Brain Imaging Data Organisation

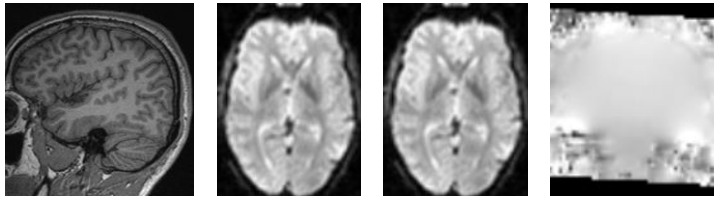
Dace Apšvalka
[Datza]

COGNESTIC, 2025

Complexity of Brain Imaging Data

- A single study often involves:
  - Multiple imaging protocols
  - Many participants, sometimes across sessions
- Produces numerous files in diverse formats:
  - From simple text files (logs, metadata)
  - To large multidimensional images
- Datasets become large, diverse, and complex

Complexity of Brain Imaging Data



Subject	Trial	Event Type	Code	trial_code	Time	T
197602009f	0	Pulse	100		1280995	
197602009f	1	Picture	nBack Intr	nBack Intr	1281027	
197602009f	1	Pulse	100		1298915	
197602009f	1	Pulse	100		1316915	
197602009f	2	Picture	Intro 0 n-B	Intro 0 n-B	1321861	
197602009f	2	Pulse	100		1334995	
197602009f	2	Pulse	100		1352915	
197602009f	3	Picture	T 10,5,No	T 10	1361195	
197602009f	3	Pulse	100		1370915	
197602009f	3	Response	2		1372643	
197602009f	3	Pulse	100		1388995	
197602009f	4	Picture	T 10,5,No	T 10	1393329	
197602009f	4	Response	2		1397415	
197602009f	4	Pulse	100		1406915	
197602009f	5	Picture	T 10,2,No	T 10	1423863	
197602009f	5	Pulse	100		1424915	

```
scans/
├── subj01/
│   ├── s01_t1w.nii.gz
│   ├── s01_bold1_run1.nii.gz
│   ├── s01_bold1_run2.nii.gz
│   └── s01_fieldmap.nii.gz
├── subj02/
│   ├── s02_t1w.nii.gz
│   ├── s02_bold1_run1.nii.gz
│   ├── s02_bold1_run2.nii.gz
│   └── s02_fieldmap.nii.gz
...
├── subj100/
│   ├── s100_t1w.nii.gz
│   ...
logs/
├── subj01/
│   ├── events_taskA_run1.tsv
│   ├── events_taskA_run2.tsv
├── subj02/
│   ├── events_taskA_run1.tsv
│   ├── events_taskA_run2.tsv
...
├── subj100/
│   ...
protocols/
...
```

and even more files
(more modalities, sessions, tasks)

Complexity of Brain Imaging Data

```
scans/
├── subj01/
│   ├── s01_t1w.nii.gz
│   ├── s01_bold1_run1.nii.gz
│   ├── s01_bold1_run2.nii.gz
│   └── s01_fieldmap.nii.gz
├── subj02/
│   ├── s02_t1w.nii.gz
│   ├── s02_bold1_run1.nii.gz
│   ├── s02_bold1_run2.nii.gz
│   └── s02_fieldmap.nii.gz
...
├── subj100/
│   ├── s100_t1w.nii.gz
...
logs/
├── subj01/
│   ├── events_taskA_run1.tsv
│   ├── events_taskA_run2.tsv
├── subj02/
│   ├── events_taskA_run1.tsv
│   ├── events_taskA_run2.tsv
...
├── subj100/
│   ...
protocols/
...
```

Many possible ways to name and organise the data

(more modalities, sessions, tasks)

Inconsistent Data Organisation

- Difficult for others (and you!) to understand data and track changes
- Scripts must be adapted, can’t be easily reused
- Huge effort to automate workflows, no way to automatically validate datasets
- Increased risk of errors (wrong files, outdated versions)
- Time wasted searching and reorganising
- Harder to reproduce results and collaborate

Inconsistent Data Organisation

- Difficult for others (and you!) to understand data and track changes
- Scripts must be adapted, can't be easily reused
- Huge effort to automate workflows, no way to automatically validate datasets
- Increased risk of errors (wrong files, outdated versions)
- Time wasted searching and reorganising
- Harder to reproduce results and collaborate

Wouldn't it be much easier if everybody organised the files in the same way?

Brain **I**maging **D**ata **S**tructure (BIDS)

- A standard for organising data and metadata across various neuroscience modalities (MRI, MEG, EEG, PET)
- Enables easier sharing, reuse, and application of automated pipelines and quality assurance protocols

Brain **I**maging **D**ata **S**tructure (BIDS)

- A standard for organising data and metadata across various neuroscience modalities (MRI, MEG, EEG, PET)
- Enables easier sharing, reuse, and application of automated pipelines and quality assurance protocols

BIDS
BRAIN IMAGING DATA STRUCTURE

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

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]

```
sub-15
└── ses-mri
    ├── anat
    │   ├── sub-15_ses-mri_T1w.json
    │   └── sub-15_ses-mri_T1w.nii.gz
    ├── fmap
    │   ├── sub-15_ses-mri_acq-func_magnitude1.json
    │   ├── sub-15_ses-mri_acq-func_magnitude1.nii.gz
    │   ├── sub-15_ses-mri_acq-func_magnitude2.json
    │   ├── sub-15_ses-mri_acq-func_magnitude2.nii.gz
    │   ├── sub-15_ses-mri_acq-func_phasediff.json
    │   └── sub-15_ses-mri_acq-func_phasediff.nii.gz
    ├── func
    │   ├── sub-15_ses-mri_task-facerecognition_run-01_bold.json
    │   ├── sub-15_ses-mri_task-facerecognition_run-01_bold.nii.gz
    │   ├── sub-15_ses-mri_task-facerecognition_run-01_events.tsv
    │   ├── sub-15_ses-mri_task-facerecognition_run-02_bold.json
    │   ├── sub-15_ses-mri_task-facerecognition_run-02_bold.nii.gz
    │   ├── sub-15_ses-mri_task-facerecognition_run-02_events.tsv
    │   ├── sub-15_ses-mri_task-facerecognition_run-03_bold.json
    │   ├── sub-15_ses-mri_task-facerecognition_run-03_bold.nii.gz
    │   ├── sub-15_ses-mri_task-facerecognition_run-03_events.tsv
    │   ├── sub-15_ses-mri_task-facerecognition_run-04_bold.json
    │   ├── sub-15_ses-mri_task-facerecognition_run-04_bold.nii.gz
    │   ├── sub-15_ses-mri_task-facerecognition_run-04_events.tsv
    │   ├── sub-15_ses-mri_task-facerecognition_run-05_bold.json
    │   ├── sub-15_ses-mri_task-facerecognition_run-05_bold.nii.gz
    │   ├── sub-15_ses-mri_task-facerecognition_run-05_events.tsv
    │   ├── sub-15_ses-mri_task-facerecognition_run-06_bold.json
    │   ├── sub-15_ses-mri_task-facerecognition_run-06_bold.nii.gz
    │   ├── sub-15_ses-mri_task-facerecognition_run-06_events.tsv
    │   ├── sub-15_ses-mri_task-facerecognition_run-07_bold.json
    │   ├── sub-15_ses-mri_task-facerecognition_run-07_bold.nii.gz
    │   ├── sub-15_ses-mri_task-facerecognition_run-07_events.tsv
    │   ├── sub-15_ses-mri_task-facerecognition_run-08_bold.json
    │   ├── sub-15_ses-mri_task-facerecognition_run-08_bold.nii.gz
    │   ├── sub-15_ses-mri_task-facerecognition_run-08_events.tsv
    │   ├── sub-15_ses-mri_task-facerecognition_run-09_bold.json
    │   ├── sub-15_ses-mri_task-facerecognition_run-09_bold.nii.gz
    │   └── sub-15_ses-mri_task-facerecognition_run-09_events.tsv
    └── sub-15_ses-mri_scans.tsv
```

Today’s plan

1

MRI Data Organisation

BIDS
BRAIN IMAGING DATA STRUCTURE

2

MRI Data Manipulation

```
((([[ 0.,  0.,  0., ...,  0.,  0.,  0.],
   [ 0.,  0.,  0., ...,  0.,  0.,  0.],
   [ 0.,  0.,  0., ...,  0.,  0.,  0.],
   ...,
   [ 0.,  0.,  0., ...,  0.,  0.,  0.],
   [ 0.,  0.,  0., ...,  0.,  0.,  0.],
   [ 0.,  0.,  0., ...,  0.,  0.,  0.]],

  [[ 0.,  0.,  0., ...,  0.,  0.,  0.],
   [ 0., 25., 23., ..., 23., 32.,  0.],
   [ 0., 28., 21., ..., 25., 25.,  0.],
   ...,
   [ 0., 26., 24., ..., 40., 20.,  0.],
   [ 0., 44., 28., ..., 30., 21.,  0.],
   [ 0.,  0.,  0., ...,  0.,  0.,  0.]],

  [[ 0.,  0.,  0., ...,  0.,  0.,  0.],
   [ 0., 28., 26., ..., 31., 29.,  0.],
   [ 0., 32., 30., ..., 22., 21.,  0.],
   ...,
   [ 0., 27., 24., ..., 31., 30.,  0.],
   [ 0., 30., 23., ..., 37., 22.,  0.],
   [ 0.,  0.,  0., ...,  0.,  0.,  0.]],

  ...,
```