



MRC Cognition
and Brain
Sciences Unit



UNIVERSITY OF
CAMBRIDGE

Functional Magnetic Resonance Imaging



GitHub https://github.com/dcdace/fMRI_training

Dace [datza] Apšvalka
February 2026

Outline

- Introduction
- Experimental design
- **Data management**
- Pre-processing
- Statistical analysis
- Practical demo

www.nature.com/scientificdata

SCIENTIFIC DATA



OPEN

SUBJECT CATEGORIES

- » Electroencephalography
- EEG
- » Brain imaging
- » Functional magnetic resonance imaging
- » Cognitive neuroscience

Received: 07 April 2014
Accepted: 05 January 2015
Published: 20 January 2015

A multi-subject, multi-modal human neuroimaging dataset

Daniel G. Wakeman^{1,2} & Richard N. Henson²

We describe data acquired with multiple functional and structural neuroimaging modalities on the same nineteen healthy volunteers. The functional data include Electroencephalography (EEG), Magnetoencephalography (MEG) and functional Magnetic Resonance Imaging (fMRI) data, recorded while the volunteers performed multiple runs of hundreds of trials of a simple perceptual task on pictures of familiar, unfamiliar and scrambled faces during two visits to the laboratory. The structural data include T1-weighted MPRAGE, Multi-Echo FLASH and Diffusion-weighted MR sequences. Though only from a small sample of volunteers, these data can be used to develop methods for integrating multiple modalities from multiple runs on multiple participants, with the aim of increasing the spatial and temporal resolution above that of any one modality alone. They can also be used to integrate measures of functional and structural connectivity, and as a benchmark dataset to compare results across the many neuroimaging analysis packages. The data are freely available from <https://openfmri.org/>.

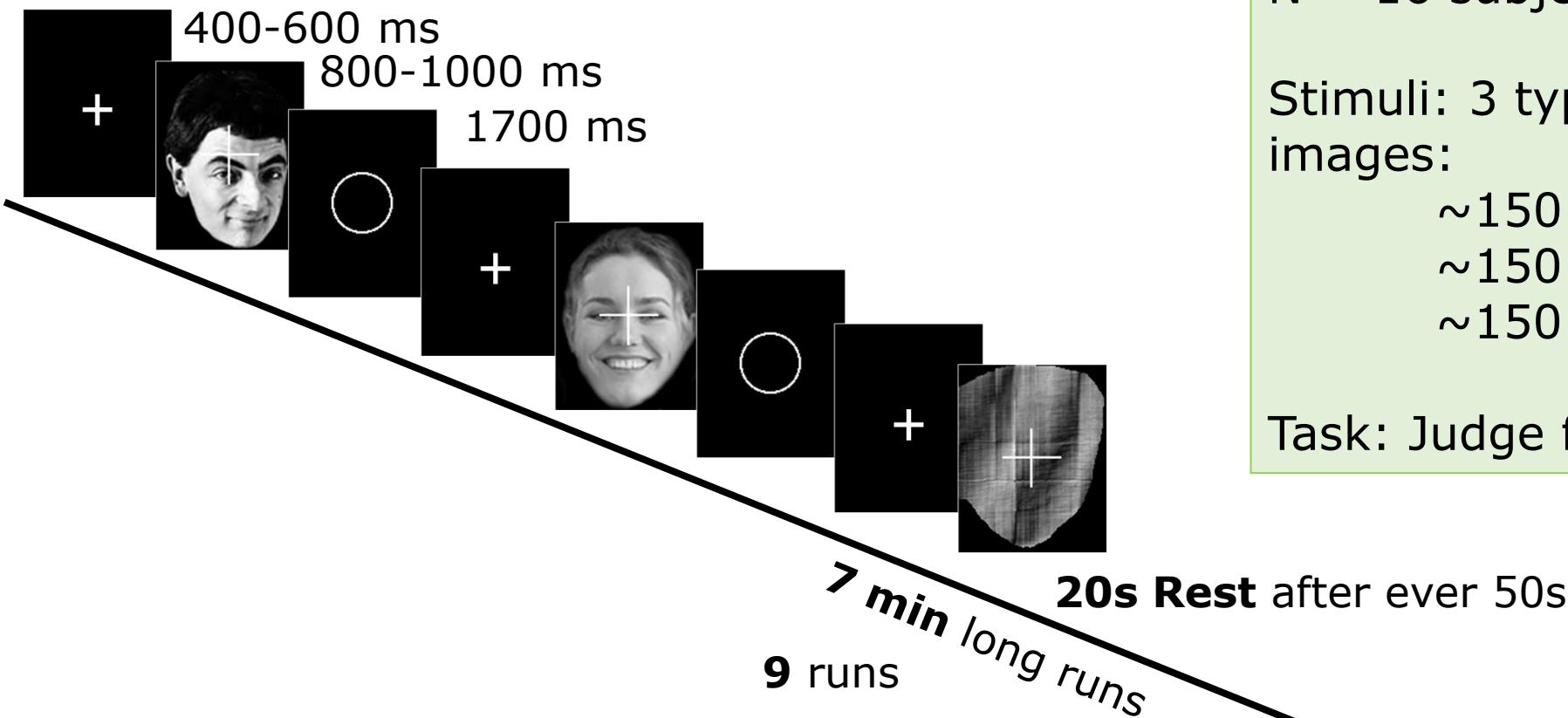
Wakeman & Henson (2015), Scientific Data, <http://www.nature.com/articles/sdata20151>

Example Dataset

<https://openneuro.org/datasets/ds000117/>

- Been used in many methods projects and publications, as well as tutorials (e.g. “multimodal” dataset in SPM12 manual)
- Here we will analyse it from the very root – the raw DICOM images

Example Experiment: Face Recognition



N = 16 subjects

Stimuli: 3 types of greyscale face images:

- ~150 x Familiar
- ~150 x Unfamiliar
- ~150 x Scrambled

Task: Judge face symmetry

Each image was presented twice, with the second presentation occurring either immediately after (**Immediate Repeats**), or after 5-15 intervening stimuli (**Delayed Repeats**), with 50% of each type of repeat.

Environment

Data

Organise & Manage

Pre-process

Analyse

Report

Environment

PROGRAMMING LANGUAGES



BASH & Shell Scripts

A low-level programming language providing a command line user interface for Unix-like operating systems (e.g., Linux, macOS).

Used to automate repetitive tasks and manage system processes and resources.



A high-level, general-purpose programming language.
License-free – good for reproducible & open code.



A high-level programming language designed for engineers and scientists.

Requires a license. Provides loads of useful resources for Neuroimaging analysis.

Environment

PROGRAMMING LANGUAGES



BASH & Shell Scripts

A low-level programming language providing a command line user interface for Unix-like operating systems (e.g., Linux, macOS).

Used to automate repetitive tasks and manage system processes and resources.



A high-level, general-purpose programming language.
License-free – good for reproducible & open code.

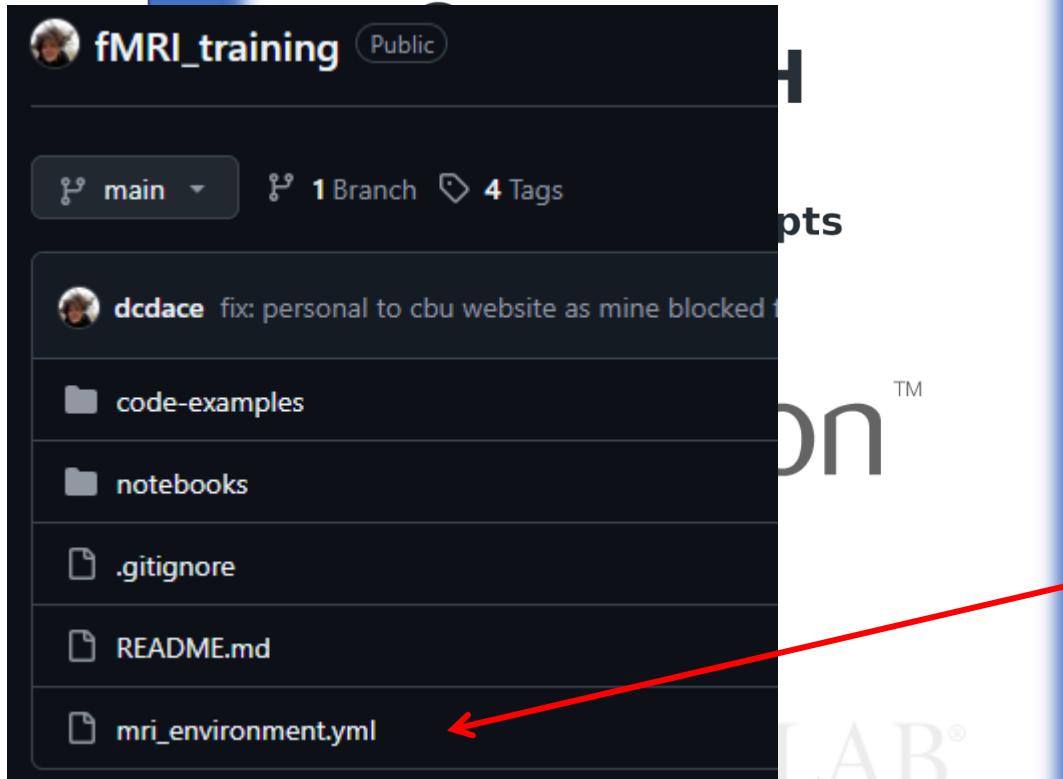


A high-level programming language designed for engineers and scientists.

Requires a license. Provides loads of useful resources for Neuroimaging analysis.

Environment

PROGRAMMING LANGUAGES



PACKAGE MANAGER

CONDA

Conda is an open-source, cross-platform, language-agnostic package manager and environment management system.

With conda, you can use environments that have different versions of Python and packages installed in them.

You can, for example, create your **MRI analysis environment** that includes packages needed for your analysis work.

```
1 name: mri
2 channels:
3   - conda-forge
4   - defaults
5 dependencies:
6   - dcm2niix=1.0.20240202
7   - heudiconv=1.1.6
8   - nipype=1.8.5
9   - pip=24.2
10  - pytest=8.3.2
11  - python=3.11.10 # dipy v1.9.0 dependency
12  - seaborn=0.13.2
13  - traits=6.4.3
14  - wheel=0.44.0
15  - pip:
16    - antspyx==0.5.3
17    - atlasreader==0.3.2
18    - dipy==1.9.0
19    - fury==0.11.0 # dipy v1.9.0 viualisa
20    - ipykernel==6.29.3
21    - ipython==8.22.1
22    - ipywidgets==8.1.5
23    - jupyter==1.1.1
24    - matplotlib==3.8.3
```

LANGUAGES

SH

scripts

CONDA™

Environment

PACKAGE MANAGER



Conda is an open-source, cross-platform, language-agnostic package manager and environment management system.

With conda, you can use environments that have different versions of Python and packages installed in them.

You can, for example, create your **MRI analysis environment** that includes packages needed for your analysis work.

<https://conda.io/projects/conda/en/latest/user-guide/tasks/manage-environments.html>

Creating an environment from a .yml file

`conda env create -f mri_environment.yml`

Environment

PROGRAMMING LANGUAGES



BASH
&
Shell Scripts



PACKAGE MANAGER



CODE EDITOR

Code editors simplify and speed up typing of source code. They also support running and debugging the code.

The best editors are cross-language, cross-platform and support version control.

Environment

PROGRAMMING LANGUAGES



PACKAGE MANAGER



A screenshot of the Visual Studio Code interface. It shows two code editors open: one containing Python script code and another containing a shell script. Below the editors is a terminal window displaying a command-line session. The interface includes a sidebar with file navigation and a bottom status bar.

The Python script code (first_level_script.py) contains:

```
#!/usr/bin/env python3
# -*- coding: utf-8 -*-
#
# =====
# Dace Apšvalka (MRC CBU 2024)
# First level fMRI analysis using Nilearn
# This script requires step08_first_level_analysis.sh, unless
# and sID here manually
#
# =====
# IMPORT RELEVANT PACKAGES
import os
import sys
import pandas as pd
import numpy as np
from bids.layout import BIDSLayout
from nilearn.glm.first_level import FirstLevelModel
import time
import warnings
warnings.filterwarnings("ignore")
```

The shell script code (step08_first_level_analysis.sh) contains:

```
#!/bin/bash
#
# =====
# The script will run the first_level_script.py script for each
# subject in the dataset. You will need a conda environment that contains the required
# packages. Usage:
# Configure the variables below and run the script: ./step08_
#
# =====
# [FILL IN THE VARIABLES BELOW]
# -----
# Your project's root directory
PROJECT_PATH=/data/fMRI/derivatives/func/2024-CBU'
# Location of the first-level analysis script
SCRIPT_PATH=$PROJECT_PATH/first_level_script.py
# Location of the job logs
LOGS_PATH=$PROJECT_PATH/first_level_job_logs
# Define the CONDA_ENV variable
CONDA_ENV=func
```

The terminal window shows a command-line session:

```
(fMRI) [da05@login-j05 code]$ squeue -u da05
JOBID PARTITION NAME USER ST TIME NODES NODENAME(REASON)
3348692 Main fmriprep da05 R 24:33 1 node-i01
3348691 Main fmriprep da05 R 25:28 1 node-j01
3348690 Main fmriprep da05 R 26:06 1 node-j01
3348689 Main fmriprep da05 R 26:51 1 node-h01
3348688 Main fmriprep da05 R 30:31 1 node-j17
3348687 Main fmriprep da05 R 31:08 1 node-j17
3348686 Main fmriprep da05 R 44:43 1 node-k02
3348685 Main fmriprep da05 R 45:02 1 node-k02
```

Bottom status bar: Ln 28, Col 73 Spaces: 4 UTF-8 LF Python 3.10.7 64-bit



MRC Cognition
and Brain
Sciences Unit



UNIVERSITY OF
CAMBRIDGE

Harnessing Visual Studio Code for your research

0:00 / 1:26:29 • Researcher's analysis environment >



In this video

Timeline

Chapters

Transcript



A preview of what we do with VS Code

7:48



Installing VS Code

11:30



An intro to VS Code interface

15:29



Connecting to a remote cluster through
SSH and setting up a typical Python...

24:43



Debugging a Python script

40:08



Using Jupyter Notebooks

47:05



Writing Markdown documents

53:35



JASON files

54:17

Sync to video time
Debugging a Python script with inputs

Harnessing Visual Studio Code for Your Research

<https://www.youtube.com/watch?v=h9cQEoWUrm4>

Environment

PROGRAMMING LANGUAGES



BASH
&
Shell Scripts



python™



MATLAB®

PACKAGE MANAGER

CONDA



Visual Studio Code



CODE EDITOR

VERSION CONTROL

The practice of tracking and managing changes to software code - your analysis code.
It allows to revert selected files back to a previous state, revert the entire project back to a previous state, compare changes over time, do collaborative coding etc.



MRC Cognition
and Brain
Sciences Unit



UNIVERSITY OF
CAMBRIDGE

How to version control your scientific code using Git and GitHub



Máté Aller

MRC Cognition and Brain Sciences Unit

18/11/2024



0:00 / 1:07:41



How to version control your scientific code using Git and GitHub (Mate Aller)

<https://www.youtube.com/watch?v=ohTW4FJdmeQ>

Environment

PROGRAMMING LANGUAGES



BASH
&
Shell Scripts



python™



MATLAB®

PACKAGE MANAGER

CODE EDITOR

VERSION CONTROL

ANALYSIS NOTEBOOK

CONDA



Visual Studio Code



Analysis notebooks

A screenshot of a GitHub repository interface. The repository name is `fMRI_training`. The main menu includes Code, Issues, Pull requests, Agents, Actions, Projects, Wiki, and Settings. The left sidebar shows the file structure: `main`, `code-examples`, `notebooks` (which is selected), `html`, and several IPython notebook files: `nb01_Data-Organisation.ipynb`, `nb02_MRI_data_manipulation.ipynb`, `nb03_Quality-Control-and-Prep...`, `nb04_Subject-Level-Analysis.ipynb`, `nb05_Group-Level-Analysis.ipynb`, `nb06_ROI_analysis.ipynb`, `.gitignore`, `README.md`, and `mri_environment.yml`. The right pane displays the contents of the `notebooks` folder, which includes a folder named `..` and files: `nb01_Data-Organisation.ipynb`, `nb02_MRI_data_manipulation.ipynb`, `nb03_Quality-Control-and-Preprocessing.ipynb`, `nb04_Subject-Level-Analysis.ipynb`, `nb05_Group-Level-Analysis.ipynb`, and `nb06_ROI_analysis.ipynb`. A commit message from `dcdace` is visible: `feat: add html notebooks`.



https://github.com/dcdace/fMRI_training

Example data (~3GB):

<https://cloud.mrc-cbu.cam.ac.uk/index.php/s/rDzG98znXae9Zsy>
(email me for password)



Environment

Data

Organise & Manage

Pre-process

Analyse

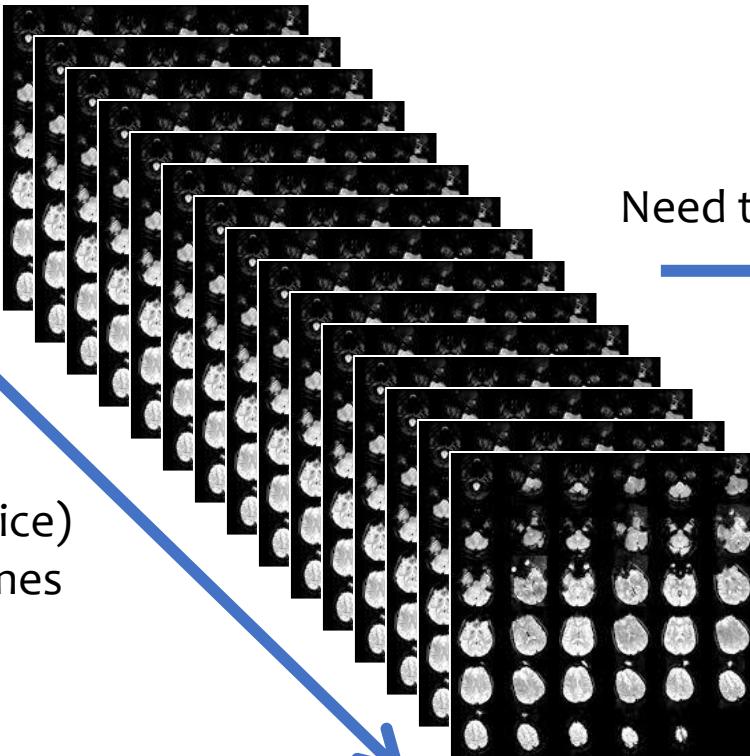
Report

fMRI file formats

Collect the data



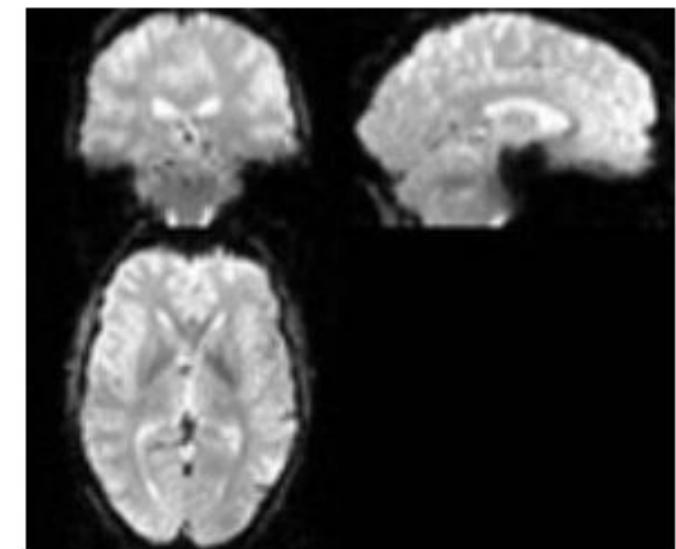
DICOM format



Need to convert to NIfTI

functional scan

A brain volume (slice-by-slice)
scanned every 2s > 100 times



DICOM

Anatomical (T1w) scans

Name
20090826_164150
Series_001_CBU_Localiser
Series_002_CBU_MPRAge
Series_003_CBU_DWEPI_BOLD210
Series_004_CBU_DWEPI_BOLD210
Series_005_CBU_DWEPI_BOLD210
Series_006_CBU_DWEPI_BOLD210
Series_007_CBU_DWEPI_BOLD210
Series_008_CBU_DWEPI_BOLD210
Series_009_CBU_DWEPI_BOLD210
Series_010_CBU_DWEPI_BOLD210
Series_011_CBU_DWEPI_BOLD210
Series_012_CBU_FieldMapping
Series_013_CBU_FieldMapping
DCM File (192)
1.3.12.2.1107.5.2.32.35119.2009082616480248824672575.dcm
1.3.12.2.1107.5.2.32.35119.2009082616480263974172579.dcm
1.3.12.2.1107.5.2.32.35119.2009082616480266173672581.dcm
1.3.12.2.1107.5.2.32.35119.2009082616480282359572587.dcm
1.3.12.2.1107.5.2.32.35119.2009082616480283389772589.dcm
1.3.12.2.1107.5.2.32.35119.2009082616480291722972597.dcm
1.3.12.2.1107.5.2.32.35119.2009082616480294149872599.dcm
1.3.12.2.1107.5.2.32.35119.2009082616480299170772601.dcm
1.3.12.2.1107.5.2.32.35119.2009082616480312524272609.dcm
1.3.12.2.1107.5.2.32.35119.2009082616480314408372615.dcm
1.3.12.2.1107.5.2.32.35119.2009082616480323872972617.dcm
1.3.12.2.1107.5.2.32.35119.2009082616480324494972619.dcm
1.3.12.2.1107.5.2.32.35119.2009082616480334100772625.dcm

Functional (T2*/BOLD) scans

Name
20090826_164150
Series_001_CBU_Localiser
Series_002_CBU_MPRAge
Series_003_CBU_DWEPI_BOLD210
Series_004_CBU_DWEPI_BOLD210
Series_005_CBU_DWEPI_BOLD210
Series_006_CBU_DWEPI_BOLD210
Series_007_CBU_DWEPI_BOLD210
Series_008_CBU_DWEPI_BOLD210
Series_009_CBU_DWEPI_BOLD210
Series_010_CBU_DWEPI_BOLD210
Series_011_CBU_DWEPI_BOLD210
Series_012_CBU_FieldMapping
Series_013_CBU_FieldMapping
DCM File (208)
1.3.12.2.1107.5.2.32.35119.200908261649572708873570.dcm
1.3.12.2.1107.5.2.32.35119.200908261649592579973671.dcm
1.3.12.2.1107.5.2.32.35119.200908261650011521873772.dcm
1.3.12.2.1107.5.2.32.35119.200908261650033002573873.dcm
1.3.12.2.1107.5.2.32.35119.200908261650052824473974.dcm
1.3.12.2.1107.5.2.32.35119.200908261650072679074075.dcm
1.3.12.2.1107.5.2.32.35119.200908261650092563574176.dcm
1.3.12.2.1107.5.2.32.35119.20090826165011450474277.dcm
1.3.12.2.1107.5.2.32.35119.200908261650132877074378.dcm
1.3.12.2.1107.5.2.32.35119.200908261650152763374479.dcm
1.3.12.2.1107.5.2.32.35119.200908261650172643174580.dcm
1.3.12.2.1107.5.2.32.35119.200908261650192479574681.dcm
1.3.12.2.1107.5.2.32.35119.200908261650211329574782.dcm

- **DICOM - Digital Imaging and Communications in Medicine (.dcm)**

- Raw data standard for storing and communicating **medical images**
- Contains a **header** (meta data) and the actual **image** itself
- Images are stored in **2D layers**
- A separate file for each **slice** (T1w) or **volume** (BOLD)

DICOM: Anatomical

DICOM browser

Search...

All patients(Patients: 1, Images: 192)

CBU090938

CBU090938

CBU Neuroimaging(MR: 1 series)

CBU_MPRAGE

1.3.12.2.1107.5.2.32.35119.2009082609555636896484776.c
1.3.12.2.1107.5.2.32.35119.2009082609555632008284776.c
1.3.12.2.1107.5.2.32.35119.2009082609555644050484784.c
1.3.12.2.1107.5.2.32.35119.2009082609555649109184786.c
1.3.12.2.1107.5.2.32.35119.2009082609555660191584794.c
1.3.12.2.1107.5.2.32.35119.20090826095556679284792.c
1.3.12.2.1107.5.2.32.35119.2009082609555669059484802.c
1.3.12.2.1107.5.2.32.35119.2009082609555674269184804.c
1.3.12.2.1107.5.2.32.35119.2009082609555685752984810.c
1.3.12.2.1107.5.2.32.35119.2009082609555686170384812.c
1.3.12.2.1107.5.2.32.35119.2009082609555694120284816.c
1.3.12.2.1107.5.2.32.35119.200908260955569677984818.c
1.3.12.2.1107.5.2.32.35119.200908260955572374084824.d
1.3.12.2.1107.5.2.32.35119.200908260955575468484826.d
1.3.12.2.1107.5.2.32.35119.2009082609555716376884832.c
1.3.12.2.1107.5.2.32.35119.2009082609555723796484838.c
1.3.12.2.1107.5.2.32.35119.200908260955572548434840.c
1.3.12.2.1107.5.2.32.35119.200908260955573303654842.c
1.3.12.2.1107.5.2.32.35119.2009082609555743189084846.c
1.3.12.2.1107.5.2.32.35119.200908260955575436474854.c
1.3.12.2.1107.5.2.32.35119.2009082609555759470484860.c
1.3.12.2.1107.5.2.32.35119.200908260955576839974866.c
1.3.12.2.1107.5.2.32.35119.2009082609555770769984870.c
1.3.12.2.1107.5.2.32.35119.2009082609555790816784884.c
1.3.12.2.1107.5.2.32.35119.2009082609555789094684880.c
1.3.12.2.1107.5.2.32.35119.20090826095558444984892.d
1.3.12.2.1107.5.2.32.35119.200908260955581570984888.d
1.3.12.2.1107.5.2.32.35119.2009082609555821034484902.c
1.3.12.2.1107.5.2.32.35119.2009082609555816681284898.c
1.3.12.2.1107.5.2.32.35119.200908260955582696354908.c
1.3.12.2.1107.5.2.32.35119.2009082609555832137684912.c
1.3.12.2.1107.5.2.32.35119.2009082609555844678084918.c
1.3.12.2.1107.5.2.32.35119.2009082609555842377884916.c
1.3.12.2.1107.5.2.32.35119.2009082609555851923884924.c
1.3.12.2.1107.5.2.32.35119.2009082609555857064184926.c
1.3.12.2.1107.5.2.32.35119.2009082609555865957784932.c
1.3.12.2.1107.5.2.32.35119.2009082609555869132184934.c
1.3.12.2.1107.5.2.32.35119.2009082609555875856184940.c
1.3.12.2.1107.5.2.32.35119.2009082609555881020184944.c
1.3.12.2.1107.5.2.32.35119.2009082609555888987984948.c
1.3.12.2.1107.5.2.32.35119.20090826095558915829284964.c
1.3.12.2.1107.5.2.32.35119.2009082609555897357184952.c
1.3.12.2.1107.5.2.32.35119.200908260955595014584956.d
1.3.12.2.1107.5.2.32.35119.2009082609555912645334962.c
1.3.12.2.1107.5.2.32.35119.2009082609555920054584968.c
1.3.12.2.1107.5.2.32.35119.2009082609555942028984980.c
1.3.12.2.1107.5.2.32.35119.2009082609555928045984972.c
1.3.12.2.1107.5.2.32.35119.2009082609555936426884978.c
1.3.12.2.1107.5.2.32.35119.2009082609555949691084988.c
1.3.12.2.1107.5.2.32.35119.200908260955596797184998.c
1.3.12.2.1107.5.2.32.35119.200908260955597421284990.c
1.3.12.2.1107.5.2.32.35119.200908260955598515485000.c
1.3.12.2.1107.5.2.32.35119.2009082609555979689485004.c
1.3.12.2.1107.5.2.32.35119.2009082609555981592785008.c
1.3.12.2.1107.5.2.32.35119.20090826095559903585012.c
1.3.12.2.1107.5.2.32.35119.200908260955599414545016.c
1.3.12.2.1107.5.2.32.35119.20090826095600206585020.dcr
1.3.12.2.1107.5.2.32.35119.200908260956002234685022.d
1.3.12.2.1107.5.2.32.35119.200908260956007930285026.d
1.3.12.2.1107.5.2.32.35119.200908260956001331885030.c
1.3.12.2.1107.5.2.32.35119.2009082609560015539585032.c
1.3.12.2.1107.5.2.32.35119.2009082609560023177985036.c
1.3.12.2.1107.5.2.32.35119.2009082609560024350785038.c
1.3.12.2.1107.5.2.32.35119.2009082609560030415585040.c
1.3.12.2.1107.5.2.32.35119.2009082609560035837385044.v



26-August-2009 9:51:26

MRC-CBU

TrioTim

DICOM Tags

Search...

Patient Name	CBU090938
PatientID	MRO9029
Patient Birth Date	19830904
Patient Sex	M
Patient Age	25Y
Patient Weight	70.000000
Patient Address	
Study Date	26-August-2009
Study Time	9:43:37
Study ID	1
Study Modality	MR
Study Description	CBU Neuroimaging
Series Date	26-August-2009
Series Time	9:55:56
Series Description	CBU_MPRAGE

DICOM: Functional

DICOM browser Search... X

All patients(Patients: 1, Images: 208) Cbu090938

Cbu Neuroimaging(MR 1 series) CBU090938

Cbu_DWEPI_BOLD210 CBU_DWEPI_BOLD210

13.1.2.2.1107.5.2.32.35119.200908260959346867948
13.1.2.2.1107.5.2.32.35119.200908260959366845383
13.1.2.2.1107.5.2.32.35119.200908260959386733138
13.1.2.2.1107.5.2.32.35119.200908260959406836561
13.1.2.2.1107.5.2.32.35119.2009082609594268656461
13.1.2.2.1107.5.2.32.35119.20090826095944685652331
13.1.2.2.1107.5.2.32.35119.200908260959466836951
13.1.2.2.1107.5.2.32.35119.200908260959486727751
13.1.2.2.1107.5.2.32.35119.200908260959506875051
13.1.2.2.1107.5.2.32.35119.200908260959526863591
13.1.2.2.1107.5.2.32.35119.200908260959546854651
13.1.2.2.1107.5.2.32.35119.200908260959566832861
13.1.2.2.1107.5.2.32.35119.200908260959586724711
13.1.2.2.1107.5.2.32.35119.200908261000006532021
13.1.2.2.1107.5.2.32.35119.20090826100002685521
13.1.2.2.1107.5.2.32.35119.200908261000046854481
13.1.2.2.1107.5.2.32.35119.200908261000066829361
13.1.2.2.1107.5.2.32.35119.200908261000106698061
13.1.2.2.1107.5.2.32.35119.200908261000126845731
13.1.2.2.1107.5.2.32.35119.20090826100014684531
13.1.2.2.1107.5.2.32.35119.200908261000166822831
13.1.2.2.1107.5.2.32.35119.200908261000186800301
13.1.2.2.1107.5.2.32.35119.200908261000206692331
13.1.2.2.1107.5.2.32.35119.200908261000226844571
13.1.2.2.1107.5.2.32.35119.200908261000246830401
13.1.2.2.1107.5.2.32.35119.200908261000266809951
13.1.2.2.1107.5.2.32.35119.200908261000286803781
13.1.2.2.1107.5.2.32.35119.200908261000306684831
13.1.2.2.1107.5.2.32.35119.2009082610003264699401
13.1.2.2.1107.5.2.32.35119.200908261000346828561
13.1.2.2.1107.5.2.32.35119.200908261000366867901
13.1.2.2.1107.5.2.32.35119.20090826100038671901
13.1.2.2.1107.5.2.32.35119.200908261000406778351
13.1.2.2.1107.5.2.32.35119.200908261000426668311
13.1.2.2.1107.5.2.32.35119.200908261000446811981
13.1.2.2.1107.5.2.32.35119.200908261000466800411
13.1.2.2.1107.5.2.32.35119.200908261000496784121
13.1.2.2.1107.5.2.32.35119.200908261000506776241
13.1.2.2.1107.5.2.32.35119.200908261000526662551
13.1.2.2.1107.5.2.32.35119.200908261000546810231
13.1.2.2.1107.5.2.32.35119.200908261000566791881
13.1.2.2.1107.5.2.32.35119.200908261000586778611
13.1.2.2.1107.5.2.32.35119.200908261001006761571
13.1.2.2.1107.5.2.32.35119.200908261001026646311
13.1.2.2.1107.5.2.32.35119.200908261001046454191
13.1.2.2.1107.5.2.32.35119.200908261001066789611
13.1.2.2.1107.5.2.32.35119.200908261001086768091
13.1.2.2.1107.5.2.32.35119.200908261001086770191
13.1.2.2.1107.5.2.32.35119.200908261001126740631
13.1.2.2.1107.5.2.32.35119.200908261001146628781
13.1.2.2.1107.5.2.32.35119.200908261001166765151
13.1.2.2.1107.5.2.32.35119.200908261001206746451
13.1.2.2.1107.5.2.32.35119.200908261001246633771
13.1.2.2.1107.5.2.32.35119.200908261001266774791
13.1.2.2.1107.5.2.32.35119.200908261001286757091
13.1.2.2.1107.5.2.32.35119.200908261001306741711
13.1.2.2.1107.5.2.32.35119.200908261001326731881
13.1.2.2.1107.5.2.32.35119.200908261001346618321
13.1.2.2.1107.5.2.32.35119.200908261001366425231
13.1.2.2.1107.5.2.32.35119.200908261001386751891
13.1.2.2.1107.5.2.32.35119.200908261001406737591
13.1.2.2.1107.5.2.32.35119.200908261001426712471
AS
R L
ST: 3 SL: -50.170833
RT: 2000 ET: 30
FS: 3
MR
LittleEndianImplicit
Images: 1/208
Series: 3
Zoom: 327%
WL: 743 WW: 1555
Patient information All Tags Favorite Ta

DICOM Tags Search... X

Patient Name CBU090938
Patient ID MR09029
Patient Birth Date 19830904
Patient Sex M
Patient Age 25Y
Patient Weight 70.00000
Patient Address
Study Date 26-August-2009
Study Time 9:43:37
Study ID 1
Study Modality MR
Study Description CBU Neuroimaging

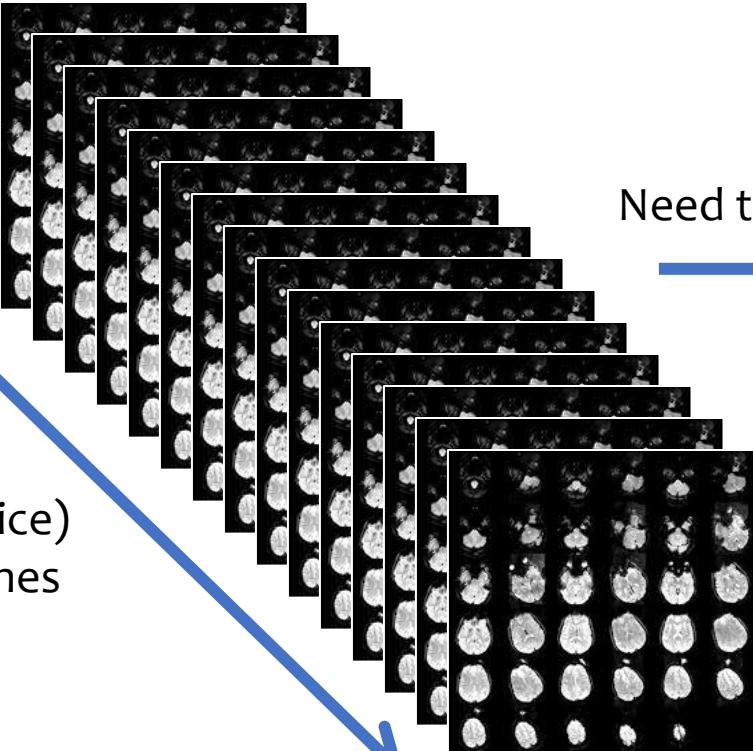
Series Date 26-August-2009
Series Time 9:59:31
Series Description CBU_DWEPI_BOLD210

Collect the data



Anatomical (T1w) image & Functional (T2*/BOLD) image

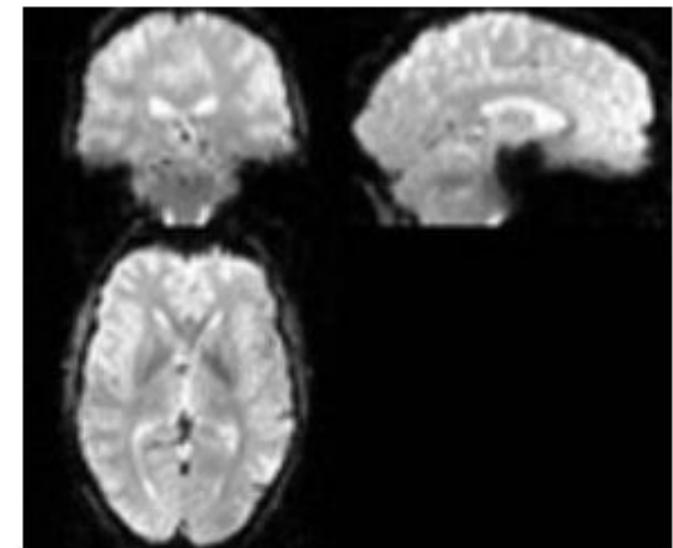
DICOM format



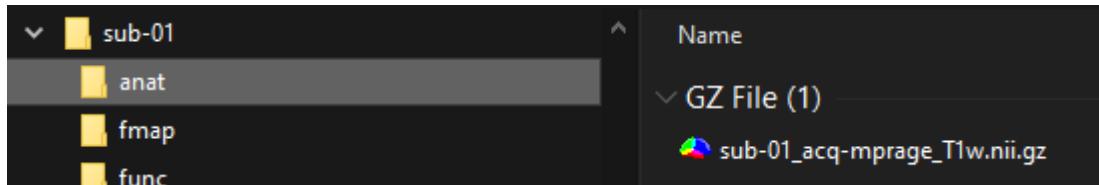
Need to convert to NIfTI

functional scan

A brain volume (slice-by-slice)
scanned every 2s > 100 times

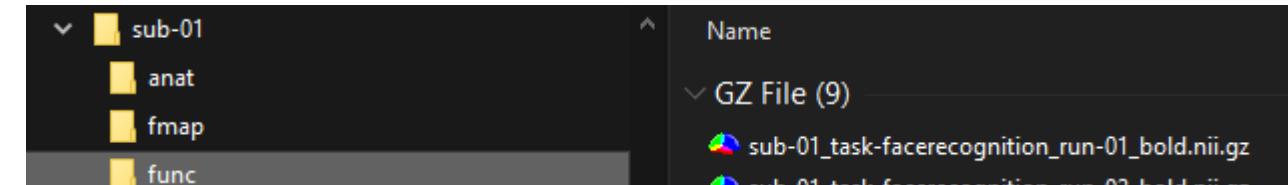


Anatomical (T1w) scans



3D file

Functional (T2*/BOLD) scans



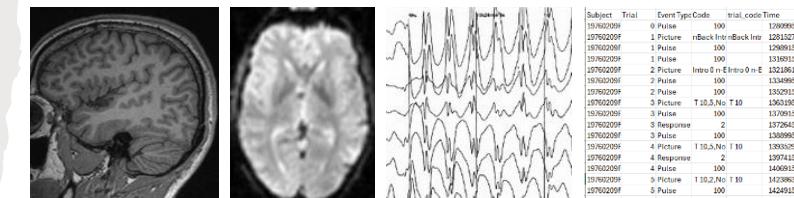
4D file

- **NIfTI – Neuroimaging Informatics Technology Initiative (.nii, .nii.gz)**
 - Standardised representation of **brain images**, cross-platform, cross-software
 - Contains **header** and **image**
 - 3D or 4D files (all slices/volumes in a single file)
- **DICOM vs NIfTI**
 - All medical vs brain images
 - 2D layer files vs 3D/4D files
 - More vs less metadata
 - When converting DICOM to NIfTI need to be aware of that and save all metadata that might be needed
- **DICOM to NIfTI**
 - Several tools are available
 - For example, **dcm2niix** Python package

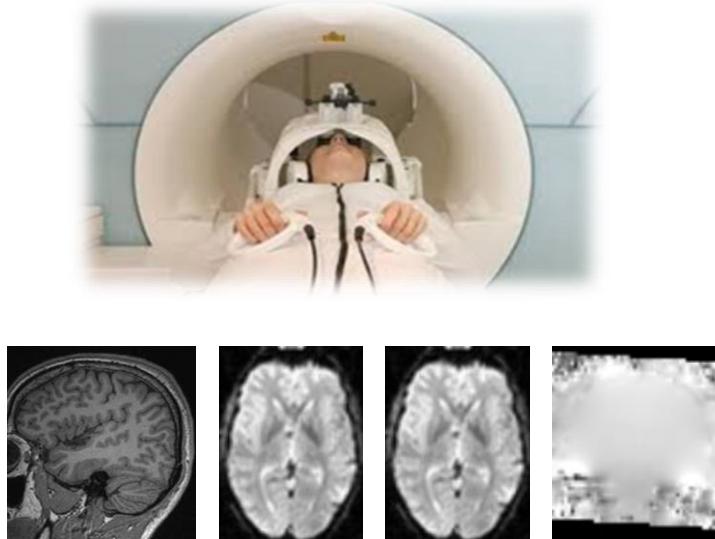
File organisation

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
19760209F	0	Pulse	100	1300000	T
19760209F	1	Picture	nBack Intra-Back Inv	1281327	
19760209F	1	Pulse	100	1280915	
19760209F	1	Pulse	100	1310915	
19760209F	2	Picture	Intra-0 n-E	1322015	
19760209F	2	Pulse	100	1330905	
19760209F	2	Pulse	100	1332915	
19760209F	3	Picture	T0,5,No T10	1363195	
19760209F	3	Pulse	100	1370915	
19760209F	3	Response	2	1377915	
19760209F	3	Pulse	100	1388995	
19760209F	4	Picture	T10,3,No 110	1393329	
19760209F	4	Response	2	1397415	
19760209F	4	Pulse	100	1407415	
19760209F	5	Picture	T110,3,Mo 110	1423983	
19760209F	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

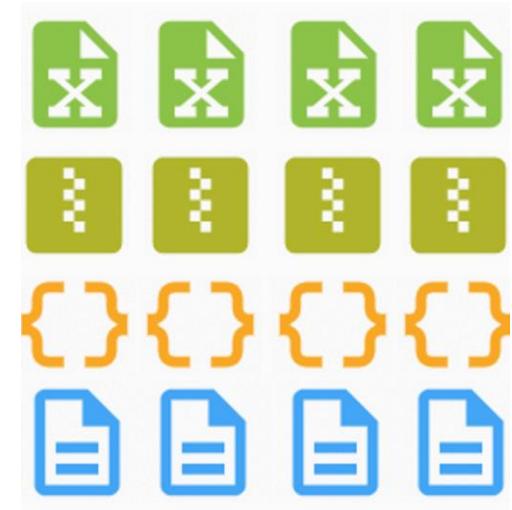


Many possible ways to name and organise the data

me and (more modalities, sessions, tasks)

Inconsistent Data Organisation

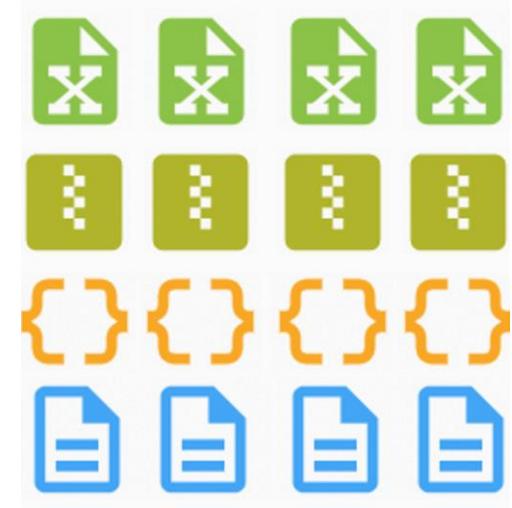
- Difficult for others (and you!) to understand data and track changes
- Scripts must be adapted, can't be easily reused
- Huge effort to automate workflows, no way to automatically validate datasets
- Increased risk of errors (wrong files, outdated versions)
- Time wasted searching and reorganising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



Inconsistent Data Organisation

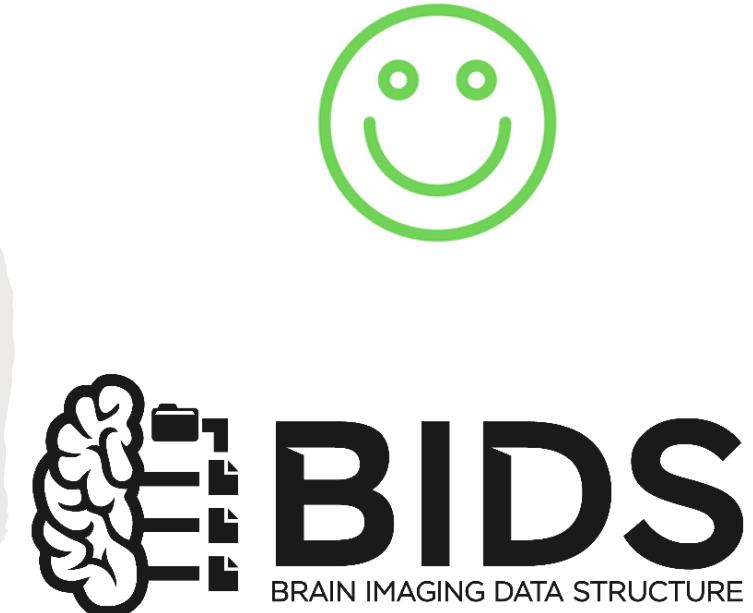


- Difficult for others (and you!) to understand data and track changes
- Scripts must be adapted, can't be easily reused
- **Wouldn't it be much easier if everybody organised the files in the same way?**
- Increased risk of errors (wrong files, outdated versions)
- Time wasted searching and reorganising
- Harder to reproduce results and collaborate



Brain Imaging Data Structure (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 Imaging Data Structure (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



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

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

Krzysztof J. Gorgolewski^{1*}, Tibor Auer², Vince D. Calhoun^{3,4}, R. Cameron Craddock^{5,6}, Samir Das⁷, Eugene P. Duff⁸, Guillaume Flandin⁹, Satrajit S. Ghosh^{10,11}, Tristan Glatard^{7,12}, Yaroslav O. Halczenko¹³, Daniel A. Handwerker¹⁴, Michael Hanke^{15,16}, David Keator¹⁷, Xiangrui Li¹⁸, Zachary Michael¹⁹, Camille Maumet²⁰, B. Nolan Nichols^{21,22}, Thomas E. Nichols^{20,23}, John Pellman⁶, Jean-Baptiste Poline²⁴, Ariel Rokem²⁵, Gunnar Schaefer^{2,26}, Vanessa Soshat²⁷, William Triplett¹, Jessica A. Turner^{3,28}, Gaël Varoquaux²⁹ & Russell A. Poldrack¹

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-Almagro², Tibor Auer³, Pierre Bellec^{4,5}, Mihai Capota⁶, M. Mallar Chakravarty^{7,8}, Nathan W. Churchill⁹, Alexander Li Cohen¹⁰, R. Cameron Craddock^{11,12}, Gabriel A. Devenyi^{7,8}, Anders Eklund^{13,14,15}, Oscar Esteban¹, Guillaume Flandin¹⁶, Satrajit S. Ghosh^{17,18}, J. Swaroop Guntupalli¹⁹, Mark Jenkins², Anisha Keshavan²⁰, Gregory Kiar^{21,22}, Franziskus Liem²³, Pradeep Reddy Raamana^{24,25}, David Raffelt²⁶, Christopher J. Steele^{7,8}, Pierre-Olivier Quirion¹⁵, Robert E. Smith²⁶, Stephen C. Strother^{24,25}, Gaël Varoquaux²⁷, Yida Wang⁶, Tal Yarkoni²⁸, Russell A. Poldrack¹

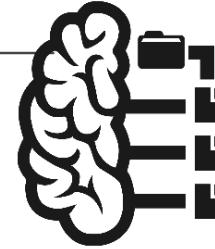


Raw DICOM format

```
mridata/  
└─ CBU090962_MR09029  
  └─ 20090902_100102  
    ├─ Series_001_CBU_Localiser  
    ├─ Series_002_CBU_MPRAGE  
    ├─ Series_003_CBU_DWEPI_BOLD210  
    ├─ Series_004_CBU_DWEPI_BOLD210  
    ├─ Series_005_CBU_DWEPI_BOLD210  
    ├─ Series_006_CBU_DWEPI_BOLD210  
    ├─ Series_007_CBU_DWEPI_BOLD210  
    ├─ Series_008_CBU_DWEPI_BOLD210  
    ├─ Series_009_CBU_DWEPI_BOLD210  
    ├─ Series_010_CBU_DWEPI_BOLD210  
    ├─ Series_011_CBU_DWEPI_BOLD210  
    └─ Series_012_CBU_FieldMapping  
      └─ Series_013_CBU_FieldMapping
```

BIDS NIfTI format

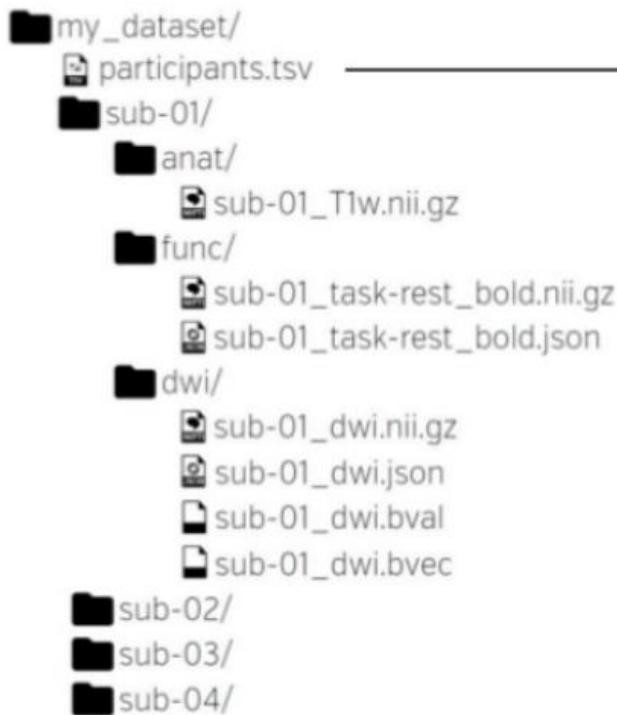
```
sub-15  
└─ ses-mri  
  └─ anat  
    └─ sub-15_ses-mri_T1w.json  
    └─ sub-15_ses-mri_T1w.nii.gz  
  └─ fmap  
    └─ sub-15_ses-mri_acq-func_magnitude1.json  
    └─ sub-15_ses-mri_acq-func_magnitude1.nii.gz  
    └─ sub-15_ses-mri_acq-func_magnitude2.json  
    └─ sub-15_ses-mri_acq-func_magnitude2.nii.gz  
    └─ 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
```



BIDS
BRAIN IMAGING DATA STRUCTURE

BIDS

- Contains participant information

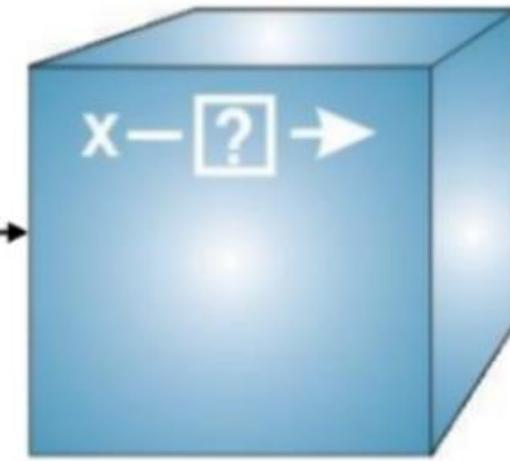
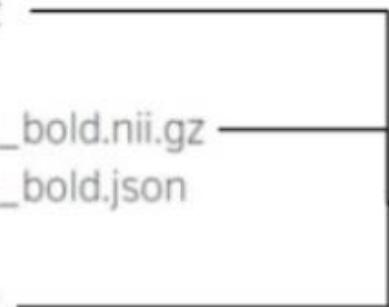


participant_id	age	sex
sub-001	34	M
sub-002	12	F
sub-003	33	F

BIDS

- Contains data files

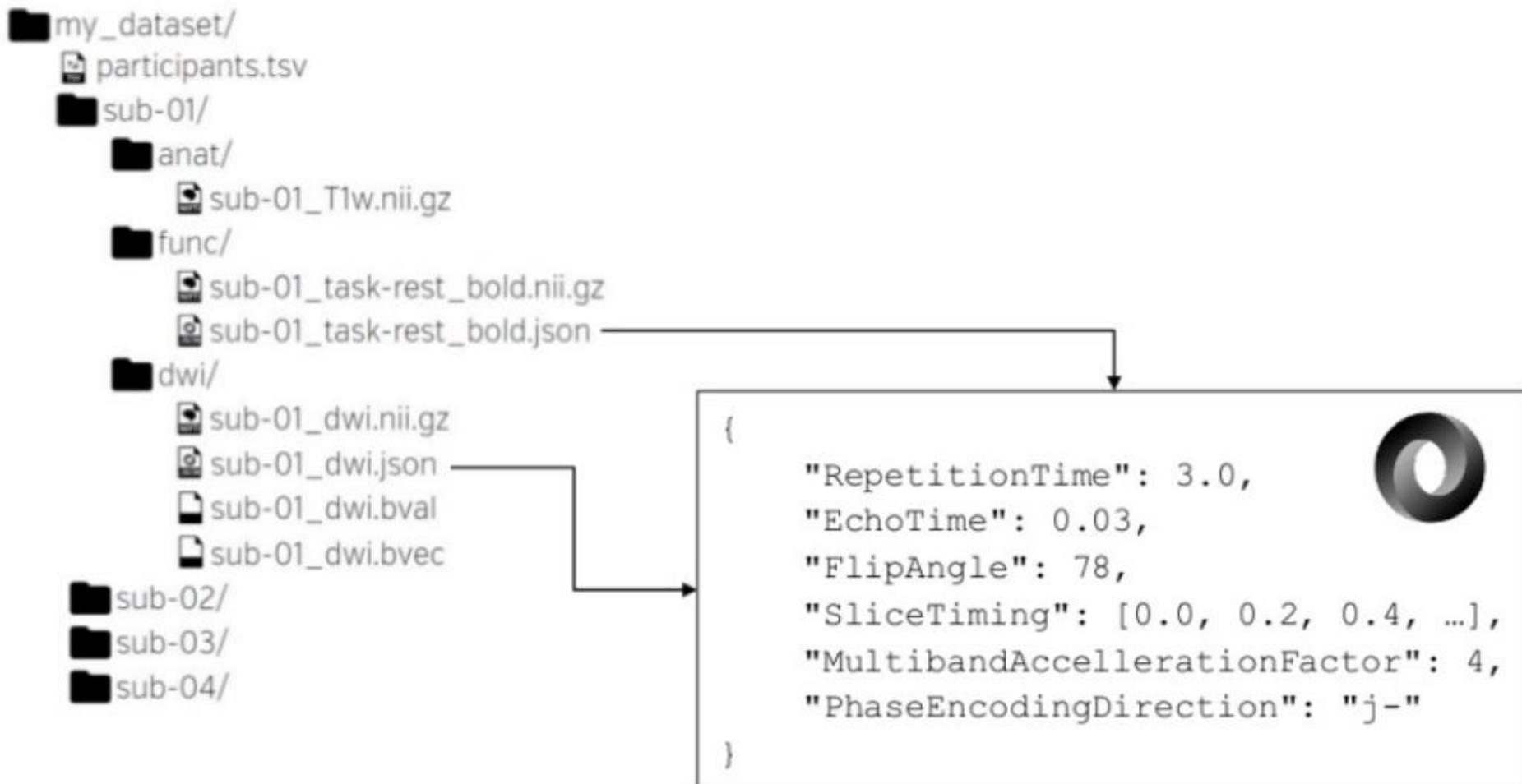
```
my_dataset/
  participants.tsv
  sub-01/
    anat/
      sub-01_T1w.nii.gz
    func/
      sub-01_task-rest_bold.nii.gz
      sub-01_task-rest_bold.json
    dwi/
      sub-01_dwi.nii.gz
      sub-01_dwi.json
      sub-01_dwi.bval
      sub-01_dwi.bvec
  sub-02/
  sub-03/
  sub-04/
```



NIfTI

BIDS

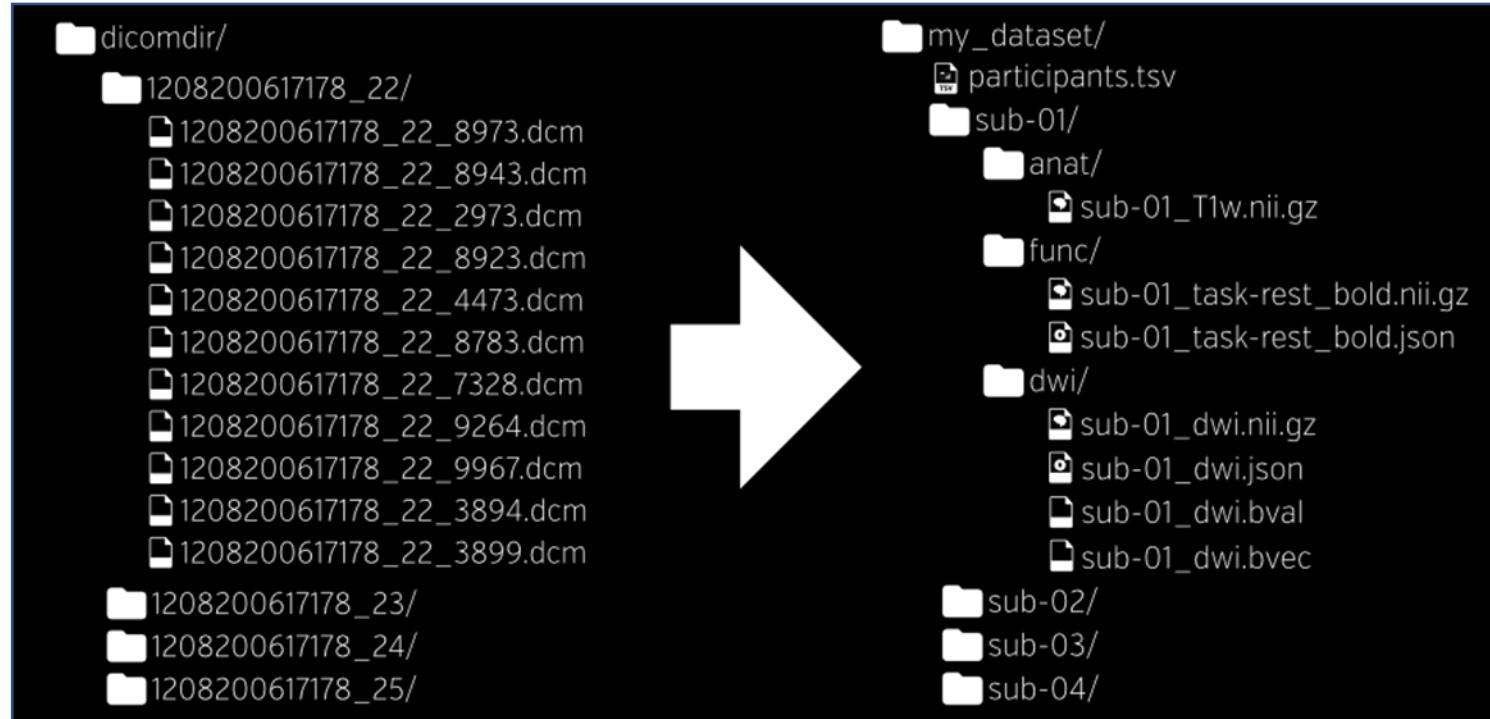
- Contains study specific JSON (metadata) files



DICOM → BIDS

Many BIDS converters available

HeuDiConv (Heuristic Dicom Conversion)
provides sophisticated and flexible creation of BIDS datasets.



The image shows a video player interface with three main sections. On the left, under 'Raw MRI data', there is a list of numerous DICOM files (e.g., Series_001_CBU_Localiser.dcm, Series_002_CBU_MPRAge.dcm) with no clear organization. In the center, under 'Organised MRI data', the same files are shown in a hierarchical BIDS directory structure: sub-01/anat, sub-01/fmap, and sub-01/func. On the right, the 'In this video' sidebar lists several BIDS-related tutorials with their titles and durations.

Raw MRI data

Organised MRI data

In this video

Timeline Chapters Transcript

Why BIDS? The Importance of Data Organization 0:00

Finding BIDS Tutorials 3:05

Installation 4:46

Getting Sample Scripts 5:32

Where Are Your Raw MRI Data & A Preview of Data in BIDS Format 9:58

A HeuDiConv Tool 16:58

Knowing/Discovering Your Scans 17:44

Creating a Heuristic File 25:05

Raw MRI to BIDS

0:00 / 1:01:28 • Why BIDS? The Importance of Data Organization >

Organising your neuroimaging data. Part 1: MRI

<https://www.youtube.com/watch?v=yQBGsZMttCc>

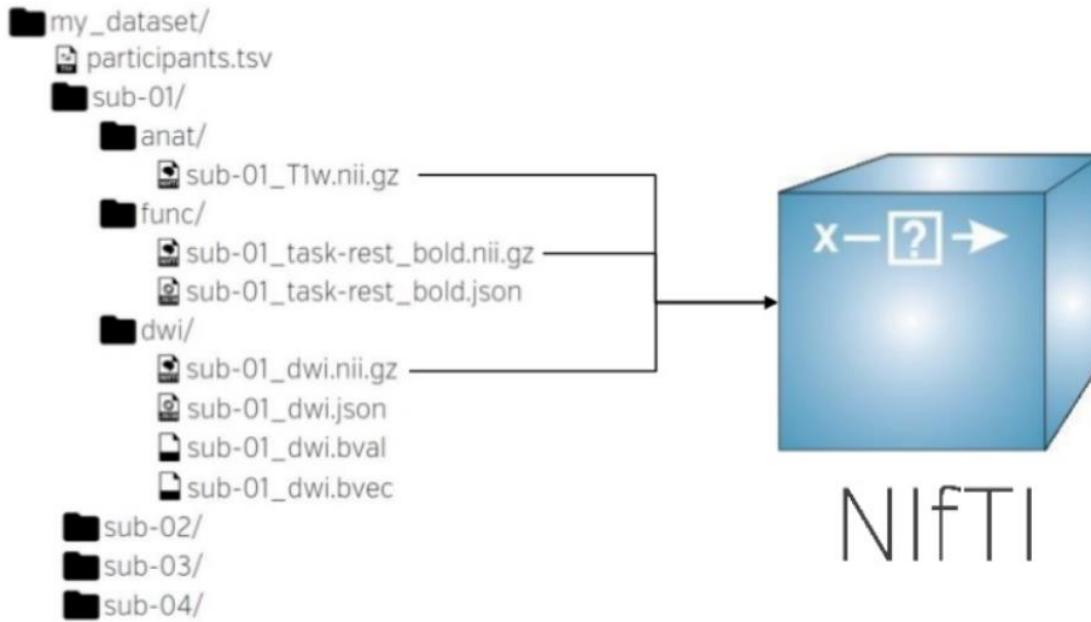
https://github.com/MRC-CBU/BIDS_conversion/tree/main/MRI

Environment

Data
Organise & Manage



Notebook: [nb01_Data-Organisation.ipynb](#)

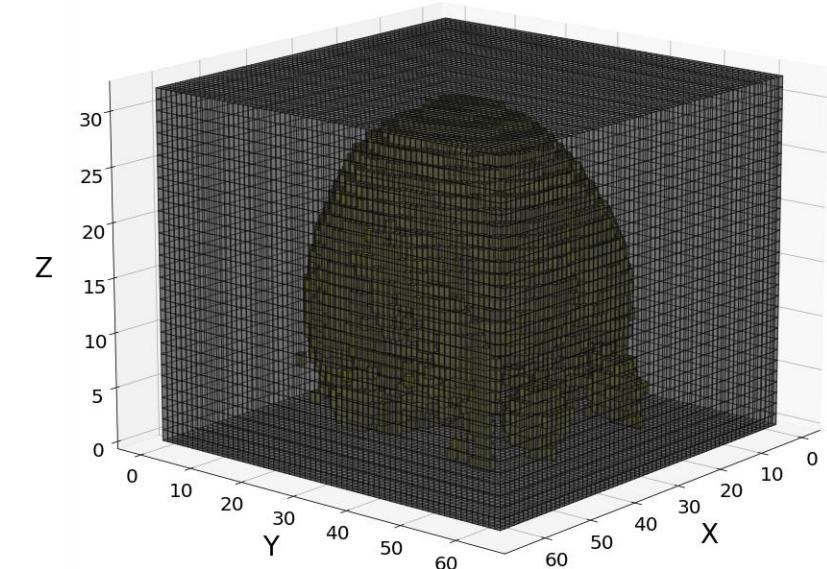
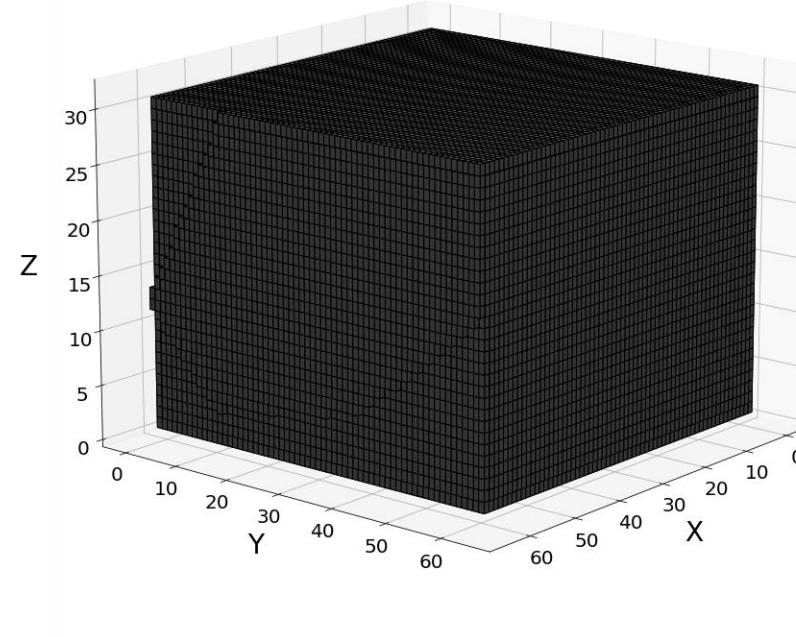


Imaging data content

MRI data structure

A 3D or 4D arrays of numbers

```
([[[ 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.]],  
 ...)
```



MRI data structure

A 3D or 4D arrays of numbers – **intensity values**

```
([[[ 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.]],  
 ...)
```

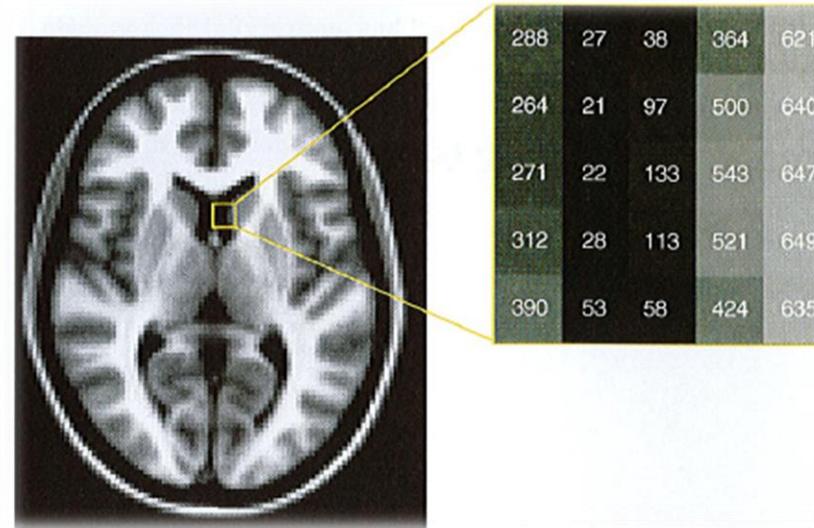
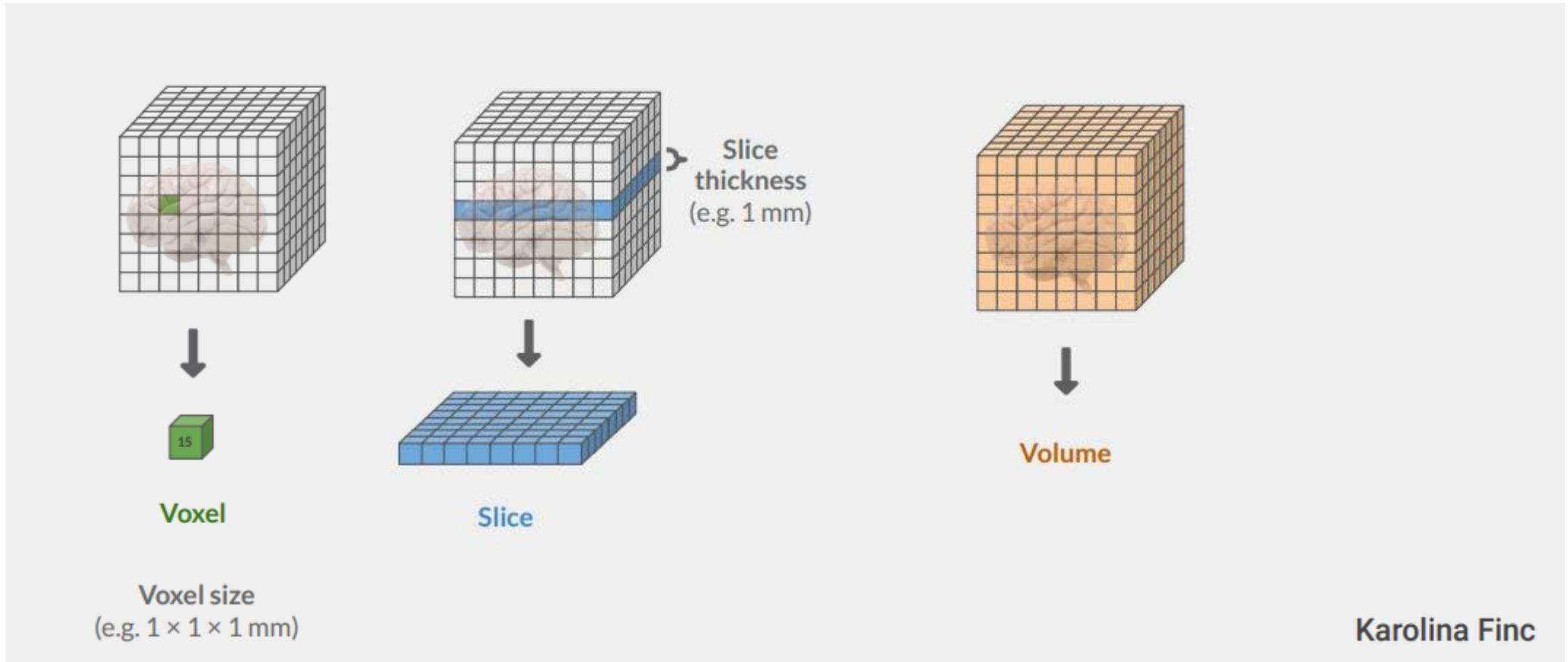


Image from Poldrack et al., 2011

MRI data structure



MRI data



MRI data

What determines the resolution?

Why can't we acquire the functional images with higher resolution?

high resolution

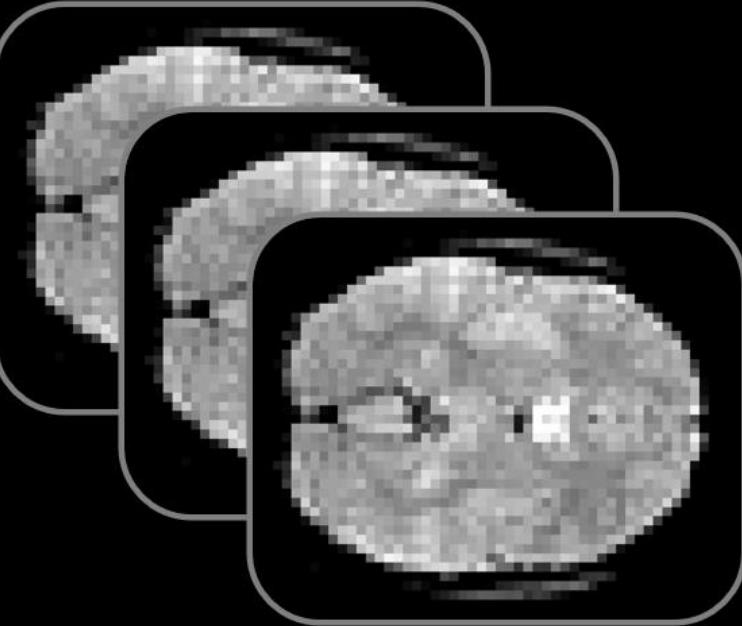
MRI



One 3D volume

fMRI

low resolution

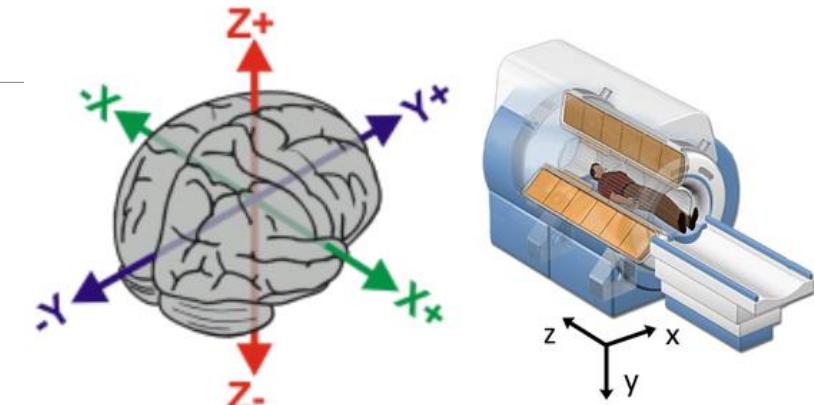
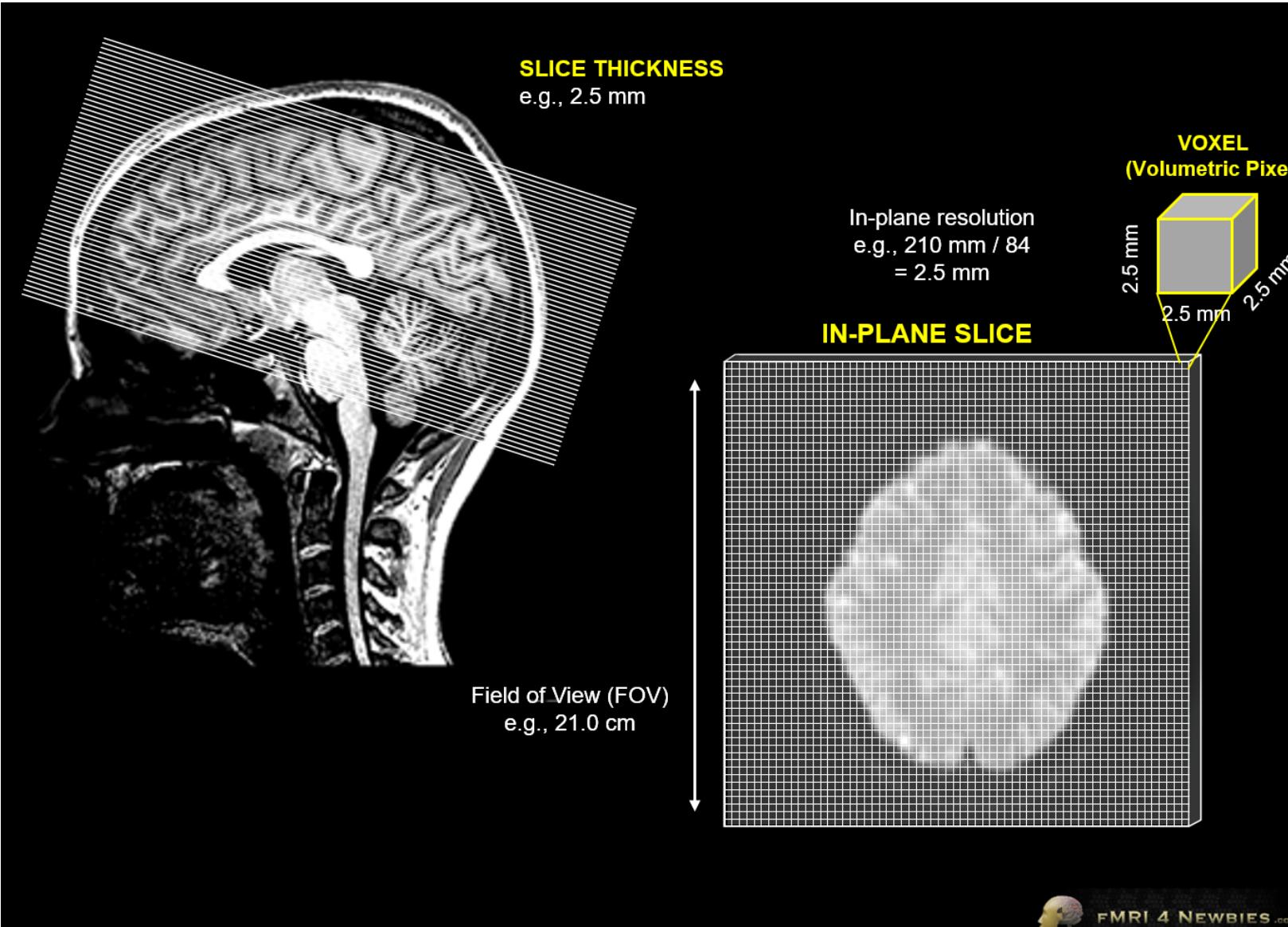


series of 3D volumes (i.e., 4D data)
(e.g., every 2 sec for 5 mins)



fMRI data

- Acquired in slices (usually axial; z-axis)

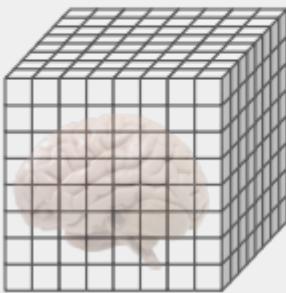


- Temporal resolution (TR), usually 1.5-3s
- Modern sequences allow acquiring multiple slices at the same time (multi-band)
- Typically, 30-50 slices acquired
- More slices = longer TR

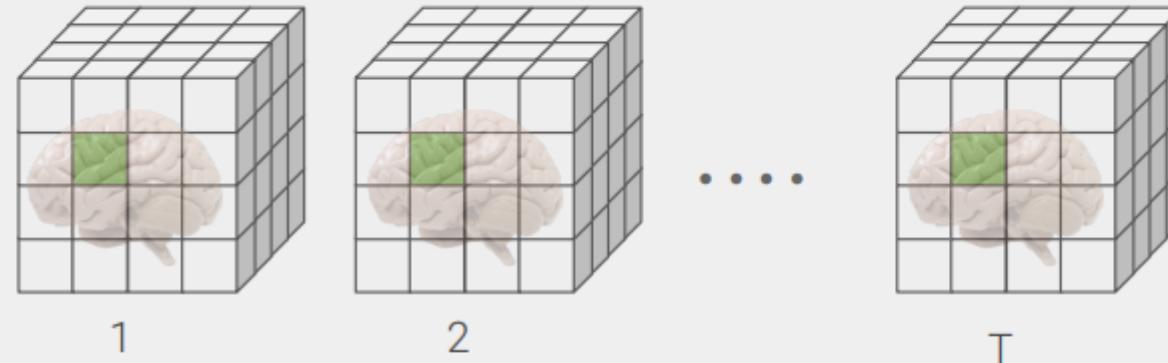


MRI data structure

Structural data

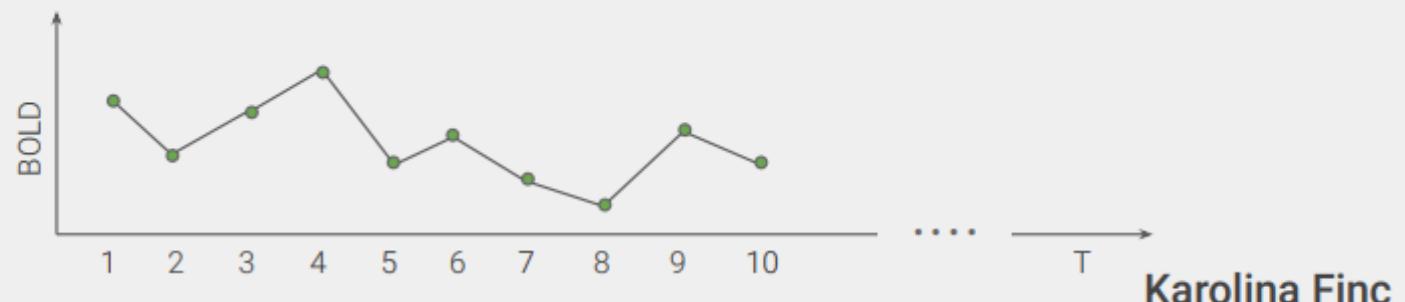


Functional data



Time series - is a series of data points listed in time order.

Every voxel has its own time-series.



Environment

Data
Organise & Manage



Notebook: [nbo2_MRI_data_manipulation.ipynb](#)

✓ Environment

✓ Data
Organise & Manage

Pre-process

Outline

- Introduction
- Experimental design
- Data management
- **Pre-processing**
- Statistical analysis
- Practical demo