for a new project or grant?
- Easy to understand!
- Growing number of applications built around BIDS!
- Contribute to public data repositories (e.g., OpenNeuro, LORIS)
- Supports reproducible science!

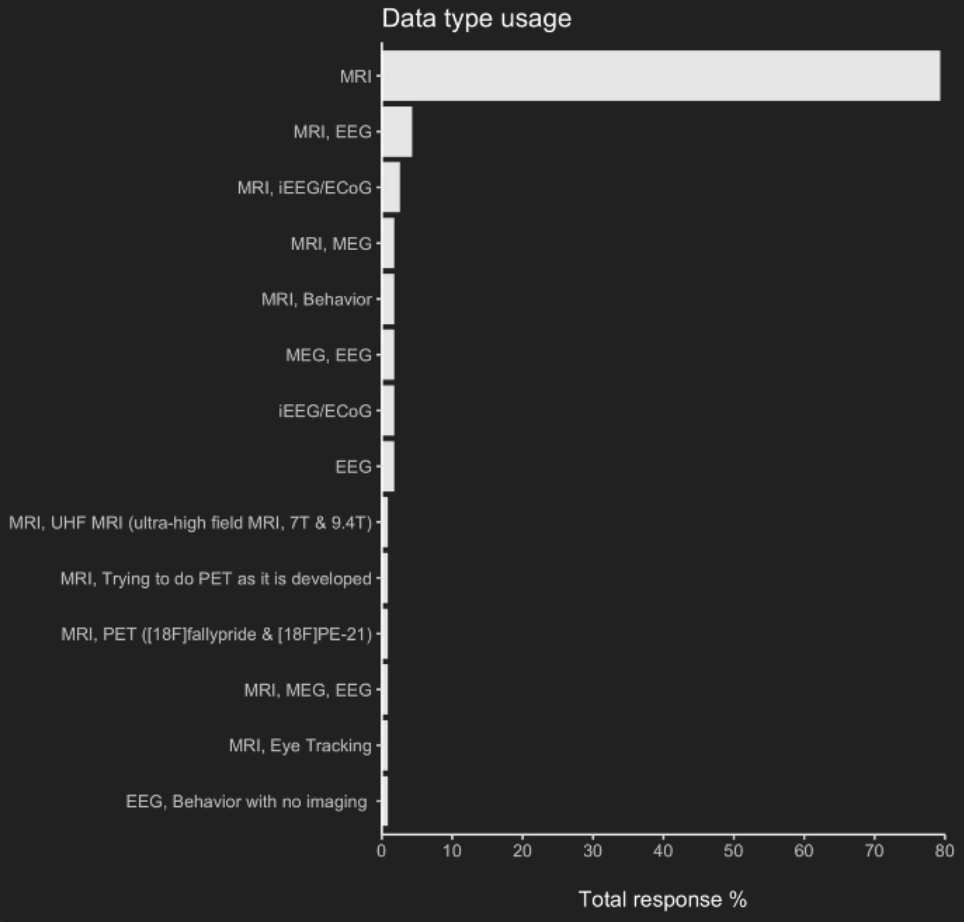
Who uses BIDS?



BIDS Usage Survey (2019)

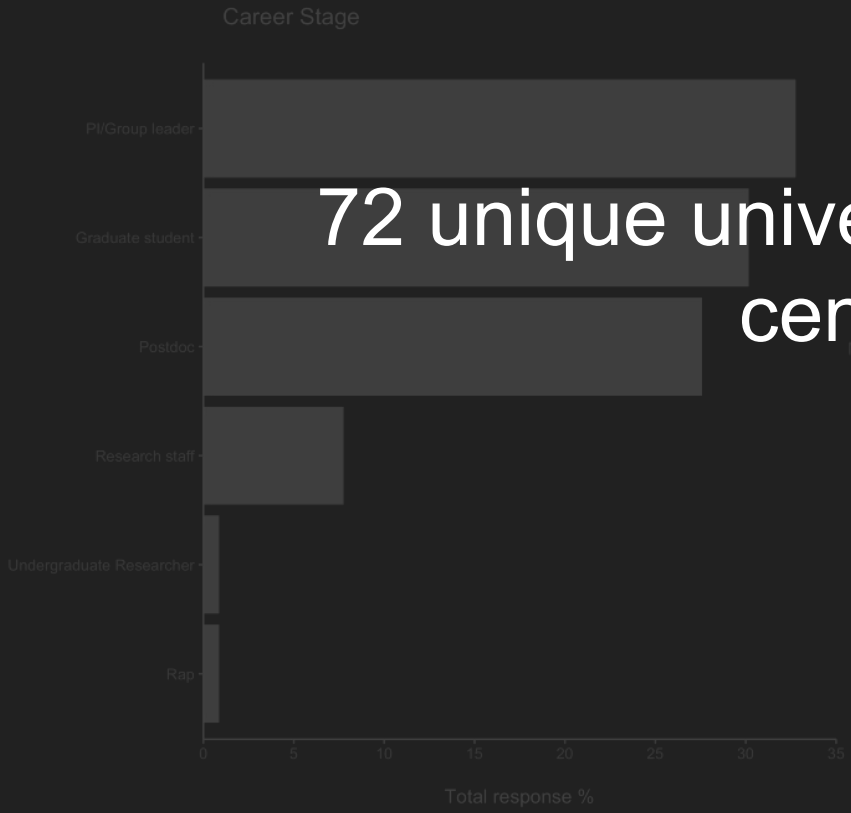
Who uses BIDS?





[BIDS Usage Survey (2019)]

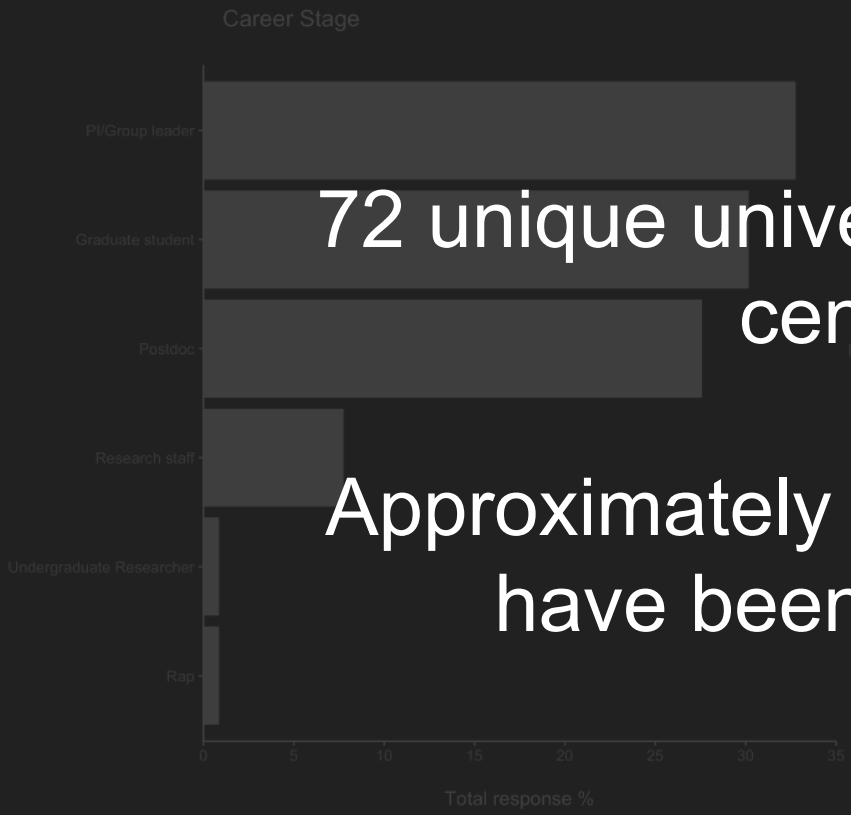
Who uses BIDS?



72 unique universities/research centers

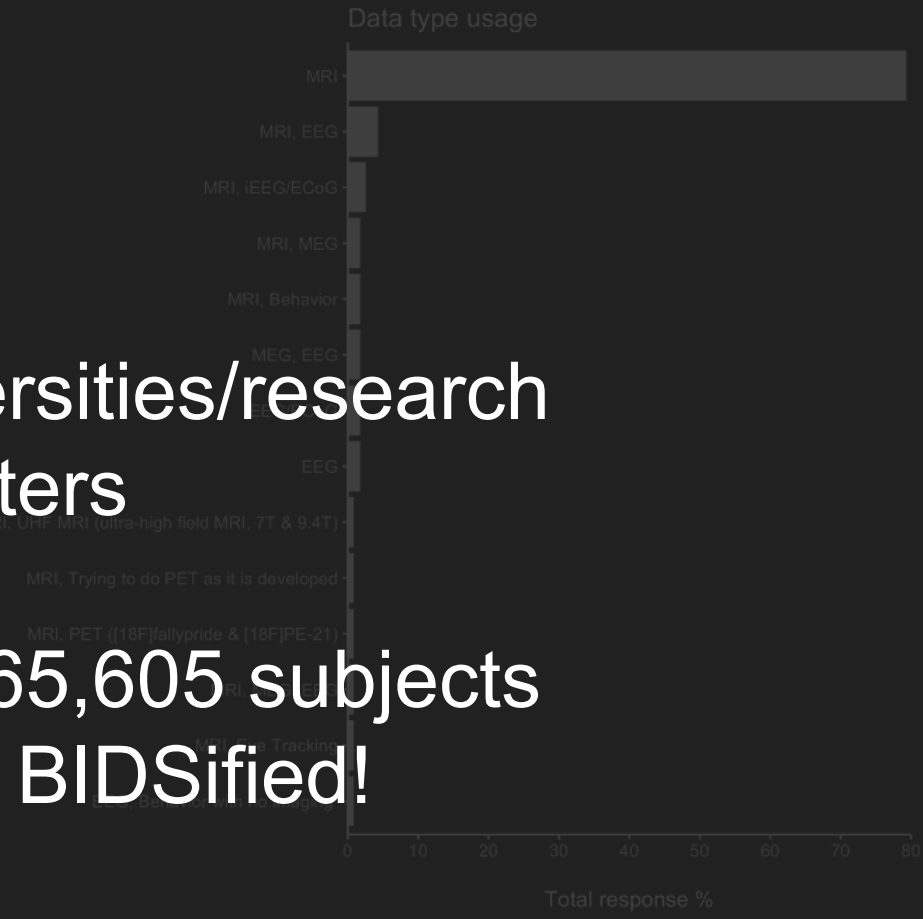
[BIDS Usage Survey (2019)]

Who uses BIDS?

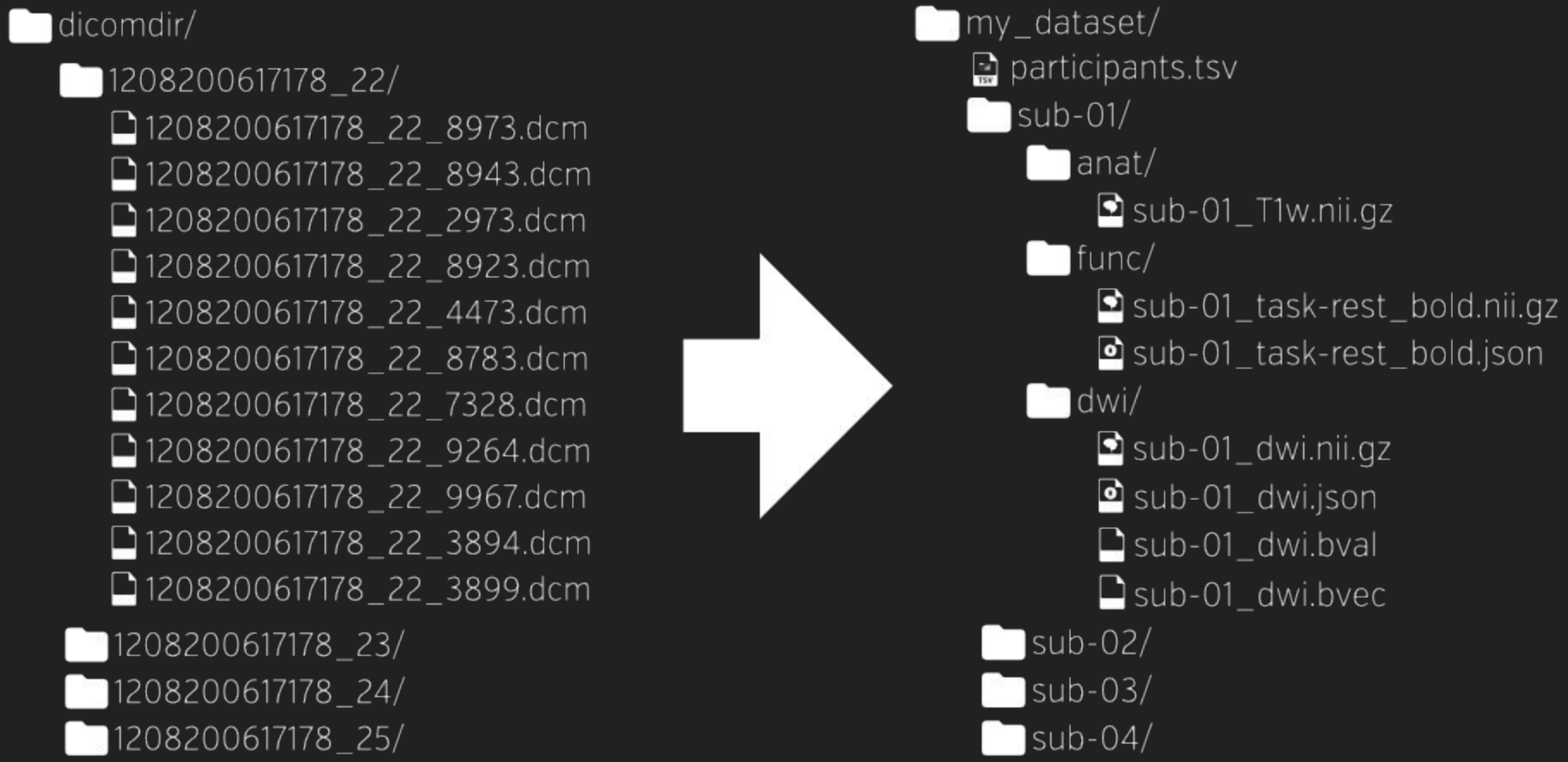


72 unique universities/research centers

Approximately 65,605 subjects have been BIDSified!



BIDS Usage Survey (2019)



https://bids.neuroimaging.io/index.html

Basic filename structure

sub-01

└─ anat

└─ func

Basic filename structure

sub-01

datatype

└─ anat

└─ func

Basic filename structure

sub-01

datatype

- anat
  - sub-01_T1w.nii.gz
  - sub-01_T1w.json
- func

Basic filename structure

sub-01

datatype

└─ anat

└─ sub-01_T1w.nii.gz

└─ sub-01_T1w.json

extension

└─ func

Basic filename structure

sub-01

└─ anat

└─ sub-01_T1w.nii.gz

└─ sub-01_T1w.json

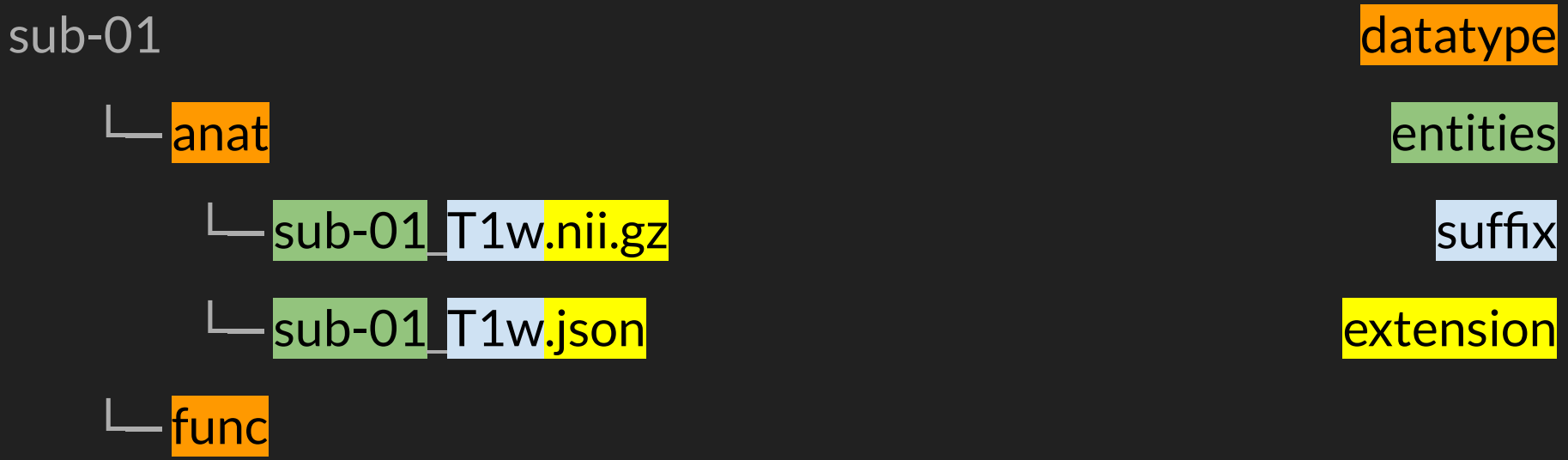
└─ func

datatype

suffix

extension

Basic filename structure



Basic filename structure

sub-01

└ anat

└ sub-01_T1w.nii.gz

└ sub-01_T1w.json

└ func

└ sub-01_task-localizer_run-01_bold.nii.gz

└ sub-01_task-localizer_run-01_bold.json

datatype

entities

suffix

extension

Basic filename structure

sub-01

- anat
  - sub-01_T1w.nii.gz
  - sub-01_T1w.json
- func
  - sub-01_task-localizer_run-01_bold.nii.gz
  - sub-01_task-localizer_run-01_bold.json

datatype

entities

suffix

extension

Basic filename structure

Anatomical	→	anat
Diffusion-weighted	→	dwi
Field mapping	→	fmap
Functional	→	func

```
dset/
├── dataset_description.json
├── participants.tsv
└── sub-01
    ├── anat
    │   ├── sub-01_T1w.json
    │   └── sub-01_T1w.nii.gz
    ├── dwi
    │   ├── sub-01_dwi.bval
    │   ├── sub-01_dwi.bvec
    │   ├── sub-01_dwi.json
    │   └── sub-01_dwi.nii.gz
    ├── fmap
    │   ├── sub-01_acq-dwi_dir-AP_epi.json
    │   ├── sub-01_acq-dwi_dir-AP_epi.nii.gz
    │   ├── sub-01_acq-dwi_dir-PA_epi.json
    │   ├── sub-01_acq-dwi_dir-PA_epi.nii.gz
    │   ├── sub-01_acq-func_dir-AP_epi.json
    │   ├── sub-01_acq-func_dir-AP_epi.nii.gz
    │   ├── sub-01_acq-func_dir-PA_epi.json
    │   └── sub-01_acq-func_dir-PA_epi.nii.gz
    └── func
        ├── sub-01_task-nback_run-01_bold.json
        ├── sub-01_task-nback_run-01_bold.nii.gz
        ├── sub-01_task-nback_run-01_events.tsv
        ├── sub-01_task-nback_run-01_sbref.json
        └── sub-01_task-nback_run-01_sbref.nii.gz
```

https://osf.io/drsx5/

Basic filename structure

```
sub-01
    └─ anat
        └─ sub-01_T1w.nii.gz
        └─ sub-01_T1w.json
```

Basic filename structure

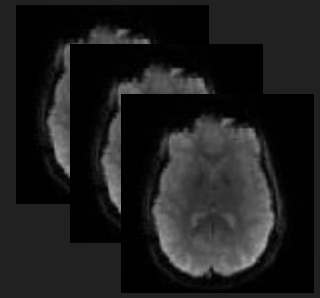
sub-01
└ anat
└ sub-01_T1w.nii.gz
└ sub-01_T1w.json

Basic filename structure

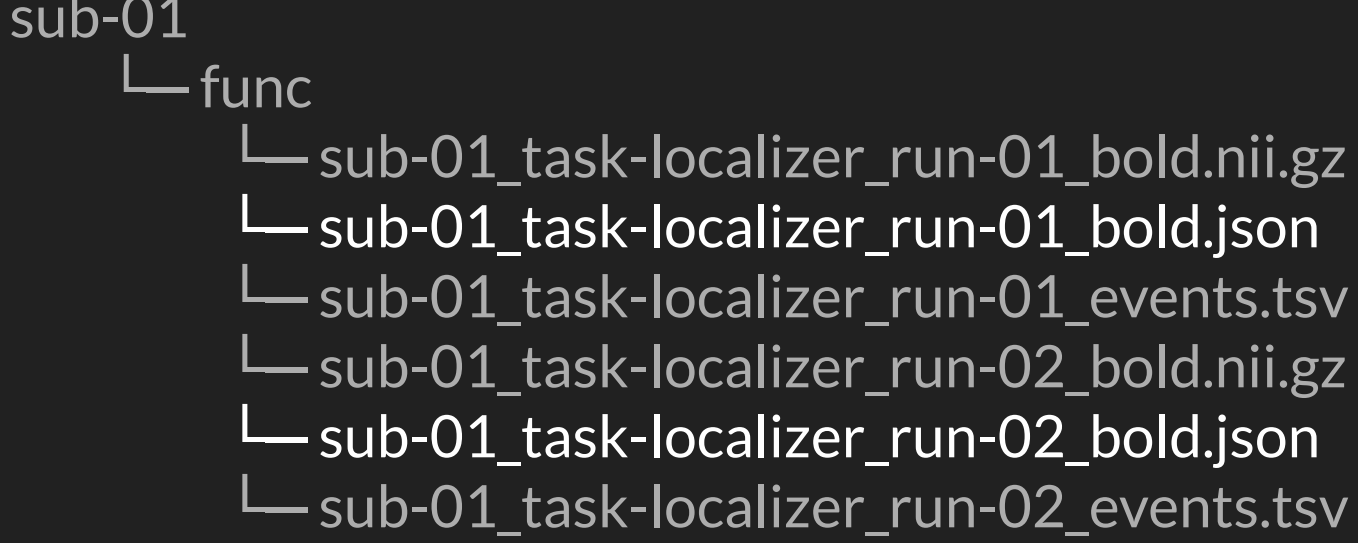
```
sub-01
    └─ func
        └─ sub-01_task-localizer_run-01_bold.nii.gz
        └─ sub-01_task-localizer_run-01_bold.json
        └─ sub-01_task-localizer_run-01_events.tsv
        └─ sub-01_task-localizer_run-02_bold.nii.gz
        └─ sub-01_task-localizer_run-02_bold.json
        └─ sub-01_task-localizer_run-02_events.tsv
```

Basic filename structure

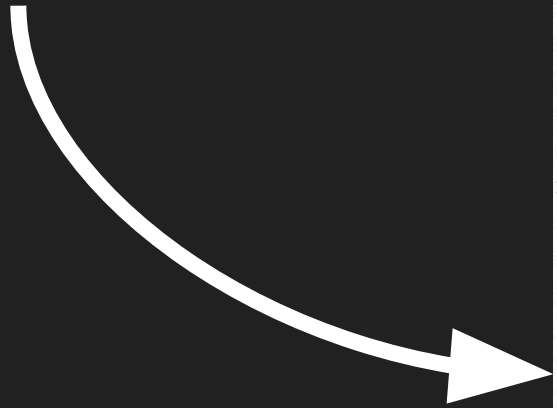
```
sub-01
    └─ func
        └─ sub-01_task-localizer_run-01_bold.nii.gz
        └─ sub-01_task-localizer_run-01_bold.json
        └─ sub-01_task-localizer_run-01_events.tsv
        └─ sub-01_task-localizer_run-02_bold.nii.gz
        └─ sub-01_task-localizer_run-02_bold.json
        └─ sub-01_task-localizer_run-02_events.tsv
```



Basic filename structure

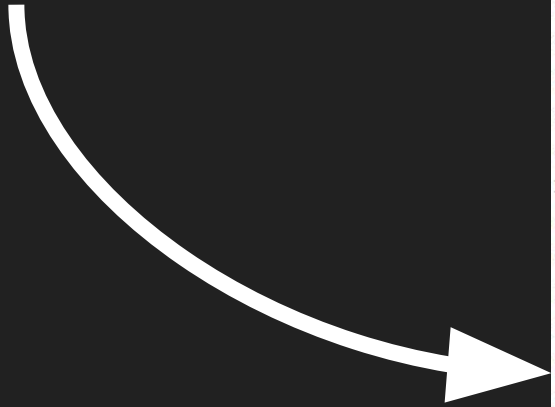


sub-01_task-localizer_run-01_bold.json



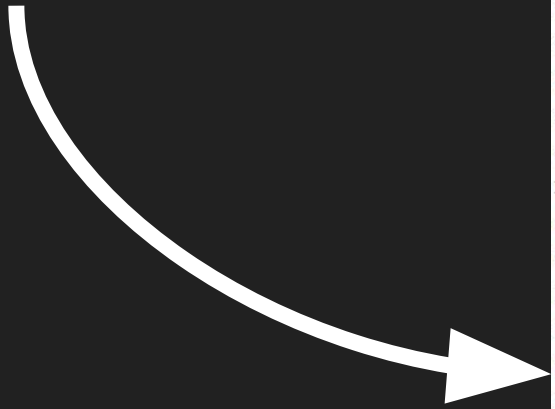

```
RepetitionTime: 1.06
MatrixCoilMode: "Auto"
FlipAngle: 15
ProcedureStepDescription: "CBU_Neuroimaging"
ConversionSoftwareVersion: "v1.0.20170130"
SliceEncodingDirection: "k"
EffectiveEchoSpacing: 0.000345002
CoilCombinationMethod: "Sum of squares"
ManufacturersModelName: "TrioTim"
ProtocolName: "3D_EPI_2mm_localiser_B380"
TaskName: "localizer"
InstitutionAddress: "15 Chaucer Road, Cambridge CB2 7EF, UK"
ImageType: [] 8 items
ReceiveCoilName: "32 Chn Head"
EchoTime: 0.03
MagneticFieldStrength: 3
SliceTiming: [] 30 items
ParallelReductionFactorInPlane: 2
InstitutionName: "MRC Cognition and Brain Sciences Unit"
PulseSequenceType: "3D Gradient Echo Planar Imaging"
PhaseEncodingDirection: "j-"
AcquisitionDateTime: "2012-09-12T12:54:43.820312"
ParallelAcquisitionTechnique: "GRAPPA"
NumberShots: 1
ConversionSoftware: "dcm2niix"
HardcopyDeviceSoftwareVersion: "VB17"
Manufacturer: "Siemens"
```

sub-01_task-localizer_run-01_bold.json



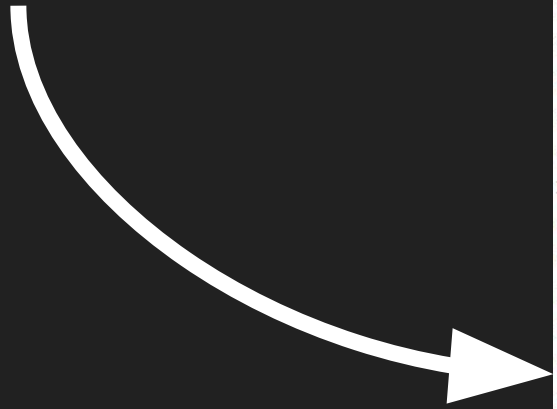

```
RepetitionTime: 1.06
MatrixCoilMode: "Auto"
FlipAngle: 15
ProcedureStepDescription: "CBU_Neuroimaging"
ConversionSoftwareVersion: "v1.0.20170130"
SliceEncodingDirection: "k"
EffectiveEchoSpacing: 0.000345002
CoilCombinationMethod: "Sum of squares"
ManufacturersModelName: "TrioTim"
ProtocolName: "3D_EPI_2mm_localiser_B380"
TaskName: "localizer"
InstitutionAddress: "15 Chaucer Road, Cambridge CB2 7EF, UK"
ImageType: [] 8 items
ReceiveCoilName: "32 Chn Head"
EchoTime: 0.03
MagneticFieldStrength: 3
SliceTiming: [] 30 items
ParallelReductionFactorInPlane: 2
InstitutionName: "MRC Cognition and Brain Sciences Unit"
PulseSequenceType: "3D Gradient Echo Planar Imaging"
PhaseEncodingDirection: "j-"
AcquisitionDateTime: "2012-09-12T12:54:43.820312"
ParallelAcquisitionTechnique: "GRAPPA"
NumberShots: 1
ConversionSoftware: "dcm2niix"
HardcopyDeviceSoftwareVersion: "VB17"
Manufacturer: "Siemens"
```

sub-01_task-localizer_run-01_bold.json




```
RepetitionTime: 1.06
MatrixCoilMode: "Auto"
FlipAngle: 15
ProcedureStepDescription: "CBU_Neuroimaging"
ConversionSoftwareVersion: "v1.0.20170130"
SliceEncodingDirection: "k"
EffectiveEchoSpacing: 0.000345002
CoilCombinationMethod: "Sum of squares"
ManufacturersModelName: "TrioTim"
ProtocolName: "3D_EPI_2mm_localiser_B380"
TaskName: "localizer"
InstitutionAddress: "15 Chaucer Road, Cambridge CB2 7EF, UK"
ImageType: [] 8 items
ReceiveCoilName: "32 Chn Head"
EchoTime: 0.03
MagneticFieldStrength: 3
SliceTiming: [] 30 items
ParallelReductionFactorInPlane: 2
InstitutionName: "MRC Cognition and Brain Sciences Unit"
PulseSequenceType: "3D Gradient Echo Planar Imaging"
PhaseEncodingDirection: "j-"
AcquisitionDateTime: "2012-09-12T12:54:43.820312"
ParallelAcquisitionTechnique: "GRAPPA"
NumberShots: 1
ConversionSoftware: "dcm2niix"
HardcopyDeviceSoftwareVersion: "VB17"
Manufacturer: "Siemens"
```

sub-01_task-localizer_run-01_bold.json

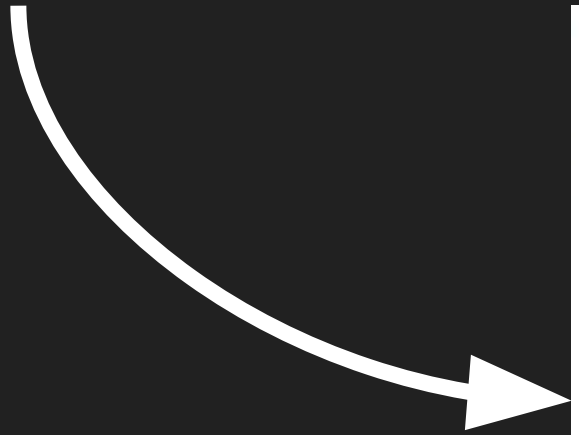



```
RepetitionTime: 1.06
MatrixCoilMode: "Auto"
FlipAngle: 15
ProcedureStepDescription: "CBU_Neuroimaging"
ConversionSoftwareVersion: "v1.0.20170130"
SliceEncodingDirection: "k"
EffectiveEchoSpacing: 0.000345002
CoilCombinationMethod: "Sum of squares"
ManufacturersModelName: "TrioTim"
ProtocolName: "3D_EPI_2mm_localiser_B380"
TaskName: "localizer"
InstitutionAddress: "15 Chaucer Road, Cambridge CB2 7EF, UK"
ImageType: [] 8 items
ReceiveCoilName: "32 Chn Head"
EchoTime: 0.03
MagneticFieldStrength: 3
SliceTiming: [] 30 items
ParallelReductionFactorInPlane: 2
InstitutionName: "MRC Cognition and Brain Sciences Unit"
PulseSequenceType: "3D Gradient Echo Planar Imaging"
PhaseEncodingDirection: "j-"
AcquisitionDateTime: "2012-09-12T12:54:43.820312"
ParallelAcquisitionTechnique: "GRAPPA"
NumberShots: 1
ConversionSoftware: "dcm2niix"
HardcopyDeviceSoftwareVersion: "VB17"
Manufacturer: "Siemens"
```

Basic filename structure

sub-01
└─ func
└─ sub-01_task-localizer_run-01_bold.nii.gz
└─ sub-01_task-localizer_run-01_bold.json
└─ sub-01_task-localizer_run-01_events.tsv
└─ sub-01_task-localizer_run-02_bold.nii.gz
└─ sub-01_task-localizer_run-02_bold.json
└─ sub-01_task-localizer_run-02_events.tsv

sub-01_task-localizer_run-01_events.tsv



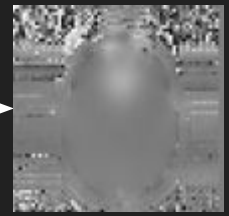
ONSET	DURATION	TRIAL_TYPE
16.01	16.00	objects
48.00	16.00	scrambled
80.00	16.00	faces
112.00	16.00	places
144.00	16.00	objects
176.00	16.00	places
208.00	16.00	faces
240.00	16.00	scrambled
272.00	16.00	scrambled
304.00	16.00	objects
336.00	16.00	faces
368.00	16.00	places

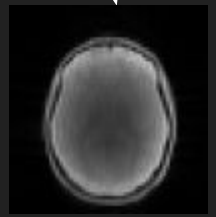
Basic filename structure

```
sub-01
    └─ fmap
        └─ sub-01_phasediff.nii.gz
        └─ sub-01_phasediff.json
        └─ sub-01_magnitude1.nii.gz
        └─ sub-01_magnitude1.json
```

Basic filename structure

sub-01
└ fmap
└ sub-01_phasediff.nii.gz
└ sub-01_phasediff.json
└ sub-01_magnitude1.nii.gz
└ sub-01_magnitude1.json





Converting MRI data to BIDS

HeuDiConv Installation

- Local: `pip install heudiconv[all]`
- Anaconda: `conda install --channel conda-forge heudiconv`
- Docker: `docker pull nipy/heudiconv:latest`
- Singularity: `singularity pull docker://nipy/heudiconv:latest`

Basic Pipeline

DICOM (*.dcm, *.IMA)
[raw data]
[no structure]

via HeuDiConv

NIFTI (*.nii.gz)
[raw data]
[BIDS format]

via fMRIPrep

NIFTI (*.nii.gz)
[preprocessed data]
[BIDS format]

Basic Pipeline

DICOM (*.dcm, *.IMA)
[raw data]
[no structure]

user-defined heuristic

via HeuDiConv

NIFTI (*.nii.gz)
[raw data]
[BIDS format]

via fMRIPrep

NIFTI (*.nii.gz)
[preprocessed data]
[BIDS format]

Creating your heuristic: `convertall.py`

```
LOCAL_DIR=/mnt/data/raw/
OUT_DIR=/mnt/data/subjects/
SUBID=01

heudiconv \
-d ${LOCAL_DIR}/{subject}/*/*IMA \
-s ${SUBID} \
-f convertall \
-c none \
-o ${OUT_DIR}
```

${OUT_DIR}/.heudiconv/{subject}/info/dicominfo.tsv

C	D	E	F	G	H	I	J	K	L	M
series_id	dcm_dir_r	unspecifie	unspecifie	dim1	dim2	dim3	dim4	TR	TE	protocol_
5-gre_fiel	GRE_FIELD	-	-	68	68	92	1	0.449	7.38	gre_field_
6-gre_fiel	GRE_FIELD	-	-	68	68	46	1	0.449	7.38	gre_field_
7-ep2d 3.(	EP2D_3_0I	-	-	68	68	46	121	2.5	30	ep2d 3.0m
9-ep2d 3.(	EP2D_3_0I	-	-	68	68	46	121	2.5	30	ep2d 3.0m
11-ep2d 3	EP2D_3_0I	-	-	68	68	46	121	2.5	30	ep2d 3.0m
13-ep2d 3	EP2D_3_0I	-	-	68	68	46	121	2.5	30	ep2d 3.0m
15-MPRAC	MPRAGE_	-	-	288	270	176	1	2.3	2.99	MPRAGE (
16-ep2d 3	EP2D_3_0I	-	-	68	68	46	241	2.5	30	ep2d 3.0m
18-ep2d 3	EP2D_3_0I	-	-	68	68	46	241	2.5	30	ep2d 3.0m
20-ep2d 3	EP2D_3_0I	-	-	68	68	46	241	2.5	30	ep2d 3.0m
26-gre_fie	GRE_FIELD	-	-	70	70	112	1	0.546	7.38	gre_field_
27-gre_fie	GRE_FIELD	-	-	70	70	56	1	0.546	7.38	gre_field_
28-ep2d 2	EP2D_2_0I	-	-	102	102	56	306	1.79	30	ep2d 2.0m

Creating your heuristic: `convert_study.py`

- Create a custom Python script (e.g., convert_study.py)
- This allows HeuDiConv to convert your data to BIDS based on your study's parameters
- See a template script here:

  https://drive.google.com/file/d/1bwMm2bp3iZs33QpfaYIk44kQ0MSder0c/view?usp=sharing

Running HeuDiConv: `convert_study.py`

```
LOCAL_DIR=/mnt/data/raw/
OUT_DIR=/mnt/data/subjects/
SUBID=01

heudiconv \
-d ${LOCAL_DIR}/{subject}/*/*IMA \
-s ${SUBID} \
-f convert_study.py \
-c dcm2niix \
-b --minmeta \
-o ${OUT_DIR}
```

Running HeuDiConv: `convert_study.py`

```
LOCAL_DIR=/mnt/data/raw/
OUT_DIR=/mnt/data/subjects/
SUBID=01
SESID=01
heudiconv \
-d ${LOCAL_DIR}/{subject}/{session}/*/*IMA \
-s ${SUBID} --ses ${SESID} \
-f convert_study.py \
-c dcm2niix \
-b --minmeta \
-o ${OUT_DIR}
```

Running HeuDiConv using Docker

```
SUBID=01
docker run --rm -it \
-v /mnt/data/raw:/raw:ro \
-v /mnt/data/subjects:/output \
nipy/heudiconv:latest \
-d /raw/{subject}/*/*IMA \
-s ${SUBID} \
-f convert_study.py \
-c dcm2niix \
-b --minmeta \
-o /output
```

Running HeuDiConv using Singularity

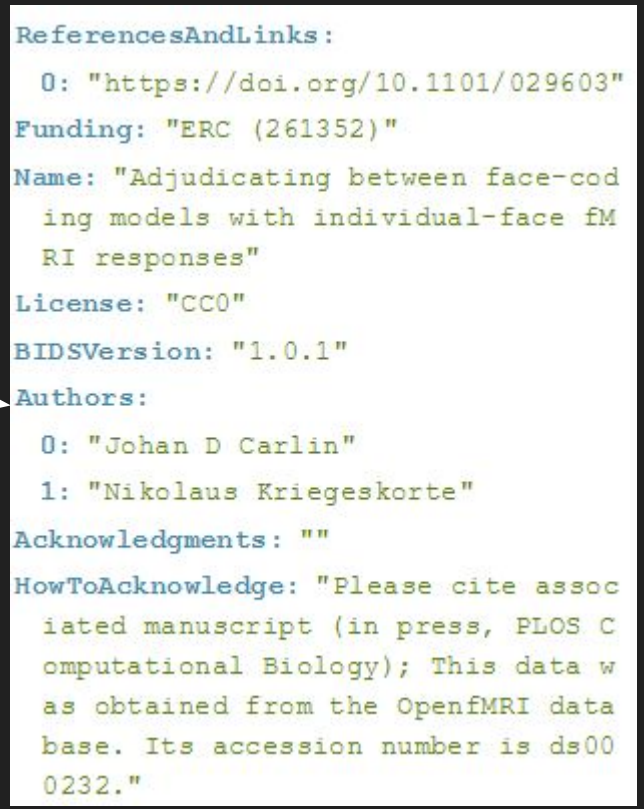
```
SUBID=01
singularity run --cleanenv \
--bind /mnt/data:/data \
~/simgs/heudiconv-latest.simg \
-d /data/raw/{subject}/*/*IMA \
-s ${SUBID} \
-f convert_study.py \
-c dcm2niix \
-b --minmeta \
-o /data/output
```

Outputs in ${OUT_DIR}/

- dataset_description.json
- participants.tsv
- README
- sub-01
    - anat
    - func
    - fmap
- sub-02
    - anat
    - func
    - fmap

Outputs in ${OUT_DIR}/

- dataset_description.json
- participants.tsv
- README
- sub-01
  - anat
  - func
  - fmap
- sub-02
  - anat
  - func
  - fmap



Outputs in ${OUT_DIR}/

- dataset_description.json
- participants.tsv
- README
- sub-01
  - anat
  - func
  - fmap
- sub-02
  - anat
  - func
  - fmap

PARTICIPAN…	AGE	SEX
sub-01	27	F
sub-02	29	F
sub-03	25	F
sub-04	30	F
sub-05	38	F
sub-06	22	F
sub-07	31	M
sub-08	36	M
sub-09	37	M
sub-10	24	F

Outputs in ${OUT_DIR}/

- dataset_description.json
- participants.tsv
- README
- sub-01
    - anat
    - func
    - fmap
- sub-02
    - anat
    - func
    - fmap

Outputs in ${OUT_DIR}/

- dataset_description.json
- participants.tsv
- README

```
sub-01
    └─ ses-01
        └─ anat
        └─ func
        └─ fmap
    └─ ses-02
        └─ anat
        └─ func
        └─ fmap
```

Some final notes

- Your .json files might contain identifiable information about subjects (data from Georgetown CFMI typically doesn't have this, but double-check)
- Anatomical data need to be de-faced
- Metadata in *_phasediff.json and *magnitude.json files may need to be manually entered for fMRIPrep (if multiple fmaps), add:

```
"IntendedFor":
["func/sub-'${SUBID}'_task-localizer_run-01_bold.nii.gz",
"func/sub-'${SUBID}'_task-localizer_run-02_bold.nii.gz",
"func/sub-'${SUBID}'_task-localizer_run-03_bold.nii.gz"],
```

More resources!

- BIDS Starter Kit
- BIDS Tutorial Series
- BIDS Core Specification Outline
- Existing BIDS Presentations
- Guide to building BIDS apps
- BIDS app template on GitHub
- BIDS Validator
- pyBIDS
- Neuroimaging Core (Read the Docs) Resources