

# *Open science methods: lessons from neuroscience*

***Tomas Knapen***

*VU Cognitive Psychology & Spinoza Centre, KNAW*

# Reproducible Science

- How to make science reproducible?
  - Work on practices:
    - no HARKing, don't keep testing until significant, preregister(?), etc.
    - Teach people Statistical/Reasoning literacy
  - Make everything check-able, meaning: **Open**:
    - Don't keep data/methods/papers to yourselves
    - Teach people **Open Science** skills and ways of thinking

*Trust me, it checks out!*

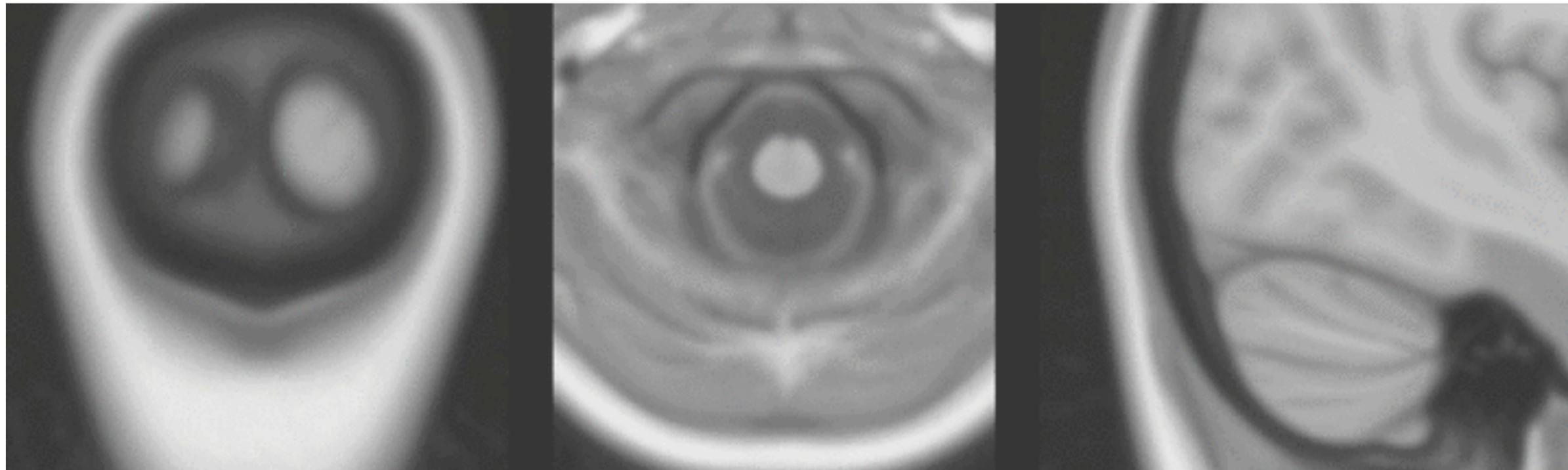
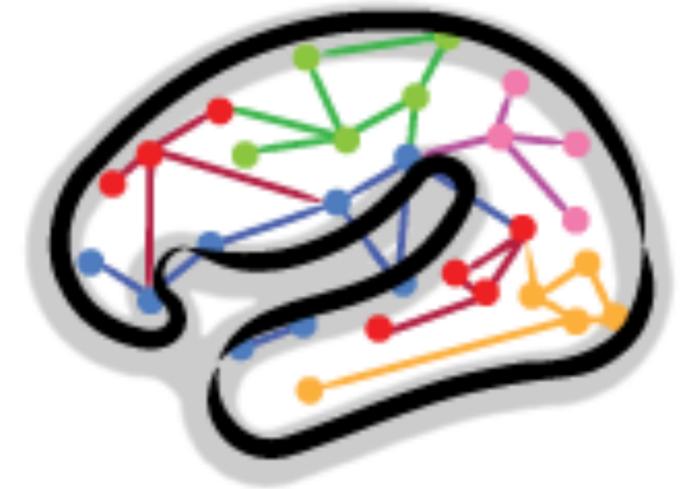
# *The start:* Open data

- Data are the bedrock of scientific results, of course.
- Save your data in a common format (not your own!), extra work!  
Many people see it as a chore
- Many people feel ownership of their data:  
*I'm not giving that away!*
- But it's not all bad:
  - It's becoming more common to share data.
  - With soo much data at your fingertips, you become science Superman/Woman!



# *An example use of open data*

- Human Connectome Project:
  - 1200 participants scanned multiple times:
    - Tasks, Resting State, Movie Watching, Retinotopy, Anatomy, DTI
  - All data freely available



- ***We discovered 5 retinotopic maps of visual space in the cerebellum.***
- Only possible by averaging 181 subjects, otherwise too hard to find: need multiple hours of data for single subject results...



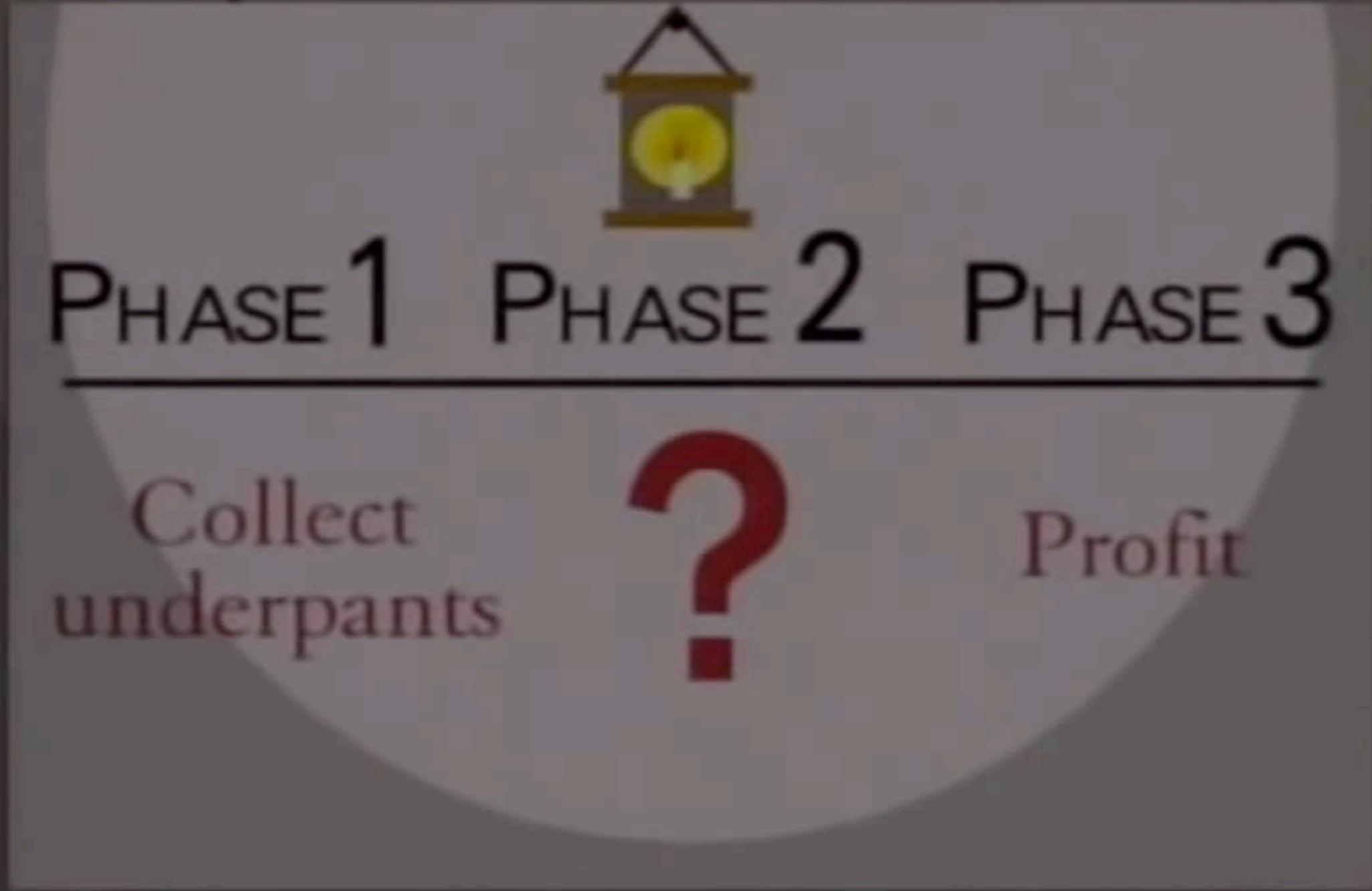


**C  
D**



**n  
hing**

Ho



...

en

ishing

# How to get from open data to open papers...

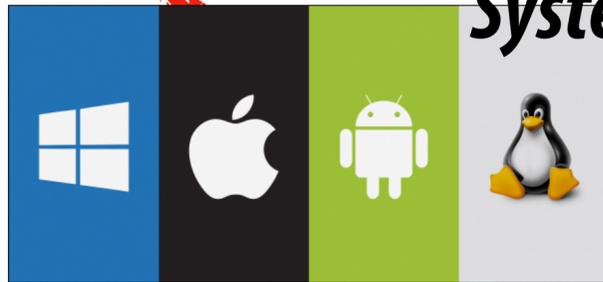
**All steps between data and publication present opportunities for non-reproducible outcomes**



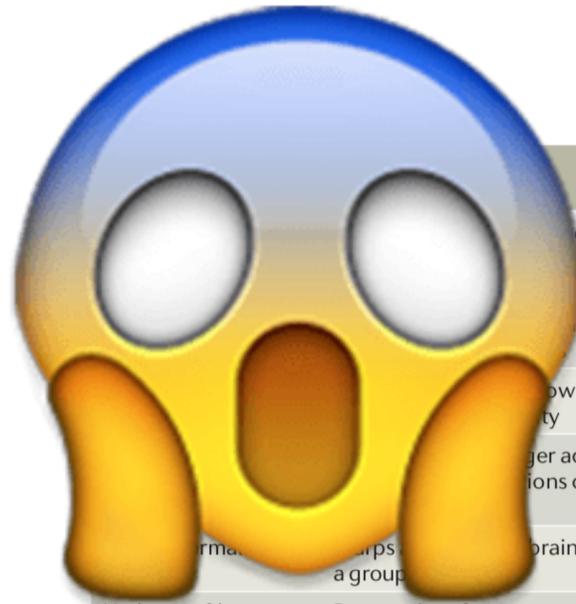
**Computer Hardware**



**Operating System**



**Preprocessing**



		Options [suboptions]	Number of plausible options
	during	• 'Interpolation' [linear or sinc] • 'Reference volume' [single or mean]	4
	acquisition	'No', 'before motion correction' or 'after motion correction'	3
	owing to	'Yes' or 'no'	2
	er activations	'FV	
	ions of GRF		
	ormat	ips	rain to match
	a group		
High-pass filter	Remove low-frequency nuisance signals from data	'Fre	120
Head motion regressors	Remove remaining signals owing to head motion via statistical model	'Ye	12
Haemodynamic response	Account for delayed nature of haemodynamic response to neuronal activity	• 'E	o
		• 'Derivatives' ['none', 'shift' or 'dispersion']	
Temporal autocorrelation model	Model for the temporal autocorrelation inherent in fMRI signals	'Yes' or 'no'	2
Multiple-comparison correction	Correct for large number of comparisons across the brain	'Voxel-based GRF', 'cluster-based GRF', 'FDR' or 'non-parametric'	4
Total possible workflows			69,120

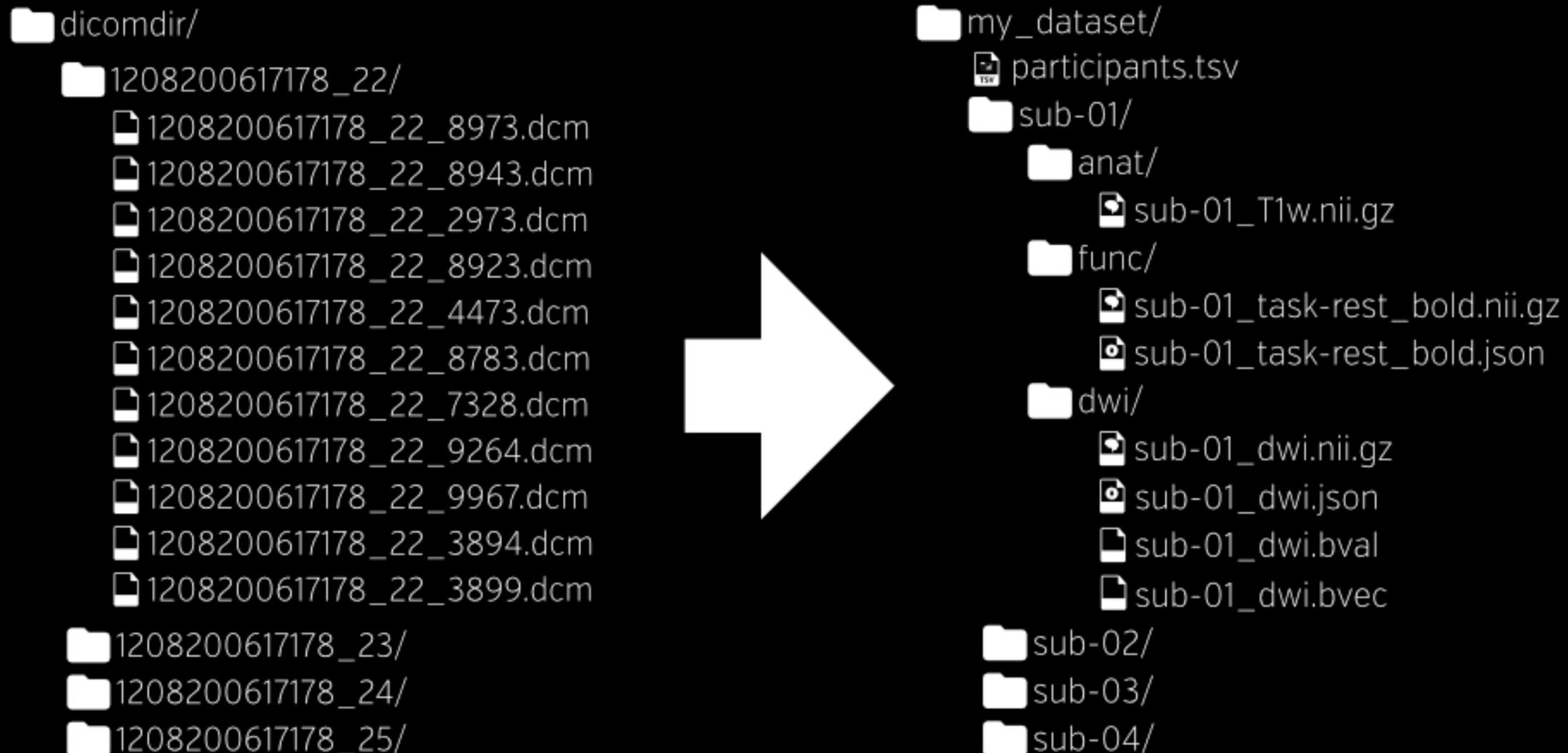
```
1 #!/usr/bin/env python
2 import sys
3 import os
4 import simpleknn
5 from bigfile import BigFile
6
7 if __name__ == "__main__":
8     trainCollection = 'toydata'
9     nimages = 2
10    feature = 'f1'
11    dim = 3
12
13    testCollection = trainCollection
14    testset = testCollection
15
16    featureDir = os.path.join(rootpath, trainCollect
```

**Your own manual data analysis**

**Your own automated data analysis**

# Brain Imaging Data Structure

- Standard format for organising neuroimaging data



# [bids.neuroimaging.io](https://bids.neuroimaging.io)

- Once you create a BIDS data format, you need to make sure it ***conforms to the specification***
- Run it through the validator!

[incf.github.io/bids-validator](https://incf.github.io/bids-validator)

The screenshot displays the BIDS Validator interface. At the top, there are three columns: 'Summary', 'Available Tasks', and 'Available Modalities'. The 'Summary' column lists: 40 Files, 18.42kB; 13 - Subjects; and 1 - Session. The 'Available Tasks' column lists: rhyme judgment. The 'Available Modalities' column lists: bold and T1w. Below this, a message states: 'Your dataset is not a valid BIDS dataset.' At the bottom, there are two buttons: a red one labeled 'view 1 error in 23 files' and a yellow one labeled 'view 1 warning in 4 files'.

Summary	Available Tasks	Available Modalities
<ul style="list-style-type: none"><li>• 40 Files, 18.42kB</li><li>• 13 - Subjects</li><li>• 1 - Session</li></ul>	<ul style="list-style-type: none"><li>• rhyme judgment</li></ul>	<ul style="list-style-type: none"><li>• bold</li><li>• T1w</li></ul>

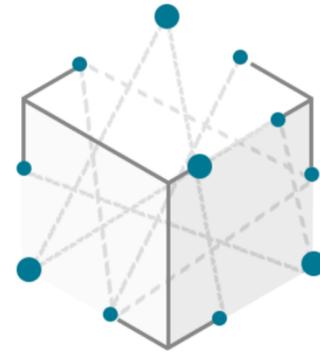
Your dataset is not a valid BIDS dataset.

[view 1 error in 23 files](#)

[view 1 warning in 4 files](#)

# BIDS for sharing

- These validated datasets can easily be shared online: [openneuro.org](https://openneuro.org)



## OpenNEURO

A free and open platform for sharing MRI, MEG, EEG, iEEG, and ECoG data



Sign in with Google

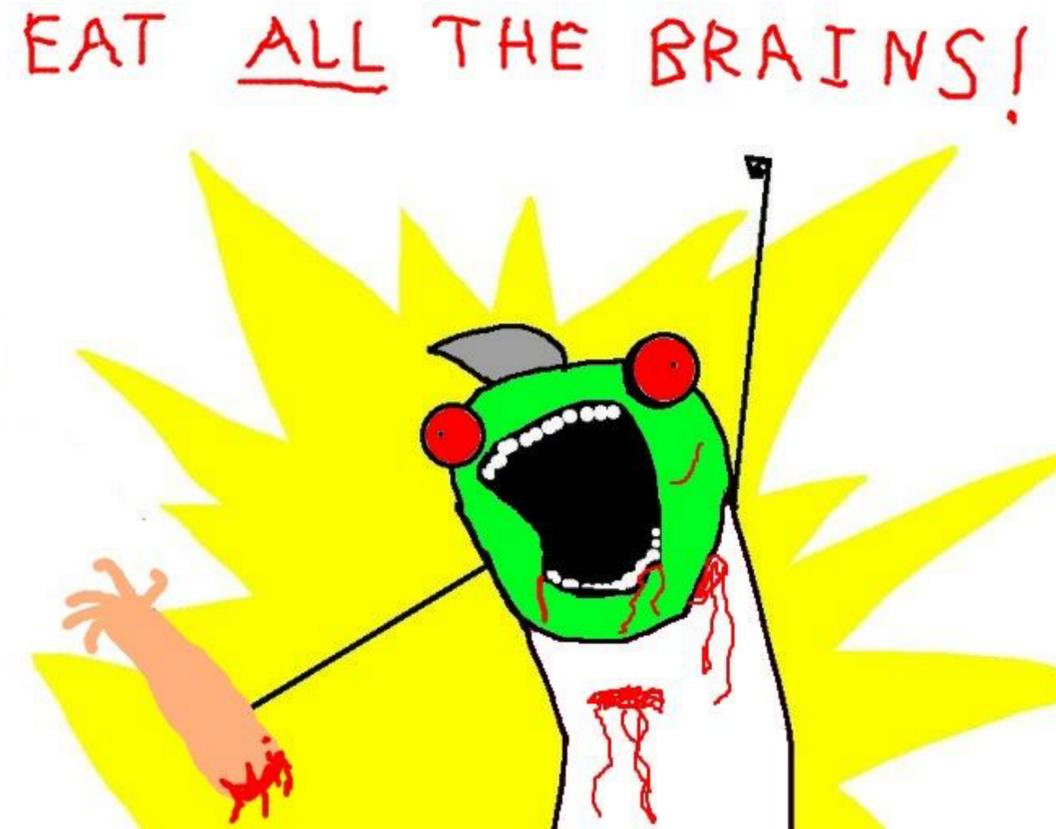


Sign in with ORCID

[Browse 247 Public Datasets](#)

# ***Now that we've got BIDS what are we going to do with it?***

- Automate everything!
- BIDS apps:
  - Operating system containers (no need to install software, run anywhere).
  - Single standardised command line interface.



# Available BIDS Apps

BIDS-Apps/example	version 0.0.7	open bug issues 0	build passing	open bug pull requests 0	docker pulls 8k	439.5MB	23 layers
BIDS-Apps/freesurfer	version v6.0.1-3	open bug issues 0	build failed		docker pulls 3k	2.6GB	52 layers
BIDS-Apps/ndmg	version v0.1.0	open bug issues 0	build passing		docker pulls 7k	920.9MB	31 layers
BIDS-Apps/BROCCOLI	version v1.0.1	open bug issues 1	build passing	open bug pull requests 0	docker pulls 256	3GB	21 layers
BIDS-Apps/FibreDensityAndCrosssection	version v0.0.1	open bug issues 0	build passing	open bug pull requests 0	docker pulls 72	576.8MB	31 layers
BIDS-Apps/SPM	version v0.0.14	open bug issues 0	build passing	open bug pull requests 0	docker pulls 872	1.6GB	24 layers
poldracklab/mriqc	version 0.10.3		build passing	open bug pull requests 0	docker pulls 16k	2.6GB	37 layers
BIDS-Apps/QAP	Image not found	open bug issues 0	build passing		docker pulls 7	Image not found	
BIDS-Apps/CPAC	version v1.0.2_di...		build passing	open bug pull requests 0	docker pulls 2k	1.4GB	38 layers
BIDS-Apps/hyperalignment	Image not found		build passing		docker pulls 3	Image not found	
BIDS-Apps/mindboggle	version 0.0.4-1	open bug issues 2	build passing		docker pulls 376	1.9GB	81 layers
BIDS-Apps/MRtrix3_connectome	version 0.2.2	open bug issues 0	build passing	open bug pull requests 0	docker pulls 345	3.4GB	56 layers
BIDS-Apps/rs_signal_extract	version 0.1		build passing	open bug pull requests 0	docker pulls 75	240MB	17 layers
BIDS-Apps/aa	version enh_vario...		build failed		docker pulls 61	3.8GB	57 layers
BIDS-Apps/niak	version latest		build passing		docker pulls 110	2.7GB	103 layers
BIDS-Apps/oppni	version v0.7.0-1		build passing		docker pulls 137	2.9GB	41 layers
poldracklab/fmriprep	version 1.0.8		build passing	open bug pull requests 0	docker pulls 30k	4.4GB	48 layers
BIDS-Apps/brainiak-srm	version latest		build failed		docker pulls 79	559.3MB	13 layers
BIDS-Apps/nipypelines	version 0.3.0		build passing		docker pulls 67	478.1MB	20 layers
BIDS-Apps/HCPPIipelines	version v3.17.0-15		build passing		docker pulls 495	2.5GB	62 layers
BIDS-Apps/MAGeTbrain	Image not found		build passing		docker pulls 149	Image not found	
BIDS-Apps/tracula	version v6.0.0-4		build passing	open bug pull requests 0	docker pulls 345	3.4GB	57 layers
BIDS-Apps/baracus	Image not found		build passing		docker pulls 809	Image not found	
BIDS-Apps/antsCorticalThickness	Image not found		build passing	open bug pull requests 0	docker pulls 18	Image not found	
BIDS-Apps/DPARSF	version v4.3.12		build passing		docker pulls 109	1.4GB	28 layers
BIDS-Apps/afni_proc	Image not found		build passing	open bug pull requests 0	docker pulls 48	Image not found	

<http://bids-apps.neuroimaging.io/tutorial/>

# *Open methods: superpowers*

- Instead of having to do a lot of things ourselves:
  - We can adopt techniques others have created
  - And use them in reproducible ways, on reproducibly stored data
- First part of making neuroscience more reproducible



**OpenNEURO**

A free and open platform for sharing MRI, MEG, EEG, iEEG, and ECoG data



Sign in with Google



Sign in with ORCID

[Browse 247 Public Datasets](#)