The benefits of BIDS

data standardization and automated processing for neuroimaging research

Samuel A. Nastase
Princeton Neuroscience Institute
snastase@princeton.edu

BIDS

BRAIN IMAGING DATA STRUCTURE

@samnastase

The benefits of BIDS

data standardization and automated processing for neuroimaging research

Samuel A. Nastase
Princeton Neuroscience Institute
snastase@princeton.edu

Thanks to the many **BIDS contributors**!

BIDS

BRAIN IMAGING DATA STRUCTURE

@samnastase

Brain Imaging Data Structure (BIDS)

What is BIDS?

—a standard for organizing neuroimaging data that facilitates **re-use** and **automated processing**

—the standard specifies machine-readable directory structure, filenames, file formats, and metadata

—consensus-based community-driven development capitalizing on existing conventions

—emphasis on simplicity, readability, and accessibility, i.e. ease of use and adoption

—developed out of the **OpenfMRI** (now **OpenNeuro**) open neuroimaging data repository

Extra credit

FAIR principles (Findable, Accessible, Interoperable, and Reusable)

Gorgolewski et al, *Sci Data*, 2016
Wilkinson et al, *Sci Data*, 2016

Brain Imaging Data Structure (BIDS)

What is BIDS?

—a standard for organizing neuroimaging data that facilitates **re-use** and **automated processing**

—the standard specifies machine-readable directory structure, filenames, file formats, and metadata

—consensus-based community-driven development capitalizing on existing conventions

—emphasis on simplicity, readability, and accessibility, i.e. ease of use and adoption

—developed out of the **OpenfMRI** (now **OpenNeuro**) open neuroimaging data repository

What's the point?

—have you ever tried to re-analyze someone else's data?

—have you ever tried to re-analyze *your own data* a year later? (¬_¬)

—have you ever felt *confident* in data you didn't collect yourself?

2016: "To date there has been no consensus about how to organize and share [neuroimaging] data, leading researchers, even those working within the same lab, to arrange their data in different and idiosyncratic ways. Lack of consensus leads to misunderstanding and time wasted on rearranging data or rewriting scripts that expect particular file formats and organization, as well as a possible cause for errors."

Gorgolewski et al, *Sci Data*, 2016
Wilkinson et al, *Sci Data*, 2016

Brain Imaging Data Structure (BIDS)

What is BIDS?

—a standard for organizing neuroimaging data that facilitates **re-use** and **automated processing**

—the standard specifies machine-readable directory structure, filenames, file formats, and metadata

—consensus-based community-driven development capitalizing on existing conventions

—emphasis on simplicity, readability, and accessibility, i.e. ease of use and adoption

—developed out of the **OpenfMRI** (now **OpenNeuro**) open neuroimaging data repository

The importance of data sharing

Public, well-curated datasets have been tremendously beneficial in fields like machine learning; e.g., ImageNet, MNIST, COCO, CIFAR.

The re-use of publicly shared fMRI datasets from the INDI consortia has saved an estimated $1.7 billion in data generation costs—and this doesn't even include HCP, OpenNeuro, ABCD, or UK Biobank.

Ferguson et al, *Nat Neurosci*, 2014
Poldrack & Gorgolewski, *Nat Neurosci*, 2014
Poldrack et al, *Nat Rev Neurosci*, 2017
Milham et al, *Nat Commun*, 2018

Brain Imaging Data Structure (BIDS)

What is BIDS?

—a standard for organizing neuroimaging data that facilitates **re-use** and **automated processing**

—the standard specifies machine-readable directory structure, filenames, file formats, and metadata

—consensus-based community-driven development capitalizing on existing conventions

—emphasis on simplicity, readability, and accessibility, i.e. ease of use and adoption

—developed out of the **OpenfMRI** (now **OpenNeuro**) open neuroimaging data repository

The importance of automated processing

Neuroimaging analysis is complex and flexible—relying on multi-stage processing workflows with many possible analysis choices at each stage (i.e. “researcher degrees of freedom”).

The machine-readable BIDS organization with rich metadata allows for adaptive, automated processing and analysis (i.e. BIDS Apps) that:

—minimize error-prone manual intervention and “procedural overfitting”

—maximize reproducible execution via containerization and content-tracking

Carp, *Front Neurosci*, 2012
Ghosh et al, *F1000Research*, 2017
Botvinik-Nezer et al, *Nature*, 2020

Brain Imaging Data Structure (BIDS)

SCIENTIFIC DATA

OPEN

SUBJECT CATEGORIES
» Data publication and archiving
» Research data

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

Krzysztof J. Gorgolewski[1], Tibor Auer[2], Vince D. Calhoun[3,4], R. Cameron Craddock[5,6], Samir Das[7], Eugene P. Duff[8], Guillaume Flandin[9], Satrajit S. Ghosh[10,11], Tristan Glatard[7,12], Yaroslav O. Halchenko[13], Daniel A. Handwerker[14], Michael Hanke[15,16], David Keator[17], Xiangrui Li[18], Zachary Michael[19], Camille Maumet[20], B. Nolan Nichols[21,22], Thomas E. Nichols[20,23], John Pellman[6], Jean-Baptiste Poline[24], Ariel Rokem[25], Gunnar Schaefer[1,26], Vanessa Sochat[27], William Triplett[1], Jessica A. Turner[3,28], Gaël Varoquaux[29] & Russell A. Poldrack[1]

Received: 18 December 2015
Accepted: 19 May 2016
Published: 21 June 2016

Brain Imaging Data Structure (BIDS)

SCIENTIFIC DATA

OPEN SCIENTIFIC DATA

SUBJECT CATEGORIES
» Data publication and archiving
» Research data

Received: 18 December 2015
Accepted: 19 May 2016
Published: 21 June 2016

OPEN

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

Guiomar Niso[1,2], Krzysztof J. Gorgolewski[3], Elizabeth Bock[1], Teon L. Brooks[3], Guillaume Flandin[4], Alexandre Gramfort[5,6], Richard N. Henson[7], Mainak Jas[5], Vladimir Litvak[4], Jeremy T. Moreau[1], Robert Oostenveld[8,9], Jan-Mathijs Schoffelen[8], Francois Tadel[1,10,11], Joseph Wexler[3] & Sylvain Baillet[1]

Received: 14 November 2017
Accepted: 3 May 2018
Published: 19 June 2018

Brain Imaging Data Structure (BIDS)

SCIENTIFIC DATA

OPEN

SUBJECT CATEGORIES
» Data publication and archiving
» Research data

Received: 18 December 2015
Accepted: 19 May 2016
Published: 21 June 2016

SCIENTIFIC DATA

OPEN

Received: 14 November 2017
Accepted: 3 May 2018
Published: 19 June 2018

SCIENTIFIC DATA

OPEN

COMMENT

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

Cyril R. Pernet[1], Stefan Appelhoff[2], Krzysztof J. Gorgolewski[3], Guillaume Flandin[4], Christophe Phillips[5], Arnaud Delorme[6,7] & Robert Oostenveld[8,9]

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

Brain Imaging Data Structure (BIDS)

SCIENTIFIC DATA

OPEN

SUBJECT CATEGORIES
» Data publication and archiving
» Research data

Received: 18 December 2015
Accepted: 19 May 2016
Published: 21 June 2016

SCIENTIFIC DATA

OPEN

Received: 14 November 2017
Accepted: 3 May 2018
Published: 19 June 2018

SCIENTIFIC DATA

OPEN

COMMENT

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

SCIENTIFIC DATA

OPEN

COMMENT

Received: 29 January 2019
Accepted: 24 May 2019
Published online: 25 June 2019

iEEG-BIDS, extending the Brain Imaging Data Structure specification to human intracranial electrophysiology

Christopher Holdgraf[1,16], Stefan Appelhoff[2], Stephan Bickel[3], Kristofer Bouchard[4], Sasha D'Ambrosio[5], Olivier David[6], Orrin Devinsky[7], Benjamin Dichter[8], Adeen Flinker[7], Brett L. Foster[9], Krzysztof J. Gorgolewski[8], Iris Groen[10], David Groppe[11], Aysegul Gunduz[12], Liberty Hamilton[13], Christopher J. Honey[14], Mainak Jas[15], Robert Knight[16], Jean-Philippe Lachaux[17], Jonathan C. Lau[18], Christopher Lee-Messer[8], Brian N. Lundstrom[19], Kai J. Miller[20], Jeffrey G. Ojemann[21], Robert Oostenveld[22], Natalia Petridou[23], Gio Piantoni[24], Andrea Pigorini[5], Nader Pouratian[25], Nick F. Ramsey[24], Arjen Stolk[16], Nicole C. Swann[26], François Tadel[6,27], Bradley Voytek[28], Brian A. Wandell[8], Jonathan Winawer[10], Kirstie Whitaker[29,32], Lyuba Zehl[30] & Dora Hermes[8,24,31]

Brain Imaging Data Structure (BIDS)

SCIENTIFIC DATA

OPEN

SUBJECT CATEGORIES
» Data publication and archiving
» Research data

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

Krzysztof J. Gorgolewski[1], Tibor Auer[2], Vince D. Calhoun[3,4], R. Cameron Craddock[5,6], Samir Das[7], Eugene P. Duff[8], Guillaume Flandin[9], Satrajit S. Ghosh[10,11], Tristan Glatard[7,12], Yaroslav O. Halchenko[13], Daniel A. Handwerker[14], Michael Hanke[15,16], David Keator[17], Xiangrui Li[18], Zachary Michael[19], Camille Maumet[20], B. Nolan Nichols[21,22], Thomas E. Nichols[20,23], John Pellman[6], Jean-Baptiste Poline[24], Ariel Rokem[25], Gunnar Schaefer[1,26], Vanessa Sochat[27], William Triplett[1], Jessica A. Turner[3,28], Gaël Varoquaux[29] & Russell A. Poldrack[1]

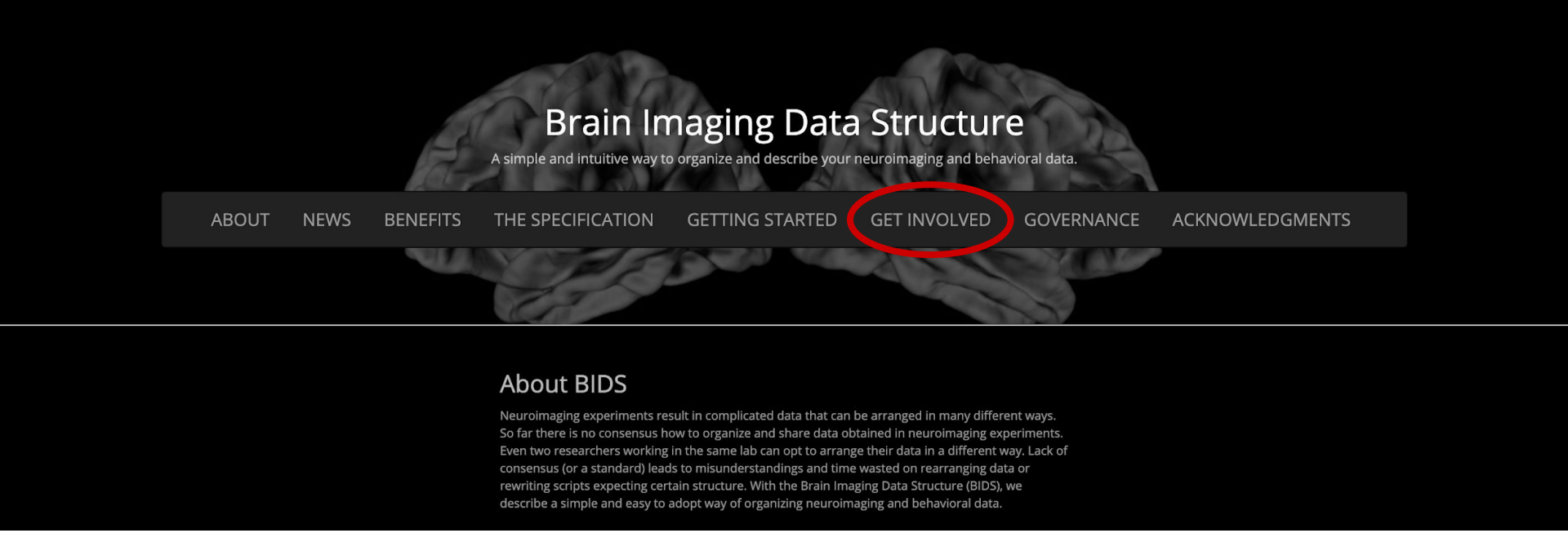
Received: 18 December 2015
Accepted: 19 May 2016
Published: 21 June 2016

Brain Imaging Data Structure

A simple and intuitive way to organize and describe your neuroimaging and behavioral data.

About BIDS

Neuroimaging experiments result in complicated data that can be arranged in many different ways. So far there is no consensus how to organize and share data obtained in neuroimaging experiments. Even two researchers working in the same lab can opt to arrange their data in a different way. Lack of consensus (or a standard) leads to misunderstandings and time wasted on rearranging data or rewriting scripts expecting certain structure. With the Brain Imaging Data Structure (BIDS), we describe a simple and easy to adopt way of organizing neuroimaging and behavioral data.

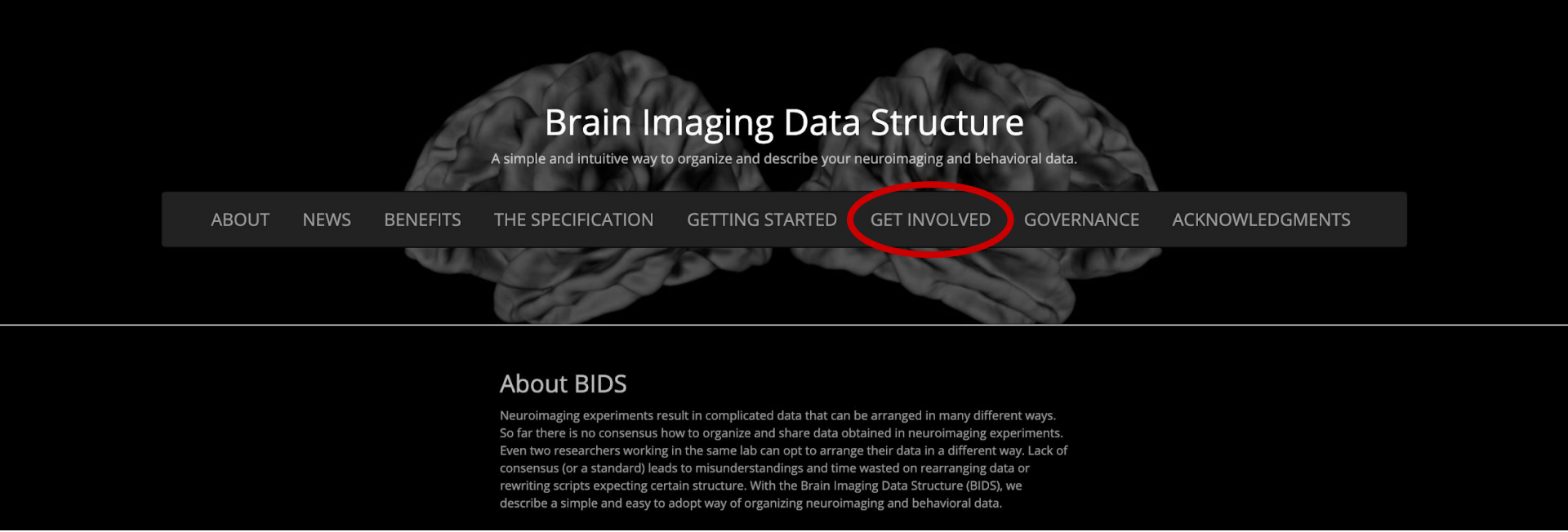


Get involved

—check out the **BIDS tag** on **NeuroStars** (INCF-funded question/answer forum)

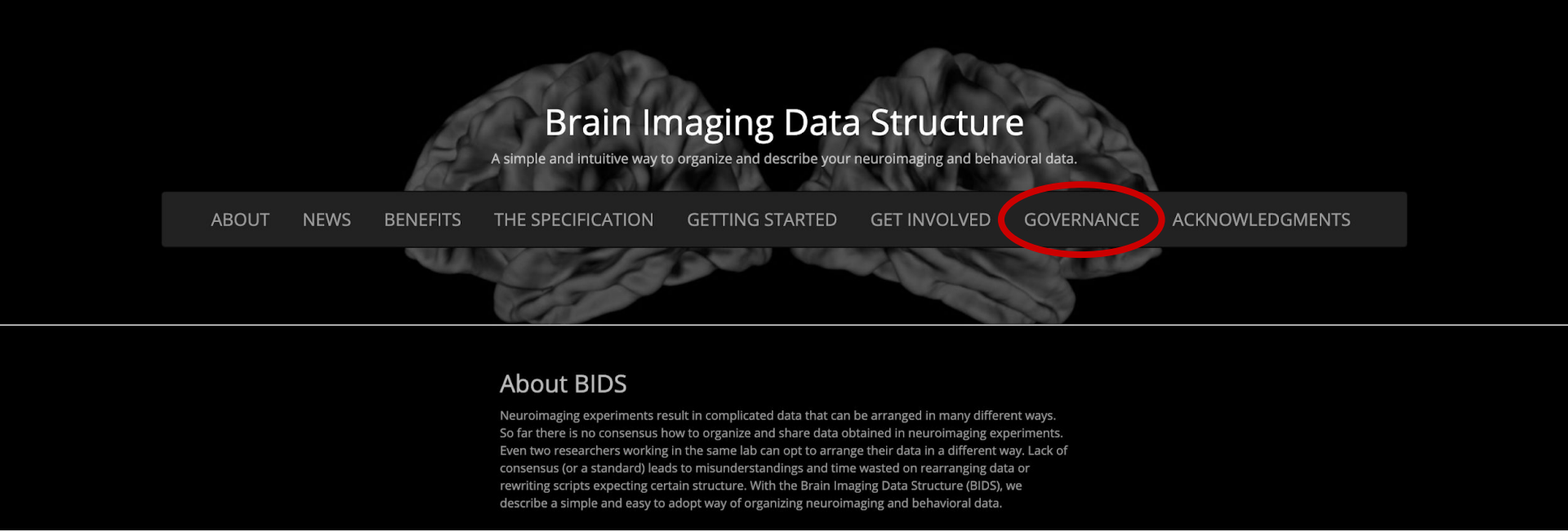
—create **issues** or **pull requests** on the **GitHub repository**

—read the **BIDS Code of Conduct**



BIDS Extension Proposals (BEPs)

- community-driven expansion of the BIDS ecosystem (read the **BEP guidelines**)
- there are currently 20+ BEPs under development, including:
  - different neuroimaging modalities (e.g. **CT**, **PET**, **NIRS**)
  - different derivatives (e.g. **functional**, **structural**, **diffusion**)
  - different analysis tools (e.g. **statistical models**, **transformations**, **execution**)



BIDS governance

Community-driven development via democratically elected steering, working, and interest groups:

—minimize complexity and facilitate adoption, reuse existing methods and technologies whenever possible

—tackle 80% of the most commonly used neuroimaging data, derivatives, and models

—adoption by the global neuroimaging community and their input during the creation of the specification

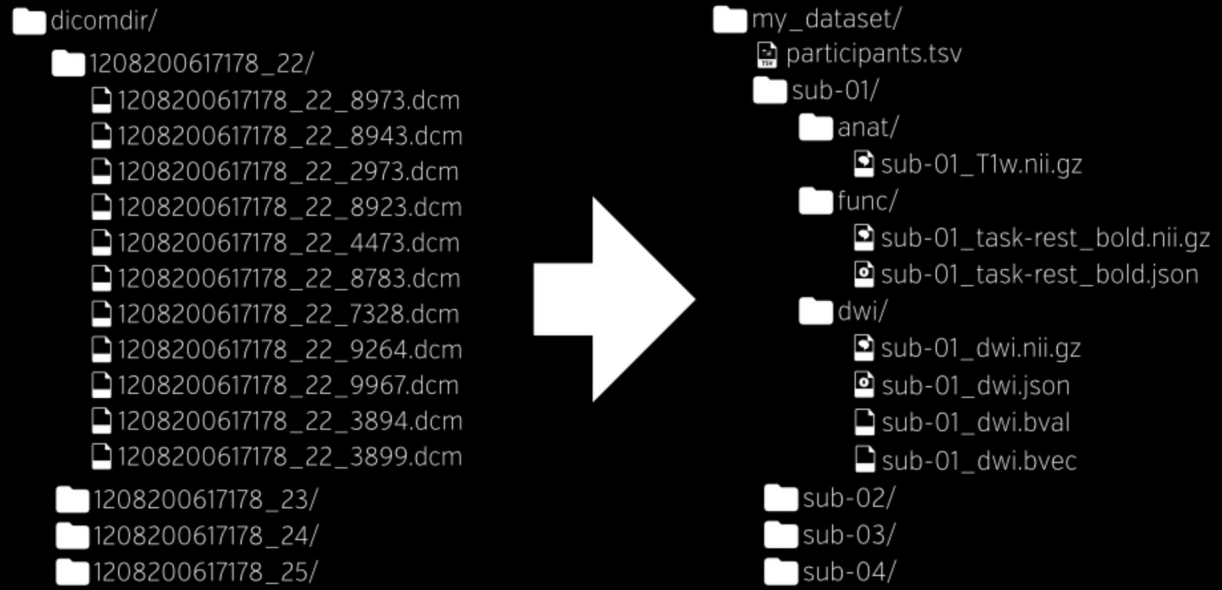
Brain Imaging Data Structure

A simple and intuitive way to organize and describe your neuroimaging and behavioral data.

ABOUT NEWS BENEFITS THE SPECIFICATION GETTING STARTED GET INVOLVED GOVERNANCE ACKNOWLEDGMENTS

About BIDS

Neuroimaging experiments result in complicated data that can be arranged in many different ways. So far there is no consensus how to organize and share data obtained in neuroimaging experiments. Even two researchers working in the same lab can opt to arrange their data in a different way. Lack of consensus (or a standard) leads to misunderstandings and time wasted on rearranging data or rewriting scripts expecting certain structure. With the Brain Imaging Data Structure (BIDS), we describe a simple and easy to adopt way of organizing neuroimaging and behavioral data.


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/

The BIDS Specification

Schematic fMRI example

```
dataset/
    participants.tsv
    dataset_description.json
    task-pieman_bold.json
    README
    CHANGES
    sub-01/
        anat/
            sub-01_t1w.nii.gz
        func/
            sub-01_task-pieman_bold.nii.gz
            sub-01_task-pieman_bold.json
            sub-01_task-pieman_events.tsv
    sub-02/
    sub-03/
    code/
    derivatives/
    stimuli/
```

The BIDS Specification

Schematic fMRI example

```
dataset/
    participants.tsv
    dataset_description.json
    task-pieman_bold.json
    README
    CHANGES
    sub-01/
        anat/
            sub-01_t1w.nii.gz
        func/
            sub-01_task-pieman_bold.nii.gz
            sub-01_task-pieman_bold.json
            sub-01_task-pieman_events.tsv
    sub-02/
    sub-03/
    code/
    derivatives/
    stimuli/
```

BIDS examples

For many lightweight example BIDS datasets, see the **bids-examples** repository on GitHub.

The BIDS Specification

Schematic fMRI example

```
dataset/
    participants.tsv
    dataset_description.json
    task-pieman_bold.json
    README
    CHANGES
    sub-01/
        anat/
            sub-01_t1w.nii.gz
        func/
            sub-01_task-pieman_bold.nii.gz
            sub-01_task-pieman_bold.json
            sub-01_task-pieman_events.tsv
    sub-02/
    sub-03/
    code/
    derivatives/
    stimuli/
```

Top-level files include tab-separated table containing participant labels and demographics, dataset description stored as key-value pairs in JSON files, and version history.

The BIDS Specification

Schematic fMRI example

```
dataset/
    participants.tsv
    dataset_description.json
    task-pieman_bold.json
    README
    CHANGES
    sub-01/
        anat/
            sub-01_t1w.nii.gz
        func/
            sub-01_task-pieman_bold.nii.gz
            sub-01_task-pieman_bold.json
            sub-01_task-pieman_events.tsv
    sub-02/
    sub-03/
    code/
    derivatives/
    stimuli/
```

Top-level files include tab-separated table containing participant labels and demographics, dataset description stored as key-value pairs in JSON files, and version history.

Subject-specific anatomical and functional NIFTI images are accompanied by JSON sidecar files describing imaging acquisition parameters, as well as tab-separated tables with event onsets.

The BIDS Specification

Schematic fMRI example

```
dataset/
    participants.tsv
    dataset_description.json
    task-pieman_bold.json
    README
    CHANGES
    sub-01/
        anat/
            sub-01_t1w.nii.gz
        func/
            sub-01_task-pieman_bold.nii.gz
            sub-01_task-pieman_bold.json
            sub-01_task-pieman_events.tsv
    sub-02/
    sub-03/
    code/
    derivatives/
    stimuli/
```

Top-level files include tab-separated table containing participant labels and demographics, dataset description stored as key-value pairs in JSON files, and version history.

Subject-specific anatomical and functional NIFTI images are accompanied by JSON sidecar files describing imaging acquisition parameters, as well as tab-separated tables with event onsets.

Loosely-structured directories for code, stimuli, and derivatives of analyses (e.g., MRIQC, fMRIPrep).

Example: OpenNeuro ds000233

Neural responses to naturalistic clips of behaving animals in two different task contexts

EDIT

uploaded by Sam Nastase on 2017-09-23 - over 2 years ago
last modified on 2019-12-14 - 6 months ago
authored by Samuel A. Nastase, Yaroslav O. Halchenko, Andrew C. Connolly, M. Ida Gobbini, James V. Haxby
81 37266

Download Analyze on brainlife.io

OpenNeuro Accession Number: ds000233
Files: 455, **Size:** 4.05GB, **Subjects:** 12, **Session:** 1
Available Tasks: beh, tax
Available Modalities: T1w, defacemask, bold, events

Nastase et al, *Front Neurosci*, 2017

The BIDS Specification

Example: OpenNeuro ds000233

Neural responses to naturalistic clips of behaving animals in two different task contexts

EDIT

uploaded by Sam Nastase on 2017-09-23 - over 2 years ago
last modified on 2019-12-14 - 6 months ago
authored by Samuel A. Nastase, Yaroslav O. Halchenko, Andrew C. Connolly, M. Ida Gobbini, James V. Haxby
81 37266

Download | Analyze on brainlife.io

OpenNeuro Accession Number: ds000233
Files: 455, **Size:** 4.05GB, **Subjects:** 12, **Session:** 1
Available Tasks: beh, tax
Available Modalities: T1w, defacemask, bold, events

Neural responses to naturalistic clips of behaving animals in two different task contexts

+ ADD FILE + ADD DIRECTORY DELETE

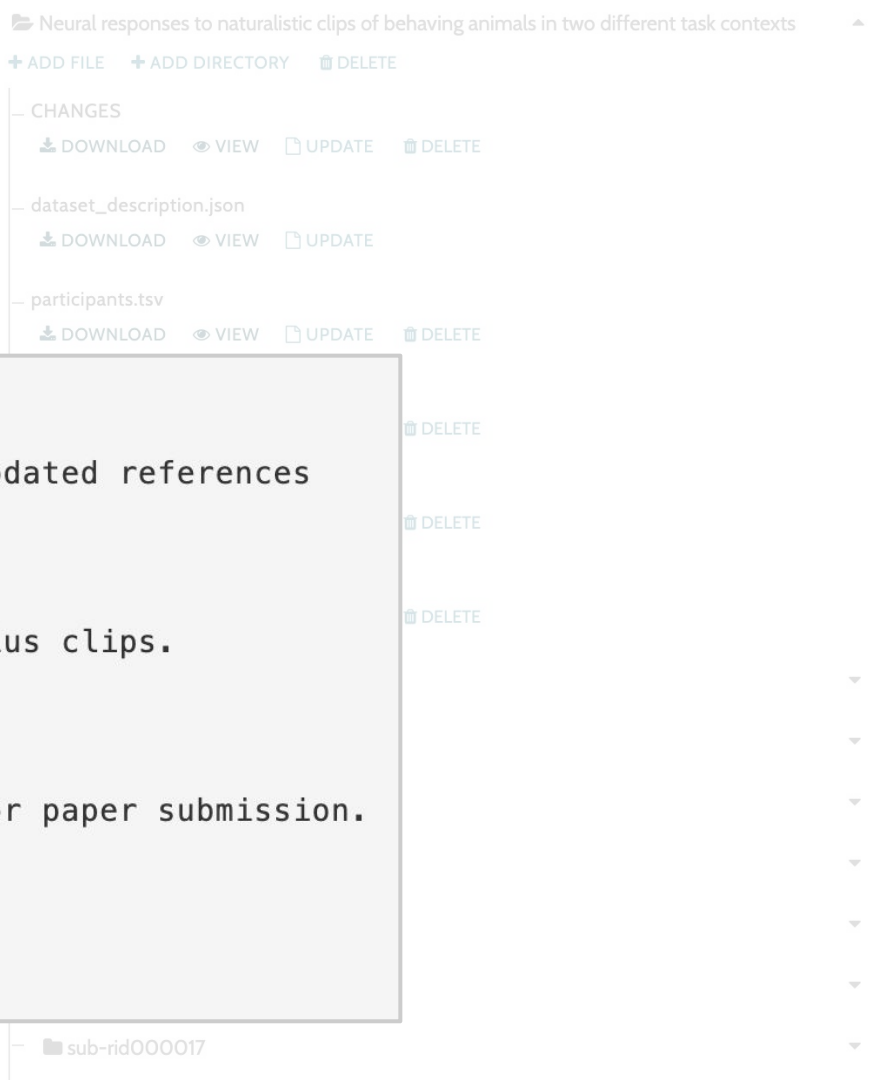
- CHANGES
  - DOWNLOAD VIEW UPDATE DELETE
- dataset_description.json
  - DOWNLOAD VIEW UPDATE
- participants.tsv
  - DOWNLOAD VIEW UPDATE DELETE
- README
  - DOWNLOAD VIEW UPDATE DELETE
- task-beh_bold.json
  - DOWNLOAD VIEW UPDATE DELETE
- task-tax_bold.json
  - DOWNLOAD VIEW UPDATE DELETE
- code
- derivatives
- sourcedata
- stimuli
- sub-rid000001
- sub-rid000012
- sub-rid000017

The BIDS Specification

Example: OpenNeuro ds000233

CHANGES

```
1.0.3   2018-05-02

  - Shared stimuli, updated *_events.tsv files, updated references

1.0.2   2018-04-12

  - Updated for paper revision; now sharing stimulus clips.

1.0.1   2018-02-20

  - Contains code, derivatives; for data descriptor paper submission.

1.0.0   2017-07-24

  - Initial release.
```



The BIDS Specification

Example: OpenNeuro ds000233

README

```
Source data (sourcedata/) may contain sensitive information and
therefore not distributed publicly. The 40 video clip stimuli
(MPEG-4 files) are distributed alongside the data in keeping with
fair use provisions for non-commercial scholarly research (see
stimuli/README.md for more information).

Note that in constructing the trial order for one run (sub-
rid0000*_task-beh_run-1_bold.nii.gz) an unintended collision
occurred where a behavior repetition event interrupted a taxonomic
repetition event. This means that one less taxonomic category
repetition occurred in this run (3 in stead of 4), and the stimulus
bird_eating_1.mp4 occurred 3 times instead of the intended 2 times
in this run.
```

The BIDS Specification

Example: OpenNeuro ds000233

participants.tsv

participant_id	age	sex	group
sub-rid000001	24	m	control
sub-rid000012	24	f	control
sub-rid000017	24	m	control
sub-rid000024	24	f	control
sub-rid000027	25	m	control
sub-rid000031	28	f	control
sub-rid000032	24	f	control
sub-rid000033	26	m	control
sub-rid000034	28	f	control
sub-rid000036	26	f	control
sub-rid000037	31	m	control
sub-rid000041	21	f	control

The BIDS Specification

Example: OpenNeuro ds000233

dataset_description.json


```
{
    "Acknowledgements": "We thank Jason Gors, Kelsey G. Wheel
er J. Swaroop Guntupalli, Matteo Visconti di Oleggio Castell
o, M. Ida Gobbini, Terry Sacket, and the rest of the DBIC (Da
rtmouth Brain Imaging Center) personnel for assistance in dat
a collection/curation.",
    "Authors": [
        "Samuel A. Nastase",
        "Yaroslav O. Halchenko",
        "Andrew C. Connolly",
        "M. Ida Gobbini",
        "James V. Haxby"
    ],
    "BIDSVersion": "1.0.2",
    "DatasetDOI": "10.18112/openneuro.ds000233.v1.0.1",
    "Funding": [
        "5R01MH075706",
        "F32MH085433-01A1",
        "NSF1129764",
        "NSF1607845"
    ],
```

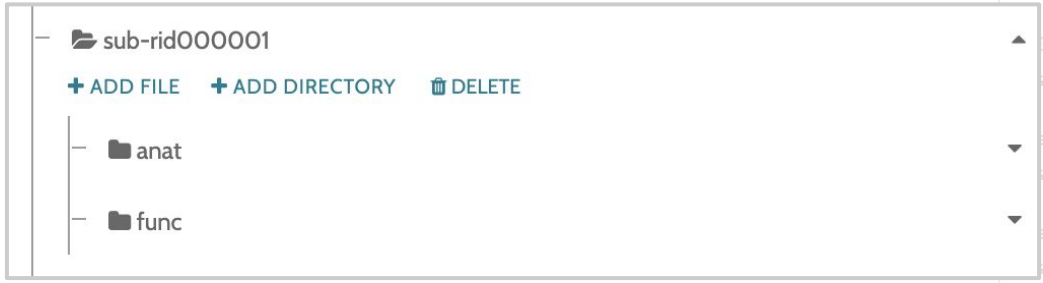
The BIDS Specification

Example: OpenNeuro ds000233

The BIDS Specification

Example: OpenNeuro ds000233

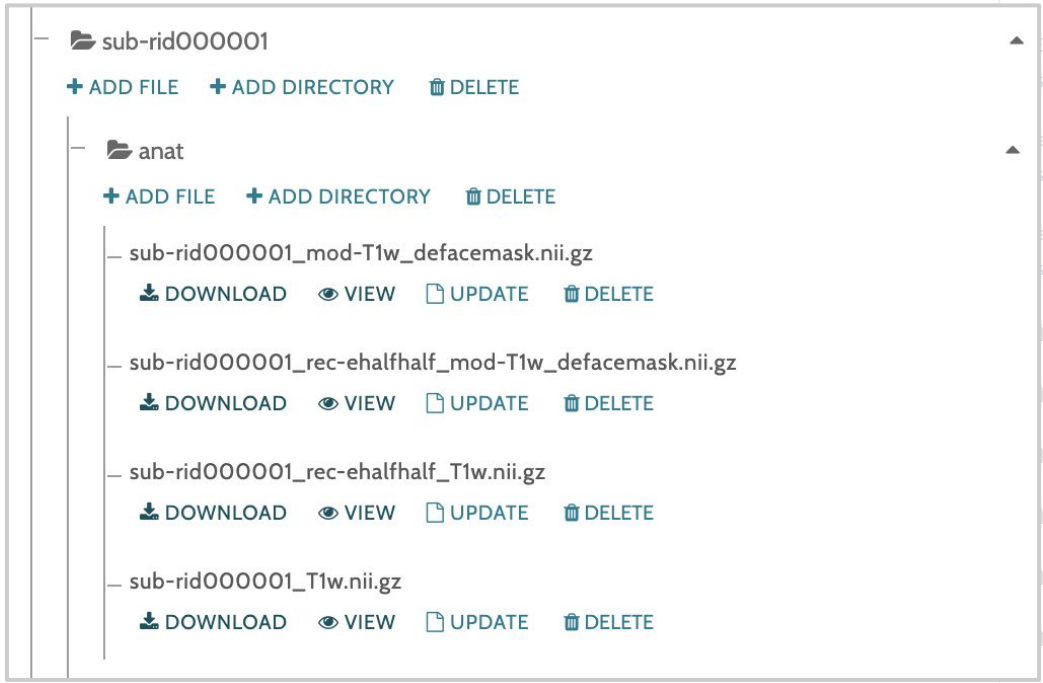
`sub-rid000001/`



The BIDS Specification

Example: OpenNeuro ds000233

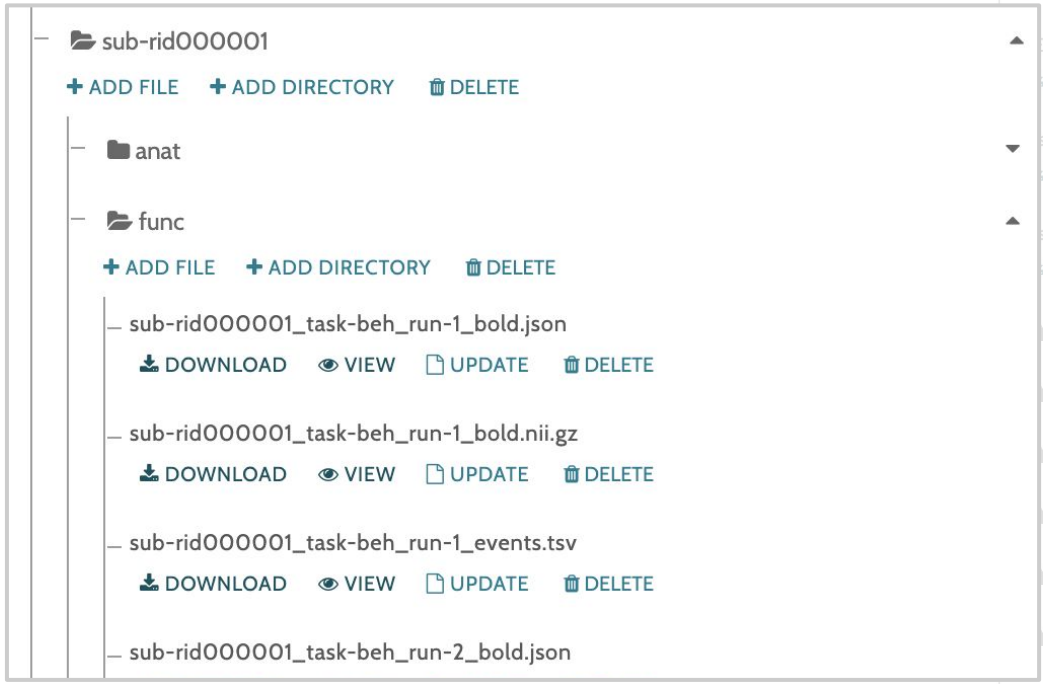
`sub-rid000001/anat/`



The BIDS Specification

Example: OpenNeuro ds000233

`sub-rid000001/func/`



The BIDS Specification

Example: OpenNeuro ds000233

sub-rid000001/func/sub-rid000001_task-beh_run-1_bold.json

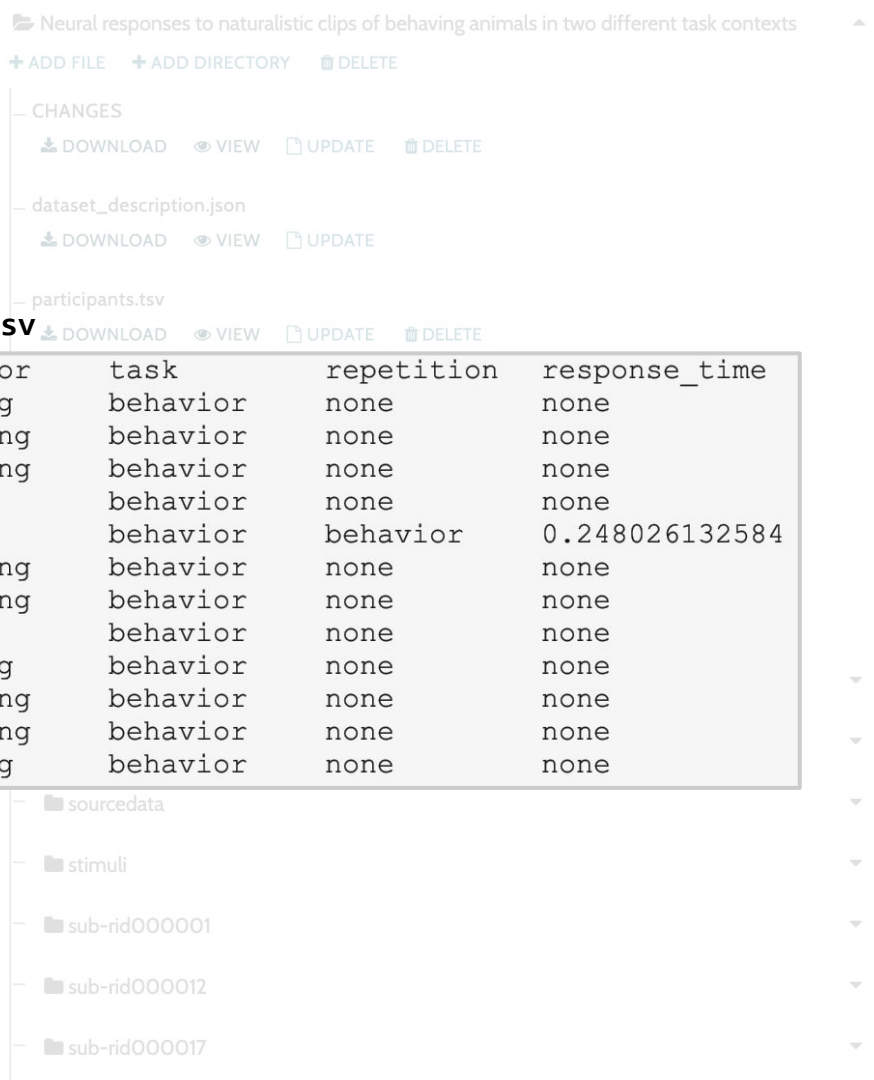

```
{
 "RepetitionTime": 2.0,
 "TaskName": "beh",
 "EchoTime": 0.035,
 "FlipAngle": 90,
 "SliceTiming": [0.000000, 0.333333, 0.666667, 1.000000, 1.33
3333, 1.666667, 0.047619, 0.380952, 0.714286, 1.047619, 1.380
952, 1.714286, 0.095238, 0.428571, 0.761905, 1.095238, 1.4285
71, 1.761905, 0.142857, 0.476190, 0.809524, 1.142857, 1.47619
0, 1.809524, 0.190476, 0.523810, 0.857143, 1.190476, 1.52381
0, 1.857143, 0.238095, 0.571429, 0.904762, 1.238095, 1.57142
9, 1.904762, 0.285714, 0.619048, 0.952381, 1.285714, 1.61904
8, 1.952381],
 "ParallelReductionFactorInPlane": 2,
 "ParellelReductionType": "SENSE",
 "Manufacturer": "Philips",
```

The BIDS Specification

Example: OpenNeuro ds000233

sub-rid000001/func/sub-rid000001_task-beh_run-1_events.tsv

onset	duration	trial_type	taxonomy	behavior	task	repetition	response_time
12.0	2.0	ungulate_running	ungulate	running	behavior	none	none
16.0	2.0	bird_fighting	bird	fighting	behavior	none	none
20.0	2.0	insect_swimming	insect	swimming	behavior	none	none
28.0	2.0	bird_eating	bird	eating	behavior	none	none
32.0	2.0	ungulate_eating	ungulate	eating	behavior	behavior	0.248026132584
36.0	2.0	primate_fighting	primate	fighting	behavior	none	none
40.0	2.0	bird_swimming	bird	swimming	behavior	none	none
44.0	2.0	ungulate_eating	ungulate	eating	behavior	none	none
48.0	2.0	bird_running	bird	running	behavior	none	none
52.0	2.0	reptile_swimming	reptile	swimming	behavior	none	none
56.0	2.0	insect_fighting	insect	fighting	behavior	none	none
60.0	2.0	primate_running	primate	running	behavior	none	none



The BIDS Specification

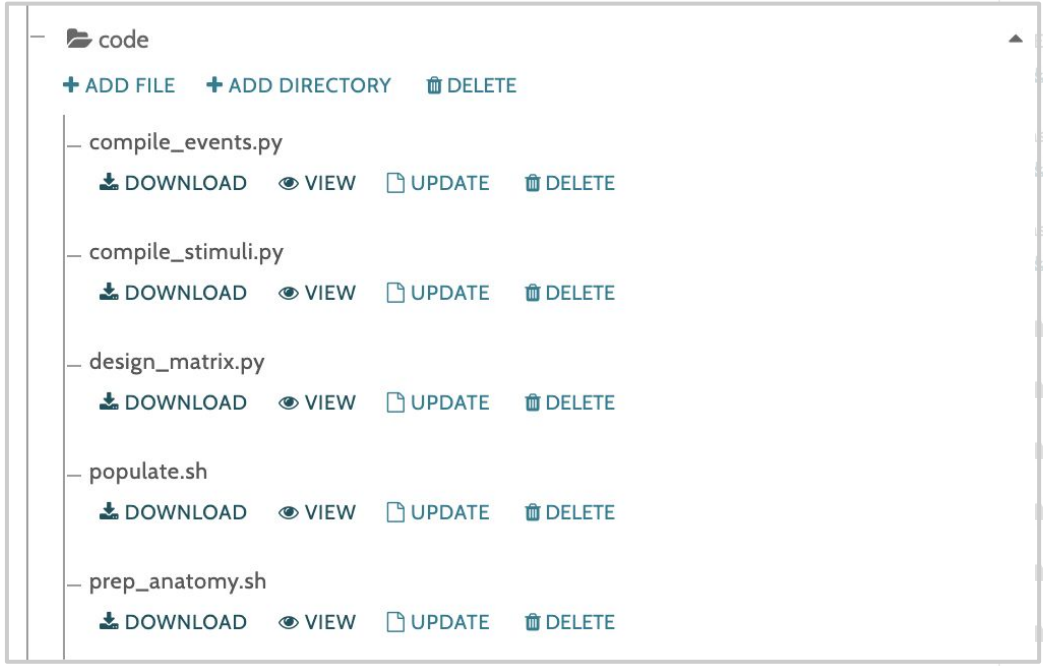
Example: OpenNeuro ds000233

`derivatives/`

The BIDS Specification

Example: OpenNeuro ds000233

code/



The BIDS Specification

Example: OpenNeuro ds000233

`stimuli/`

The BIDS Specification

Example: OpenNeuro ds000233

`stimuli/`

stimuli

+ ADD FILE + ADD DIRECTORY DELETE

bird_eating_1.mp4
DOWNLOAD VIEW UPDATE DELETE

bird_eating_2.mp4
DOWNLOAD VIEW UPDATE DELETE

bird_fighting_1.mp4
DOWNLOAD VIEW UPDATE DELETE

bird_fighting_2.mp4
DOWNLOAD VIEW UPDATE DELETE

bird_running_1.mp4
DOWNLOAD VIEW UPDATE DELETE

The BIDS Specification

Example: OpenNeuro ds000233

`stimuli/`

stimuli

+ ADD FILE + ADD DIRECTORY DELETE

bird_eating_1.mp4
DOWNLOAD VIEW UPDATE DELETE

bird_eating_2.mp4
DOWNLOAD VIEW UPDATE DELETE

bird_fighting_1.mp4
DOWNLOAD VIEW UPDATE DELETE

bird_fighting_2.mp4
DOWNLOAD VIEW UPDATE DELETE

bird_running_1.mp4
DOWNLOAD VIEW UPDATE DELETE

Neural responses to naturalistic clips of behaving animals in two different task contexts

+ ADD FILE + ADD DIRECTORY DELETE

CHANGES

DOWNLOAD VIEW UPDATE DELETE

sub-rid000017

Converting data to BIDS

The hard way...

There are many tools that can facilitate BIDS conversion:
HeuDiConv, **dcm2niix**, **PyBIDS**, **bidsify**, **bidskit**, **pyBIDSconv**, **dcm2BIDS**, etc

This is the most unpleasant part…
a single-use script manually tailored to the idiosyncrasies of each data set (☉_◎)

Converting data to BIDS

The hard way...

There are many tools that can facilitate BIDS conversion: **HeuDiConv**, **dcm2niix**, **PyBIDS**, **bidsify**, **bidskit**, **pyBIDSconv**, **dcm2BIDS**, etc.

The easy way!

Using a prespecified naming convention when creating program cards on the scanner console can allow for automated BIDS conversion—e.g. **ReproIn** (for Siemens).

This is the most unpleasant part…
a single-use script manually tailored to the idiosyncrasies of each data set (⊙_◎)

Converting data to BIDS

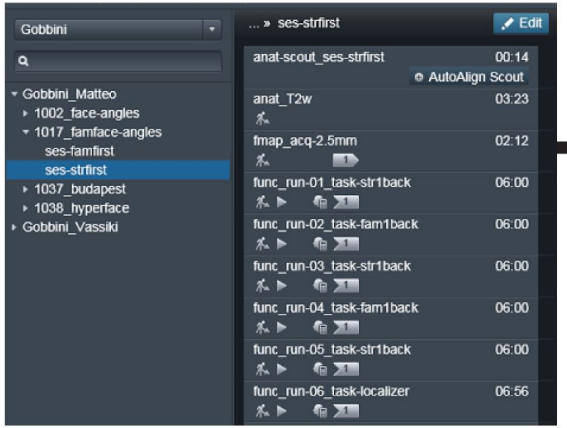
The hard way...

There are many tools that can facilitate BIDS conversion: **HeuDiConv**, **dcm2niix**, **PyBIDS**, **bidsify**, **bidskit**, **pyBIDSconv**, **dcm2BIDS**, etc.

The easy way!

Using a prespecified naming convention when creating program cards on the scanner console can allow for automated BIDS conversion—e.g. **ReproIn** (for Siemens).

Scanner



DICOM

```
001-anat-scout_ses-strfirst
005-anat_T2w
  000001.dcm
  000002.dcm
  ...
006-fmap_acq-2.5mm
007-fmap_acq-2.5mm
008-func_run-01_task-str1back
011-func_run-01_task-str1back
018-func_run-02_task-fam1back
025-func_run-03_task-str1back
032-func_run-04_task-fam1back
039-func_run-05_task-str1back
046-func_run-06_task-localizer
```

$ heudiconv

Datalad

BIDS

```
anat
  sub-sid000005_ses-strfirst_T2w.json
  sub-sid000005_ses-strfirst_T2w.nii.gz
fmap
  sub-sid000005_ses-strfirst_acq-25mm_magnitude1.json
  sub-sid000005_ses-strfirst_acq-25mm_magnitude1.nii.gz
  ...
func
  sub-sid000005_ses-strfirst_task-fam1back_run-02_bold.json
  sub-sid000005_ses-strfirst_task-fam1back_run-02_bold.nii.gz
  sub-sid000005_ses-strfirst_task-fam1back_run-02_events.tsv
  ...
sub-sid000005_ses-strfirst_scans.tsv
```

```
$ git grep TODO

CHANGES:TODOs:
README:TODO: Provide description for the dataset ...
dataset_description.json: "Acknowledgements": "TODO...",
dataset_description.json:   "TODO:",
dataset_description.json: "DatasetDOI": "TODO: ...
task-fam1back_bold.json:  "CogAtlasID": "TODO",
task-fam1back_bold.json:  "TaskName": "TODO: full task name",
task-localizer_bold.json:  "CogAtlasID": "TODO",
task-localizer_bold.json:  "TaskName": "TODO: full task name",
task-str1back_bold.json:  "CogAtlasID": "TODO",
task-str1back_bold.json:  "TaskName": "TODO: full task name",
...
```

Visconti di Oleggio Castello et al, *Zenodo*, 2020

Converting data to BIDS

The hard way...

There are many tools that can facilitate BIDS conversion: **HeuDiConv**, **dcm2niix**, **PyBIDS**, **bidsify**, **bidskit**, **pyBIDSconv**, **dcm2BIDS**, etc.

The easy way!

Using a prespecified naming convention when creating program cards on the scanner console can allow for automated BIDS conversion—e.g. **ReproIn** (for Siemens).

This is the most unpleasant part…
a single-use script manually tailored to the idiosyncrasies of each data set (☉_◎)

Moral of the story:
standardize for sharing from the start and not as an afterthought (๑>ᴗ<)ง

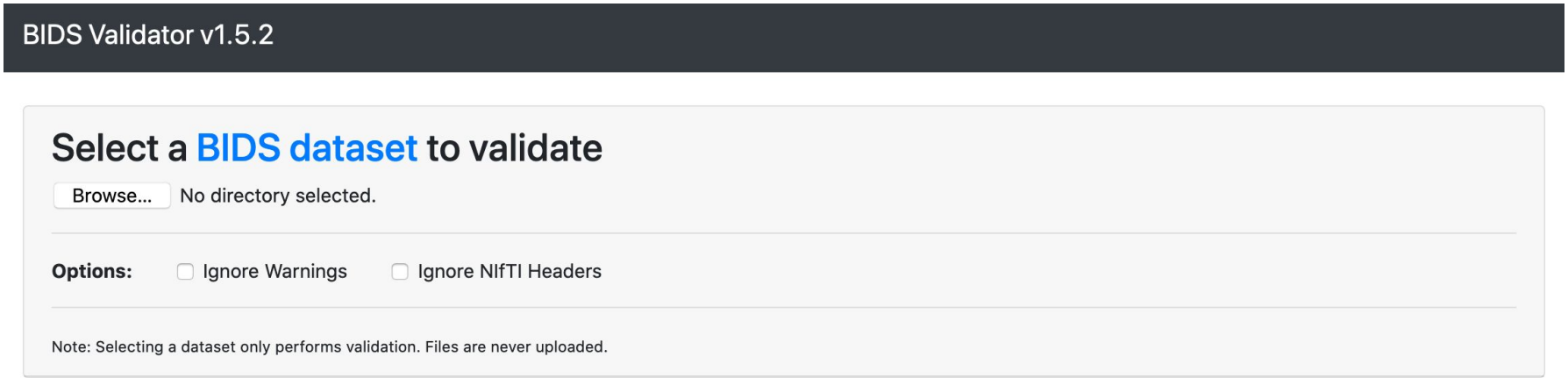
Halchenko & Hanke, *GigaScience*, 2015

BIDS Validator

Check your work

—the BIDS Validator is a lightweight tool for ensuring that your dataset is BIDS-compliant

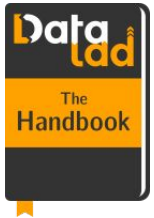
—use the **browser-based version** or run locally with via Node.js, Docker, or Python



Data provenance

Content tracking with DataLad

All stages of data processing can be version-controlled using DataLad.



http://datalad.org — discover
http://github.com/datalad/datalad — contribute
http://handbook.datalad.org — learn

```
commit ea46e14fd0dc949acd654947e136fe2c354ef780 (HEAD -> master)
Author: Sam Nastase <sam.nastase@gmail.com>
Date:   Thu Oct 17 00:40:28 2019 -0400

    Updated some missing condition labels (prettymouth)

commit a46de0940b96057d7b0297a48ab2f7a8d08a1c1e
Author: Sam Nastase <sam.nastase@gmail.com>
Date:   Tue Oct 15 18:06:08 2019 -0400

    Added audio files to stimuli subdataset

commit a1019e628f4e6a40730941d3ec467e15fbdc8560
Author: Sam Nastase <sam.nastase@gmail.com>
Date:   Tue Oct 15 15:38:40 2019 -0400

    Added subdataset for stimuli

commit e93be6b469e7e664abc2dd6c0790002a9d0171c0
Author: Sam Nastase <sam.nastase@gmail.com>
Date:   Mon Oct 14 18:45:51 2019 -0400

    Added subdatasets for code and derivatives

commit bee66631b402f4a048244c8c3cb1ebda2ad33c3a
Author: Sam Nastase <sam.nastase@gmail.com>
Date:   Mon Oct 14 18:35:07 2019 -0400

    Populated raw NIfTIs and metadata (BIDS valid)

commit daf3b89009b9bde93fabffdb2f11d2788ea44688
Author: Sam Nastase <sam.nastase@gmail.com>
Date:   Mon Oct 14 17:39:19 2019 -0400

    Instruct annex to add text files to Git

commit bc03620e50865195228e46c1ab3da3a2fa87e6cf
Author: Sam Nastase <sam.nastase@gmail.com>
Date:   Mon Oct 14 17:39:17 2019 -0400

    [DATALAD] new dataset
```

PLOS COMPUTATIONAL BIOLOGY

RESEARCH ARTICLE

BIDS apps: Improving ease of use, accessibility, and reproducibility of neuroimaging data analysis methods

Krzysztof J. Gorgolewski[1]*, Fidel Alfaro-Almagro[2], Tibor Auer[3], Pierre Bellec[4,5], Mihai Capotă[6], M. Mallar Chakravarty[7,8], Nathan W. Churchill[9], Alexander Li Cohen[10], R. Cameron Craddock[11,12], Gabriel A. Devenyi[7,8], Anders Eklund[13,14,15], Oscar Esteban[1], Guillaume Flandin[16], Satrajit S. Ghosh[17,18], J. Swaroop Guntupalli[19], Mark Jenkinson[2], Anisha Keshavan[20], Gregory Kiar[21,22], Franziskus Liem[23], Pradeep Reddy Raamana[24,25], David Raffelt[26], Christopher J. Steele[7,8], Pierre-Olivier Quirion[15], Robert E. Smith[26], Stephen C. Strother[24,25], Gaël Varoquaux[27], Yida Wang[6], Tal Yarkoni[28], Russell A. Poldrack[1]

BIDS Apps

What is a BIDS App?

—the machine-readable BIDS format enables automated processing via **BIDS Apps**

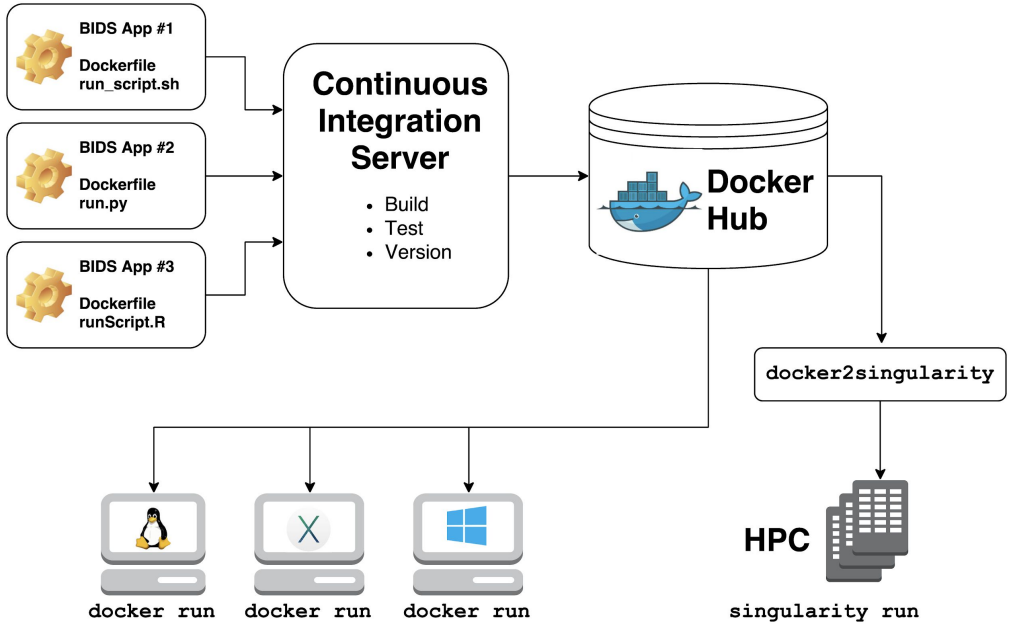
—BIDS Apps use **containerization** to facilitate portability and reproducibility

Software containers

Containerization solutions package software with all dependencies in single, encapsulated image.

Docker (for local use)

Singularity (for servers)



PLOS ONE

RESEARCH ARTICLE

MRIQC: Advancing the automatic prediction of image quality in MRI from unseen sites

Oscar Esteban[1]*, Daniel Birman[1], Marie Schaer[2], Oluwasanmi O. Koyejo[3], Russell A. Poldrack[1], Krzysztof J. Gorgolewski[1]

PLOS ONE

RESEARCH ARTI(

SCIENTIFIC DATA

OPEN

DATA DESCRIPTOR

Crowdsourced MRI quality metrics and expert quality annotations for training of humans and machines

Received: 19 September 2018
Accepted: 12 March 2019
Published online: 11 April 2019

Oscar Esteban[1], Ross W. Blair[1], Dylan M. Nielson[2], Jan C. Varada[3], Sean Marrett[3], Adam G. Thomas[2], Russell A. Poldrack[1] & Krzysztof J. Gorgolewski[1]

PLOS ONE

RESEARCH ARTI(

MRIQC:
image q

Oscar Esteban[1]
A. Poldrack[1], Kr

SCIENTIFIC DATA

OPEN

DATA DESCRIPTOR

Crowdsourced MRI quality metrics and expert quality annotations for training of humans and machines

Received: 19 September 2018
Accepted: 12 March 2019
Published online: 11 April 2019

Oscar Esteban[1], Ross W. Blair[1], Dylan M. Nielson[2], Jan C. Varada[3], Sean Marrett[3], Adam G. Thomas[2], Russell A. Poldrack[1] & Krzysztof J. Gorgolewski[1]

Getting the MRIQC Singularity image

```
singularity build mriqc-0.15.2.simg docker://poldracklab/mriqc:0.15.2
```

Ethel Franklin Betts (1908)
The Orphant Annie Book, by James Whitcomb Riley

Example derivatives: OpenNeuro ds002345

Narratives

EDIT

uploaded by Sam Nastase on 2019-12-10 - 6 months ago

last modified on 2019-12-14 - 6 months ago

authored by Samuel A. Nastase, Yun-Fei Liu, Hanna Hillman, Asieh Zadbood, Liat Hasenfratz, Neggin Keshavarzian, Janice Chen, Christopher J. Honey, Yaara Yeshurun, Mor Regev, Mai Nguyen, Claire H. C. Chang, Christopher Baldassano, Olga Lositsky, Erez Simony, Michael A. Chow, Yuan Chang Leong, Paula P. Brooks, Emily Micciche, Gina Choe, Ariel Goldstein, Yaroslav O. Halchenko, Kenneth A. Norman, Uri Hasson

33 6584

Download Analyze on brainlife.io

OpenNeuro Accession Number: ds002345

Files: 3816, **Size:** 129.67GB, **Subjects:** 315, **Session:** 1

Available Tasks: tunnel, pieman, notthefall, slumlordreach, lucy, milkyway, prettymouth, shapesphysical, shapessocial, schema, 21styear, sherlock, merlin, piemanpni, black, bronx, forgot

Available Modalities: T1w, bold, events, T2w

Nastase et al, *OpenNeuro*, 2019

Narratives

CHANGES	sub-023	sub-053	sub-083	sub-113	sub-143	sub-173	sub-203	sub-233	sub-263	sub-293
code	sub-024	sub-054	sub-084	sub-114	sub-144	sub-174	sub-204	sub-234	sub-264	sub-294
dataset_description.json	sub-025	sub-055	sub-085	sub-115	sub-145	sub-175	sub-205	sub-235	sub-265	sub-295
derivatives	sub-026	sub-056	sub-086	sub-116	sub-146	sub-176	sub-206	sub-236	sub-266	sub-296
participants.json	sub-027	sub-057	sub-087	sub-117	sub-147	sub-177	sub-207	sub-237	sub-267	sub-297
participants.tsv	sub-028	sub-058	sub-088	sub-118	sub-148	sub-178	sub-208	sub-238	sub-268	sub-298
README	sub-029	sub-059	sub-089	sub-119	sub-149	sub-179	sub-209	sub-239	sub-269	sub-299
stimuli	sub-030	sub-060	sub-090	sub-120	sub-150	sub-180	sub-210	sub-240	sub-270	sub-300
sub-001	sub-031	sub-061	sub-091	sub-121	sub-151	sub-181	sub-211	sub-241	sub-271	sub-301
sub-002	sub-032	sub-062	sub-092	sub-122	sub-152	sub-182	sub-212	sub-242	sub-272	sub-302
sub-003	sub-033	sub-063	sub-093	sub-123	sub-153	sub-183	sub-213	sub-243	sub-273	sub-303
sub-004	sub-034	sub-064	sub-094	sub-124	sub-154	sub-184	sub-214	sub-244	sub-274	sub-304
sub-005	sub-035	sub-065	sub-095	sub-125	sub-155	sub-185	sub-215	sub-245	sub-275	sub-305
sub-006	sub-036	sub-066	sub-096	sub-126	sub-156	sub-186	sub-216	sub-246	sub-276	sub-306
sub-007	sub-037	sub-067	sub-097	sub-127	sub-157	sub-187	sub-217	sub-247	sub-277	sub-307
sub-008	sub-038	sub-068	sub-098	sub-128	sub-158	sub-188	sub-218	sub-248	sub-278	sub-308
sub-009	sub-039	sub-069	sub-099	sub-129	sub-159	sub-189	sub-219	sub-249	sub-279	sub-309
sub-010	sub-040	sub-070	sub-100	sub-130	sub-160	sub-190	sub-220	sub-250	sub-280	sub-310
sub-011	sub-041	sub-071	sub-101	sub-131	sub-161	sub-191	sub-221	sub-251	sub-281	sub-311
sub-012	sub-042	sub-072	sub-102	sub-132	sub-162	sub-192	sub-222	sub-252	sub-282	sub-312
sub-013	sub-043	sub-073	sub-103	sub-133	sub-163	sub-193	sub-223	sub-253	sub-283	sub-313
sub-014	sub-044	sub-074	sub-104	sub-134	sub-164	sub-194	sub-224	sub-254	sub-284	sub-314
sub-015	sub-045	sub-075	sub-105	sub-135	sub-165	sub-195	sub-225	sub-255	sub-285	sub-315
sub-016	sub-046	sub-076	sub-106	sub-136	sub-166	sub-196	sub-226	sub-256	sub-286	
sub-017	sub-047	sub-077	sub-107	sub-137	sub-167	sub-197	sub-227	sub-257	sub-287	
sub-018	sub-048	sub-078	sub-108	sub-138	sub-168	sub-198	sub-228	sub-258	sub-288	
sub-019	sub-049	sub-079	sub-109	sub-139	sub-169	sub-199	sub-229	sub-259	sub-289	
sub-020	sub-050	sub-080	sub-110	sub-140	sub-170	sub-200	sub-230	sub-260	sub-290	
sub-021	sub-051	sub-081	sub-111	sub-141	sub-171	sub-201	sub-231	sub-261	sub-291	
sub-022	sub-052	sub-082	sub-112	sub-142	sub-172	sub-202	sub-232	sub-262	sub-292	

Narratives

28 diverse spoken story stimuli ranging from ~3 minutes to ~56 minutes (mean ≈ 10 minutes) for a total of ~5 hours of unique stimuli.

315 unique subjects (mean age = 22.4 years, range = 18–53 years; 183 reported female) participating in 788 functional scans with accompanying anatomical data.

In total, over 350,000 TRs of story-listening fMRI data with accompanying stimuli—that is, ~150 hours or 6.2 days.

PLOS ONE

RESEARCH ARTI(

MRIQC:
image q

Oscar Esteban[1]
A. Poldrack[1], Kr

SCIENTIFIC DATA

OPEN

DATA DESCRIPTOR

Crowdsourced MRI quality metrics and expert quality annotations for training of humans and machines

Received: 19 September 2018
Accepted: 12 March 2019
Published online: 11 April 2019

Oscar Esteban[1], Ross W. Blair[1], Dylan M. Nielson[2], Jan C. Varada[3], Sean Marrett[3], Adam G. Thomas[2], Russell A. Poldrack[1] & Krzysztof J. Gorgolewski[1]

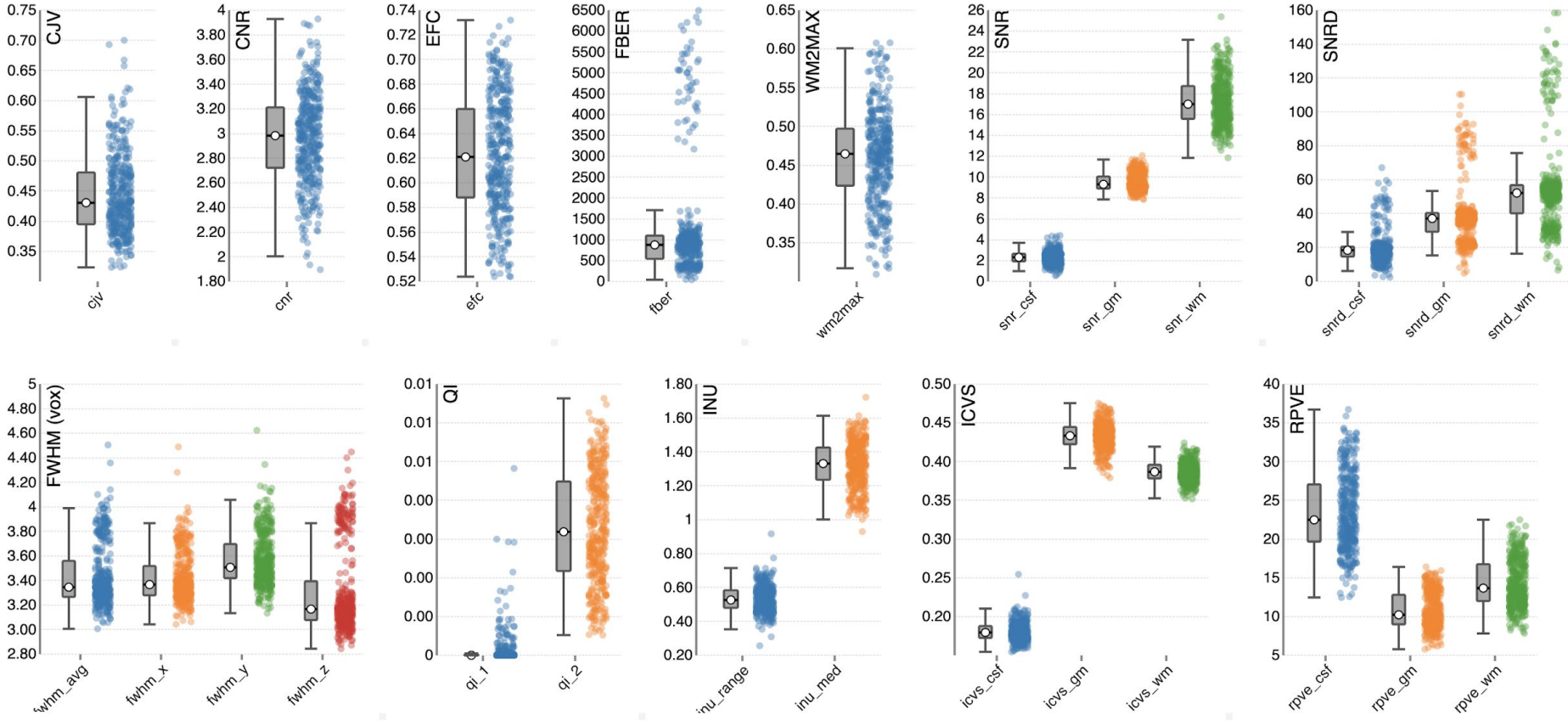
Getting the MRIQC Singularity image

```
singularity build mriqc-0.15.2.simg docker://poldracklab/mriqc:0.15.2
```

MRIQC: group T1w report

Summary

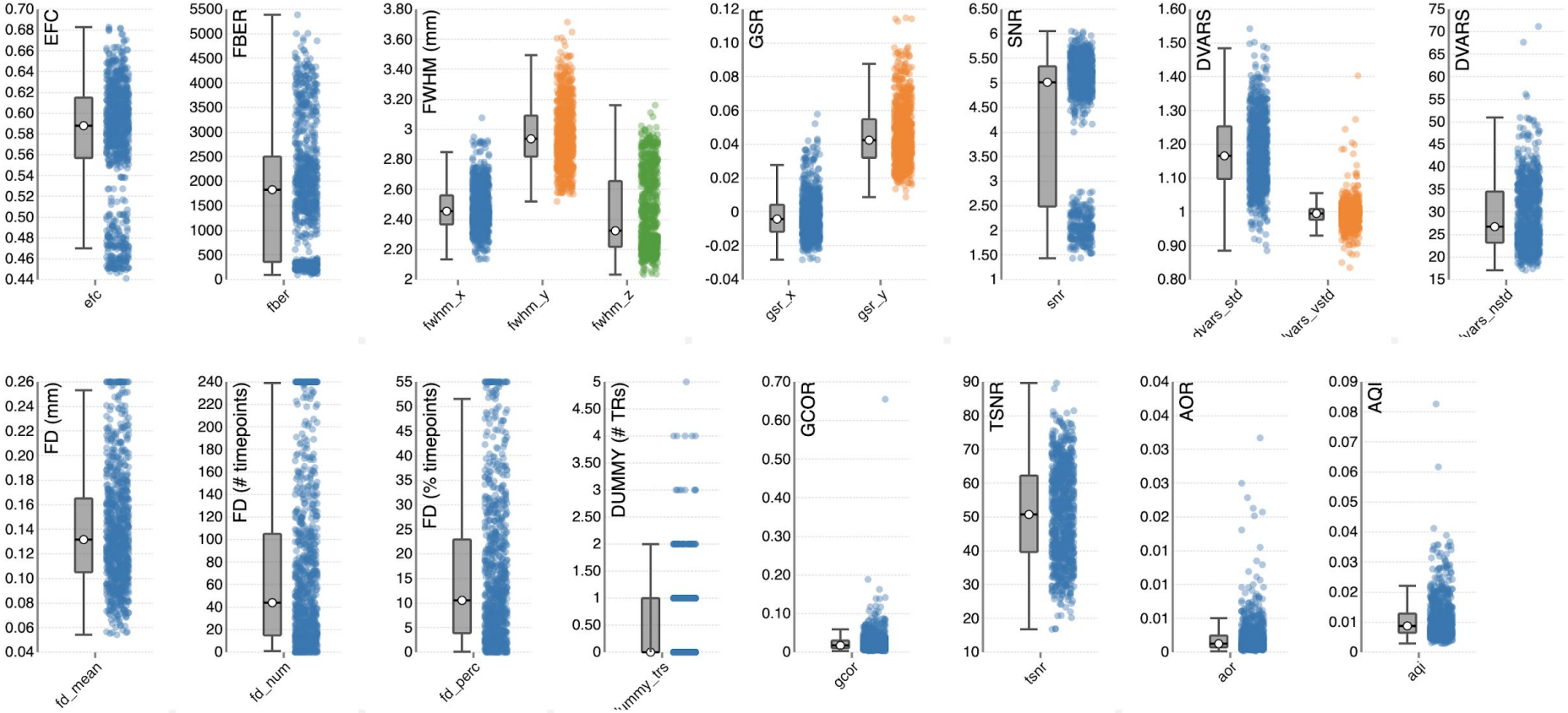
- Date and time: 2020-03-19, 19:30.
- MRIQC version: 0.15.1.



MRIQC: group bold report

Summary

- Date and time: 2020-03-19, 19:30.
- MRIQC version: 0.15.1.



nature methods

ARTICLES

https://doi.org/10.1038/s41592-018-0235-4

fMRIPrep: a robust preprocessing pipeline for functional MRI

Oscar Esteban[1*], Christopher J. Markiewicz[1], Ross W. Blair[1], Craig A. Moodie[1], A. Ilkay Isik[2], Asier Erramuzpe[3], James D. Kent[4], Mathias Goncalves[5], Elizabeth DuPre[6], Madeleine Snyder[7], Hiroyuki Oya[8], Satrajit S. Ghosh[5,9], Jessey Wright[1], Joke Durnez[1], Russell A. Poldrack[1,10] and Krzysztof J. Gorgolewski[1,10*]

nature methods

ARTICLES

https://doi.org/10.1038/s41592-018-0235-4

fMRIPrep: functional

Oscar Esteban[1]*, C
Asier Erramuzpe[3],
Hiroyuki Oya[8], Satraj
Krzysztof J. Gorgolev

nature protocols

PROTOCOL

https://doi.org/10.1038/s41596-020-0327-3

Check for updates

Analysis of task-based functional MRI data preprocessed with fMRIPrep

Oscar Esteban[1]✉, Rastko Ciric[1], Karolina Finc[2], Ross W. Blair[1], Christopher J. Markiewicz[1], Craig A. Moodie[1], James D. Kent[3], Mathias Goncalves[4], Elizabeth DuPre[5], Daniel E. P. Gomez[6], Zhifang Ye[7], Taylor Salo[8], Romain Valabregue[9], Inge K. Amlien[10], Franziskus Liem[11], Nir Jacoby[12], Hrvoje Stojić[13], Matthew Cieslak[14], Sebastian Urchs[5], Yaroslav O. Halchenko[15], Satrajit S. Ghosh[4,16], Alejandro De La Vega[17], Tal Yarkoni[17], Jessey Wright[1], William H. Thompson[1,18], Russell A. Poldrack and Krzysztof J. Gorgolewski[1]

Getting the fMRIPrep Singularity image

```
singularity build fmriprep-20.1.0.simg docker://poldracklab/fmriprep:20.1.0
```

fMRIPrep

“a functional magnetic resonance imaging (fMRI) data preprocessing pipeline that is designed to provide an easily accessible, state-of-the-art interface that is robust to variations in scan acquisition protocols and that requires minimal user input, while providing easily interpretable and comprehensive error and output reporting”

Simple command-line interface (positional arguments):

`fmriprep`	`bids_folder/`	`output_folder/`	`participant`	analysis level: `participant` (or `group`)
				output directory; e.g., `derivatives/`
				input directory; e.g., `dataset/` (BIDS formatted)
				fMRIPrep executable

fMRIPrep

“a functional magnetic resonance imaging (fMRI) data preprocessing pipeline that is designed to provide an easily accessible, state-of-the-art interface that is robust to variations in scan acquisition protocols and that requires minimal user input, while providing easily interpretable and comprehensive error and output reporting”

Full command line-interface using Singularity:

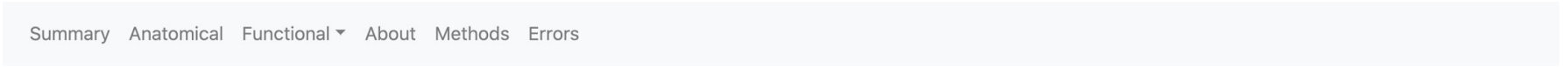
`singularity run --cleanenv \` `  --bind bids_folder/:/data \`	Singularity arguments
`  /home/snastase/singularity/fmriprep-20.1.0.simg \`	fMRIPrep Singularity image

`  --participant-label sub-001 \`
`  --nthreads 8 --omp-nthreads 8 \`
`  --output-spaces T1w fsaverage6 MNI152NLin2009cAsym \`
`  --use-syn-sdc --write-graph \`
`  --fs-license-file /data/code/license.txt \`
`  --work-dir /data/derivatives/work \`
	fMRIPrep keyword arguments

`  /data /data/derivatives participant`	fMRIPrep positional arguments

fMRIPrep

“a functional magnetic resonance imaging (fMRI) data preprocessing pipeline that is designed to provide an easily accessible, state-of-the-art interface that is robust to variations in scan acquisition protocols and that requires minimal user input, while providing easily interpretable and comprehensive error and output reporting”



Summary

- Subject ID: 001
- Structural images: 1 T1-weighted
- Functional series: 3
  - Task: pieman (2 runs)
  - Task: tunnel (1 run)
- Standard output spaces: fsaverage, MNI152NLin2009cAsym, MNI152NLin6Asym
- Non-standard output spaces: anat, func
- FreeSurfer reconstruction: Run by fMRIPrep

Brain Imaging Data Structure (BIDS)

Data standardization is the linchpin for reproducible neuroimaging

BIDS is a community-driven standard for organizing neuroimaging data that facilitates sharing and automated processing. BIDS Apps are portable software containers that capitalize on the BIDS format to reproducibly analyze data with minimal manual intervention.

Accompanying materials

GitHub: **ohbm-traintrack-bids**

Zenodo: **DOI**

Other resources

Stanford BIDS tutorials: **part 1a**, **part 1b**, **part 2a**

Other BIDS presentations: **OSF**