MRC Cognition and Brain Sciences Unit

UNIVERSITY OF CAMBRIDGE

fMRI analysis

Dace Apšvalka

[Datza]

http://www.dcdace.net

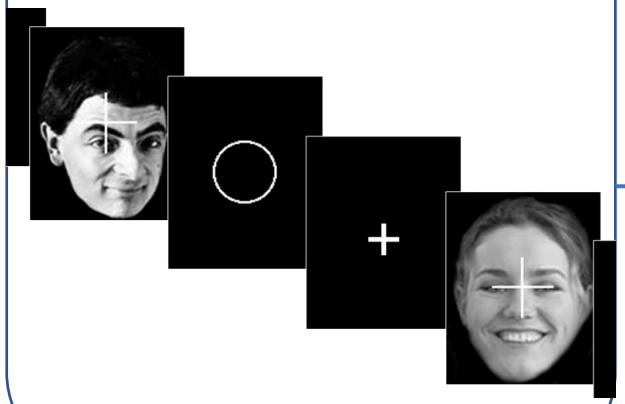
Winter, 2023

Materials

https://github.com/dcdace/fMRI_training

Famous vs **Unfamiliar** faces are processed differently in the brain

Design an experiment

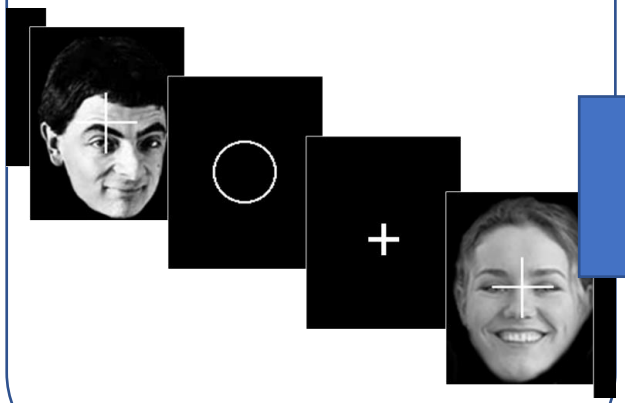


Collect the MRI data

What do we do now?

Famous vs **Unfamiliar** faces are processed differently in the brain

Design an experiment



Data

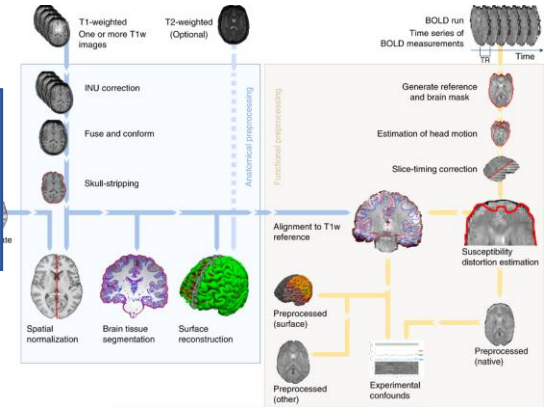
Stimuli
Timing

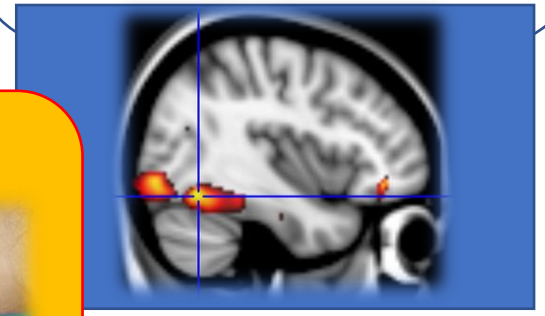
Collect the MRI data

Data

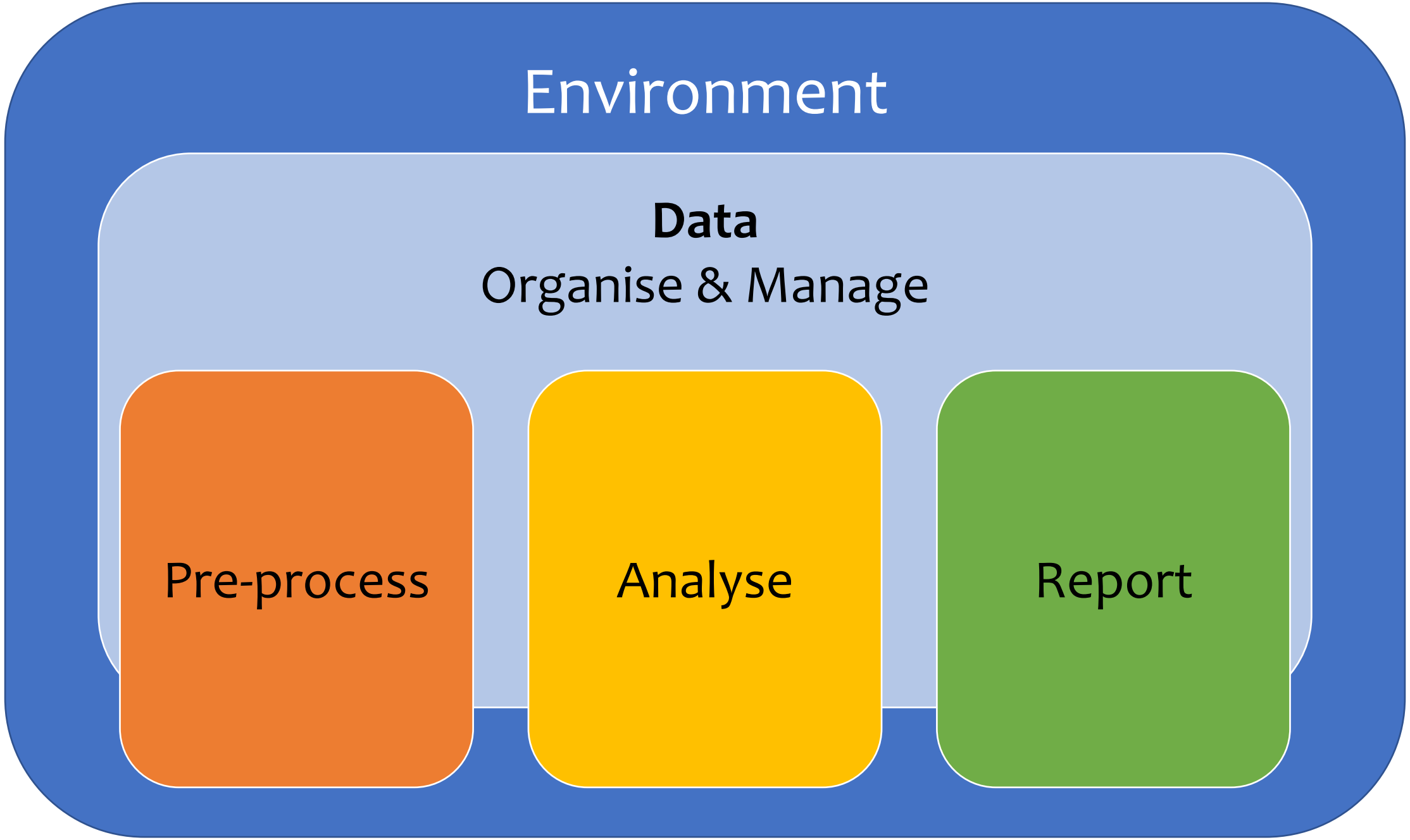
Anatomical image
Functional images
Event details

Pre-process & Analyse





The final push


Environment
Data
Organise & Manage
Pre-process
Analyse
Report

Example Dataset

https://openneuro.org/datasets/ds000117/versions/1.0.5

www.nature.com/scientificdata

SCIENTIFIC DATA

OPEN

SUBJECT CATEGORIES

» Electroencephalography -EEG

» Brain imaging

» Functional magnetic resonance imaging

» Cognitive neuroscience

A multi-subject, multi-modal human neuroimaging dataset

Daniel G. Wakeman[1,2] & Richard N. Henson[2]

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/.

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

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

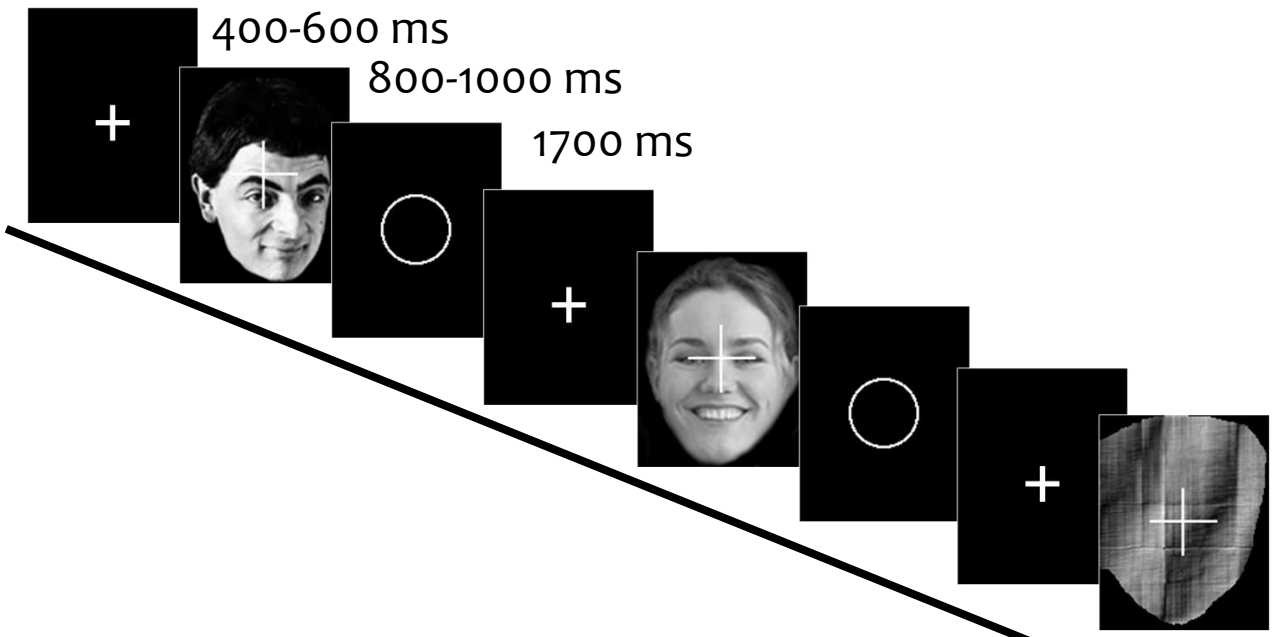
Example Dataset

https://openneuro.org/datasets/ds000117/versions/1.0.5

- 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

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

Experiment: Face Recognition



9 runs

7 min *long runs*

20s Rest after ever 50s

N = 16 subjects

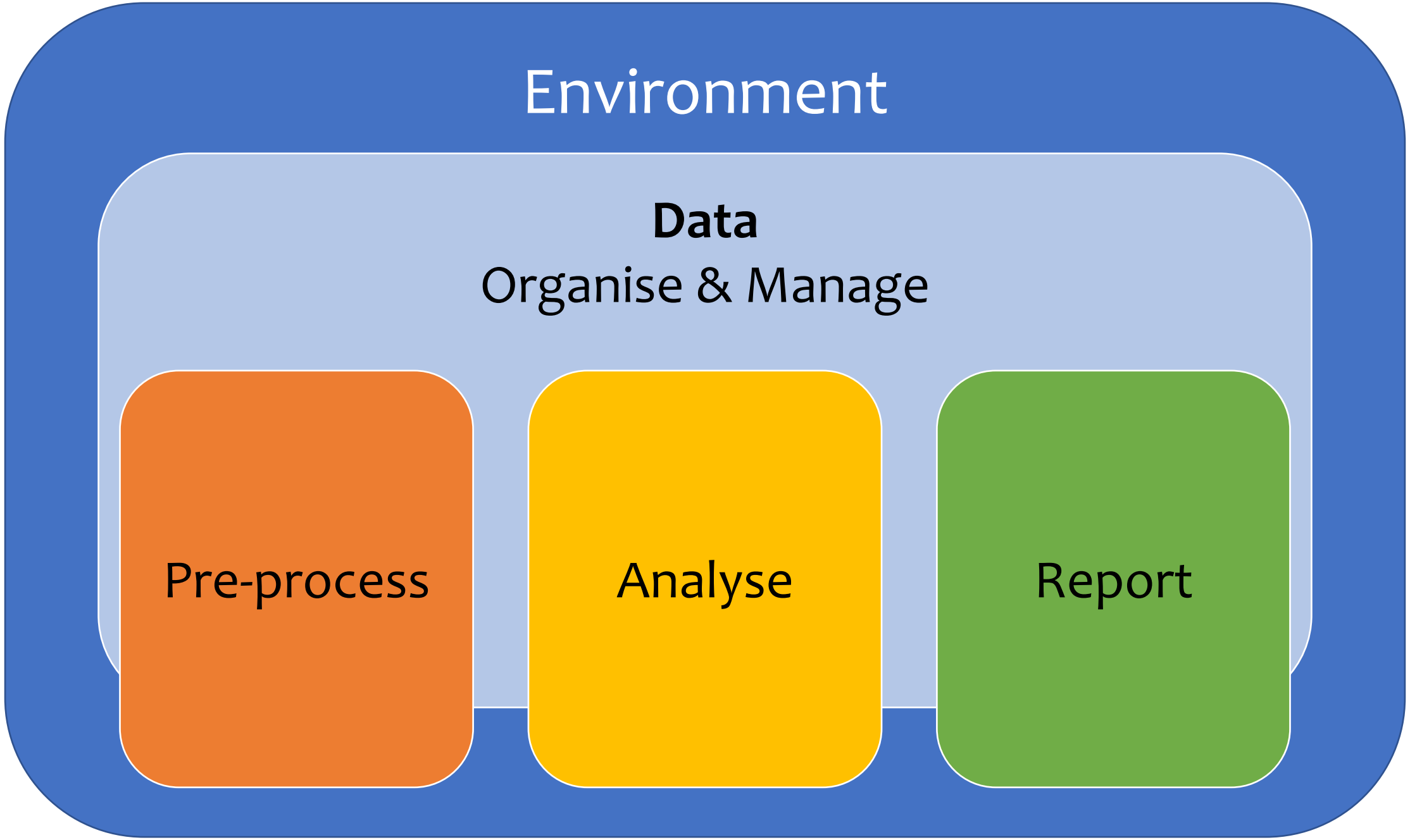
Stimuli: 3 types of greyscale face images:

~150 x Famous

~150 x Unfamiliar

~150 x Scrambled

Task: Judge face symmetry


Environment
Data
Organise & Manage
Pre-process
Analyse
Report

Environment

PROGRAMMING LANGUAGES

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

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

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 fMRI analysis environment that includes packages needed for your analysis work.

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

BASH
&
Shell Scripts

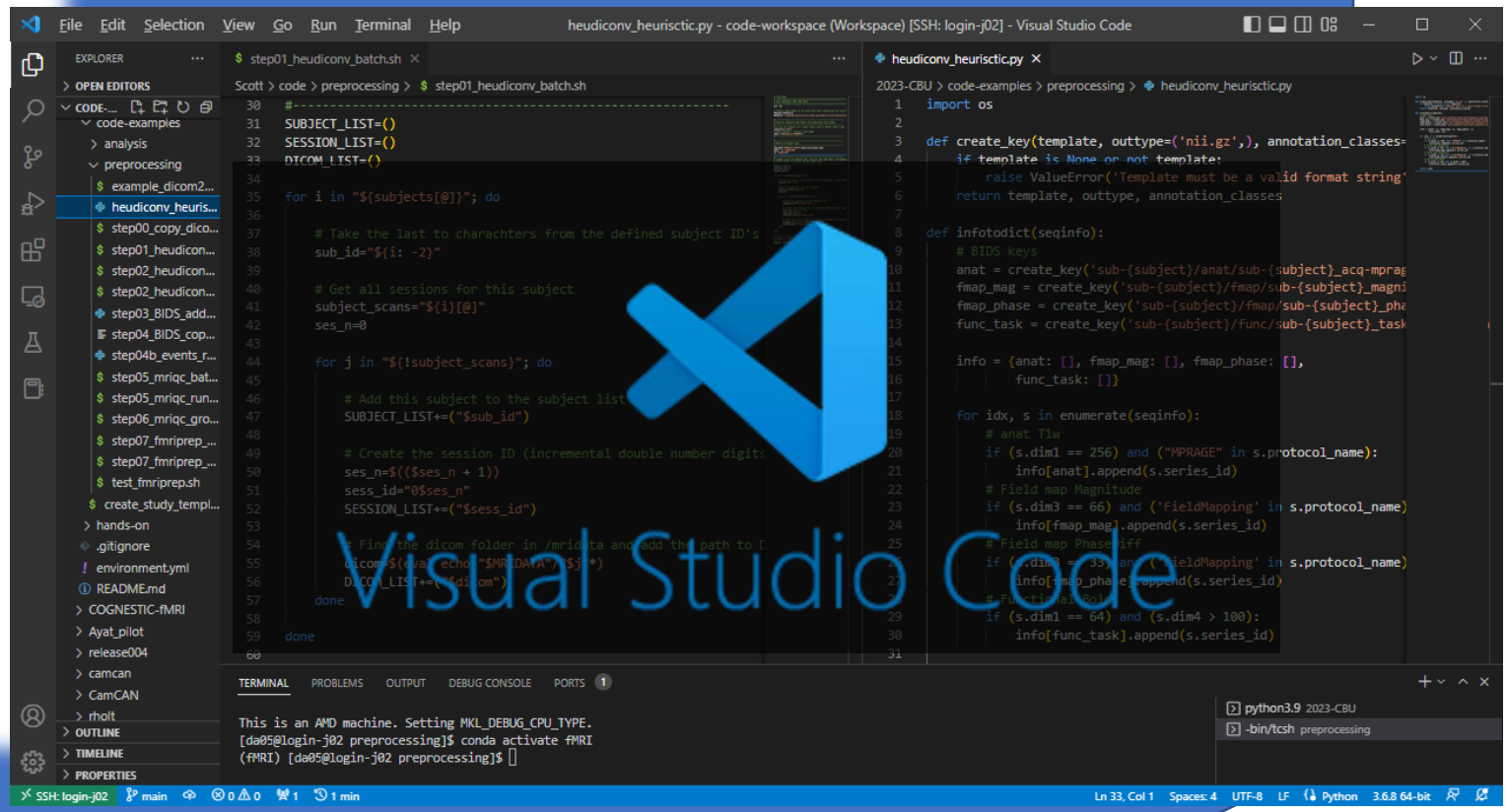
python™

MATLAB®

PACKAGE MANAGER

CONDA

CODE EDITOR



Environment

PROGRAMMING LANGUAGES

BASH
&
Shell Scripts

PACKAGE MANAGER

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.

Environment

PROGRAMMING LANGUAGES

BASH
&
Shell Scripts

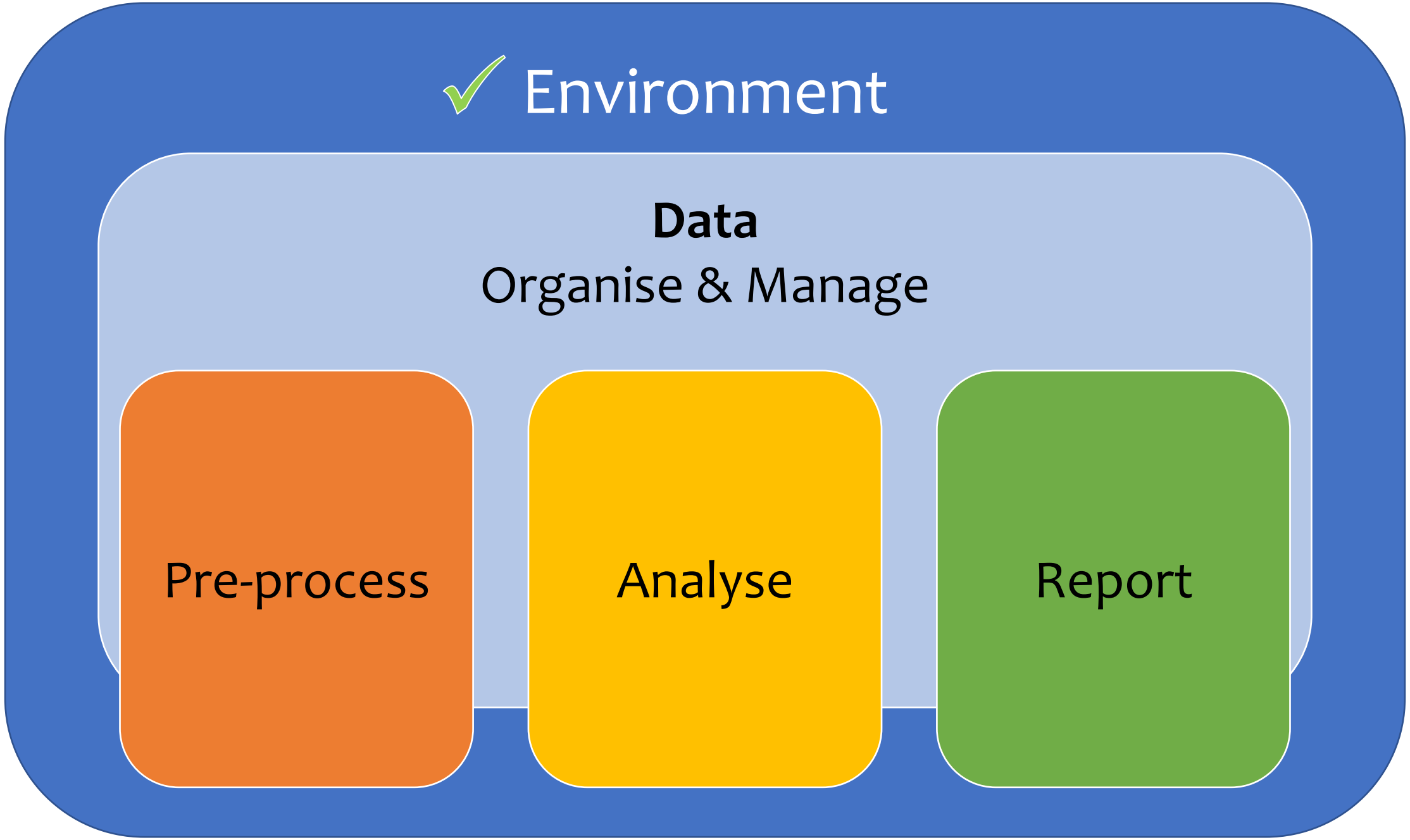
PACKAGE MANAGER

CODE EDITOR

VERSION CONTROL

ANALYSIS NOTEBOOK

01_Analysis_Environment.ipynb notebook

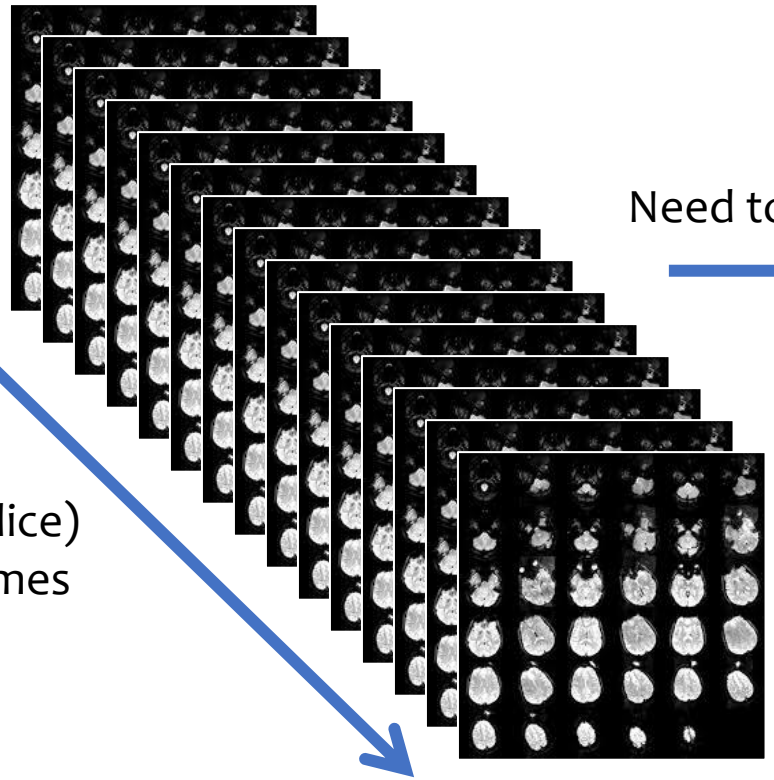

Environment
Data
Organise & Manage
Pre-process
Analyse
Report


Environment
Data
Organise & Manage

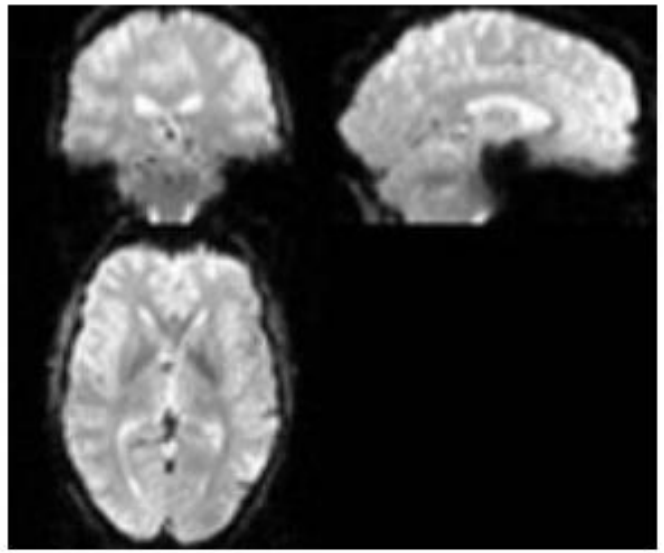
fMRI file formats

Collect the data

DICOM format



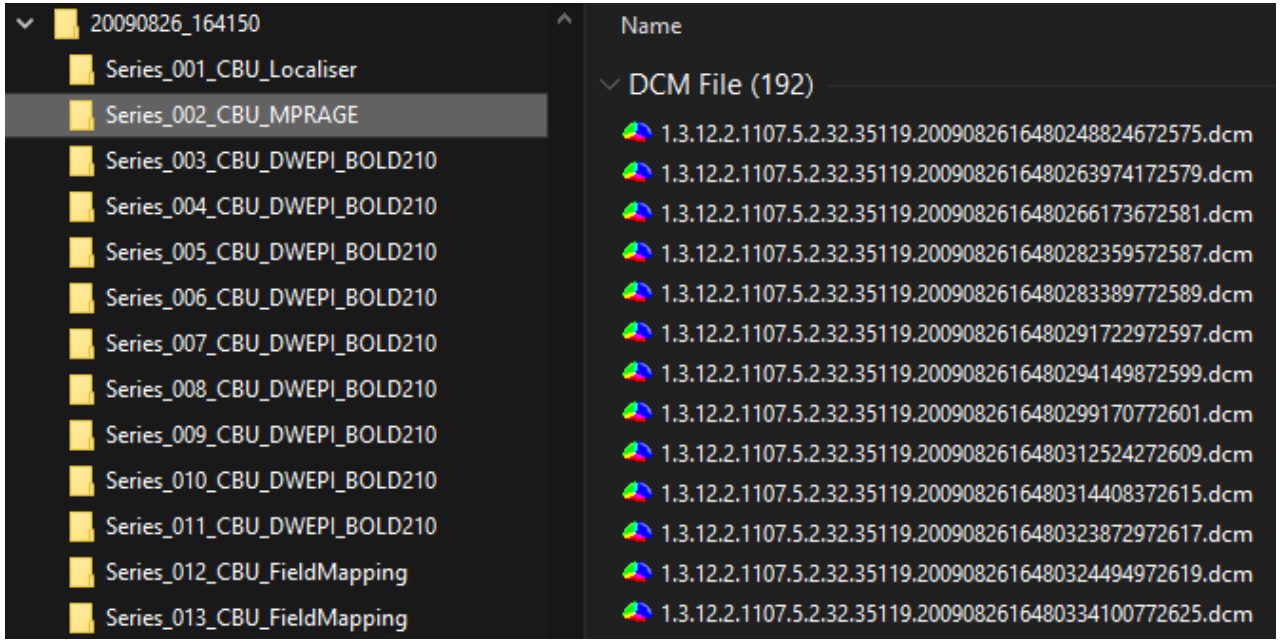
Need to convert to **NIfTI**



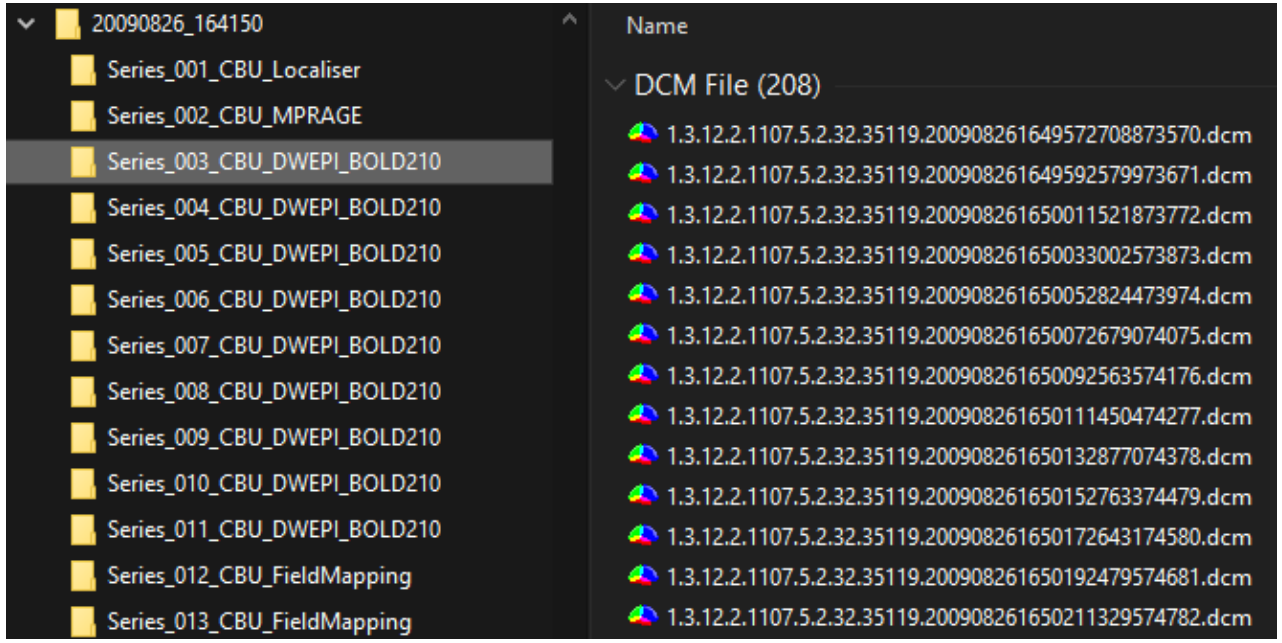
functional scan
A brain volume (slice-by-slice)
scanned every 2s > 100 times

DICOM

Anatomical (T1w) scans



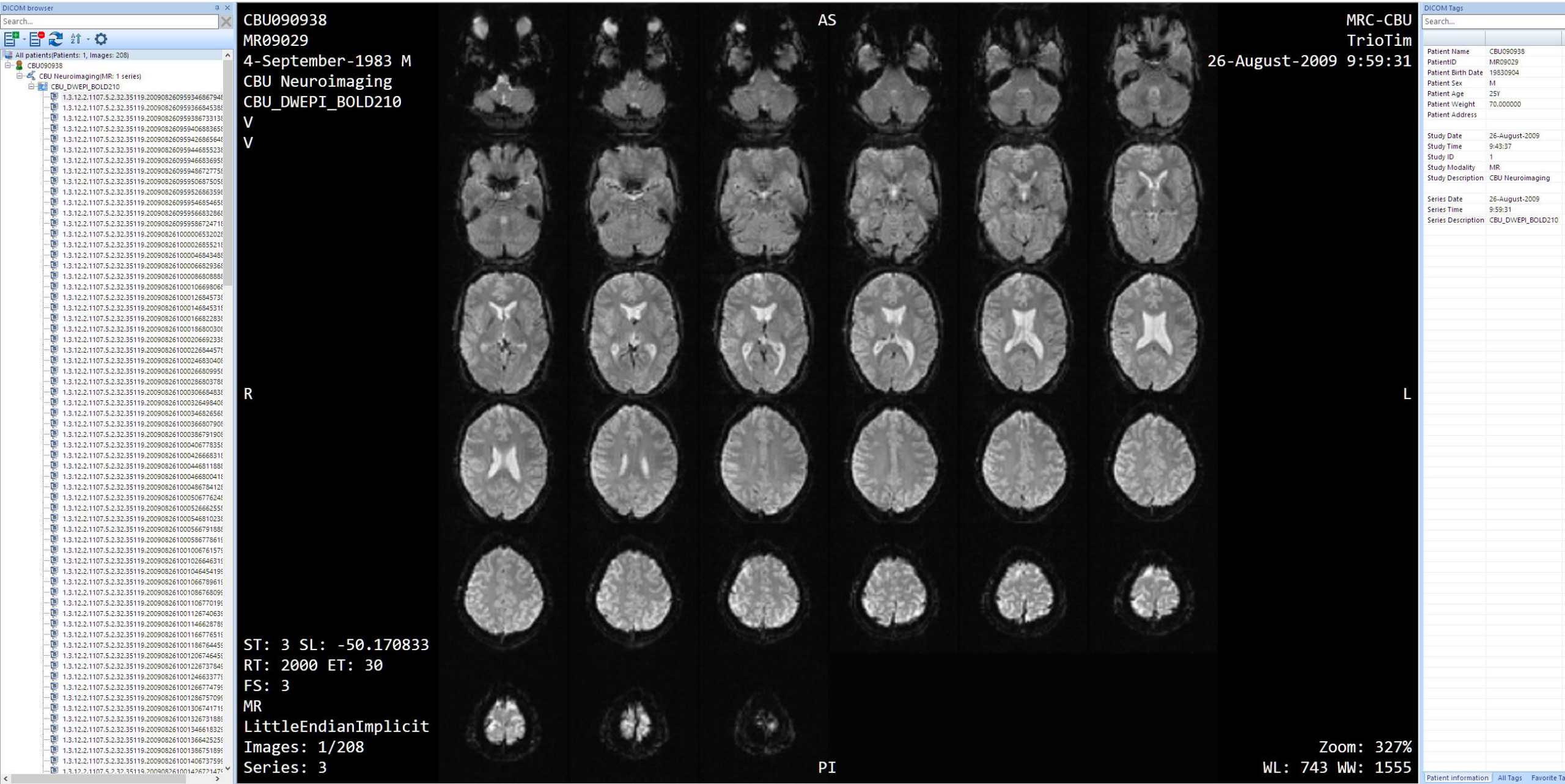
Functional (T2*/BOLD) scans



- **DICOM** - **D**igital **I**maging and **Co**mmunications in **M**edicine (.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 or volume

DICOM: Anatomical

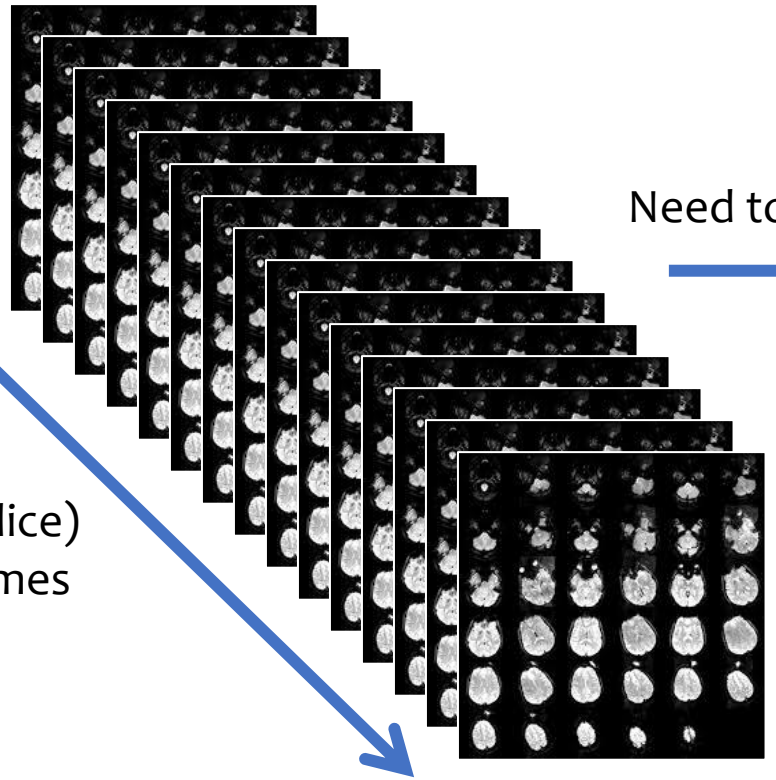
DICOM: Functional



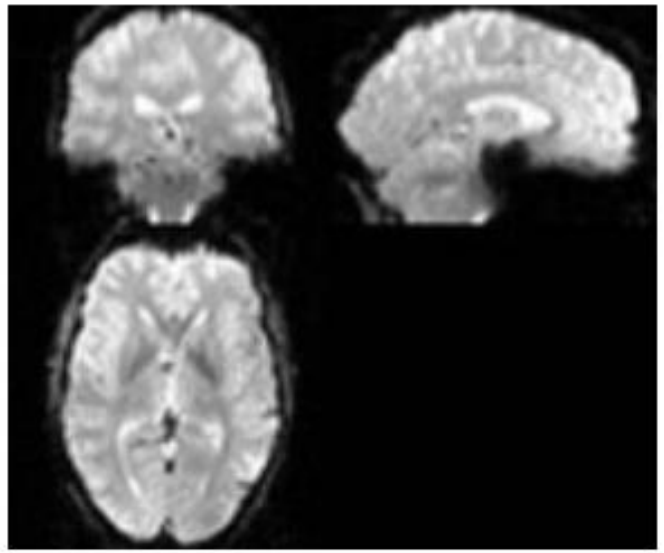
Anatomical (T1w) image & **Functional** (T2*/BOLD) image

Collect the data

DICOM format



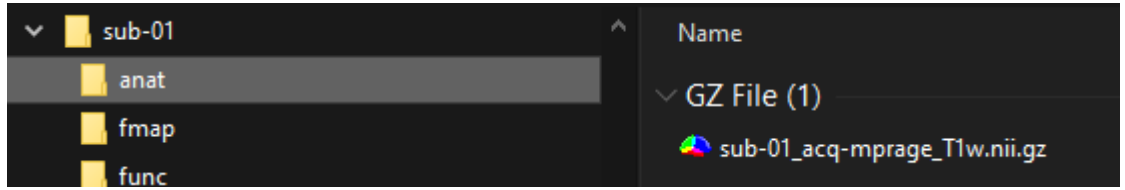
Need to convert to **NIfTI**



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

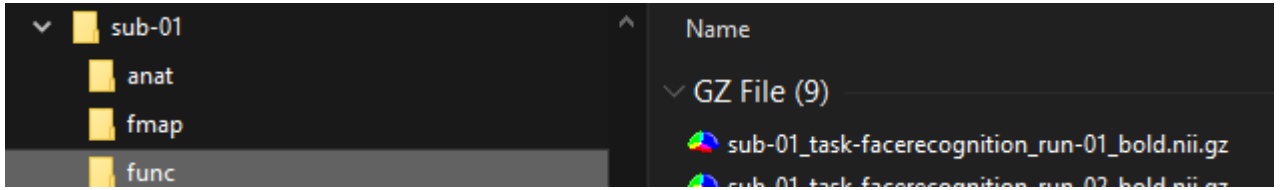
NIfTI

Anatomical (T1w) scans



3D file

Functional (T2*/BOLD) scans



4D file

- **NIfTI – N**euroimaging **I**n**f**ormatics **T**echnology **I**nitiative (.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

Collect the data

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

- sub-01_T1w.nii
- sub-01_run-01_bold.nii
- sub-01_run-02_bold.nii
- sub-02_T1w.nii
- sub-02_run-01_bold.nii
- sub-02_run-02_bold.nii
- …
- sub-100_T1w.nii
- sub-100_run-01_bold.nii
- sub-100_run-02_bold.nii
- and even more files (more sessions and tasks)

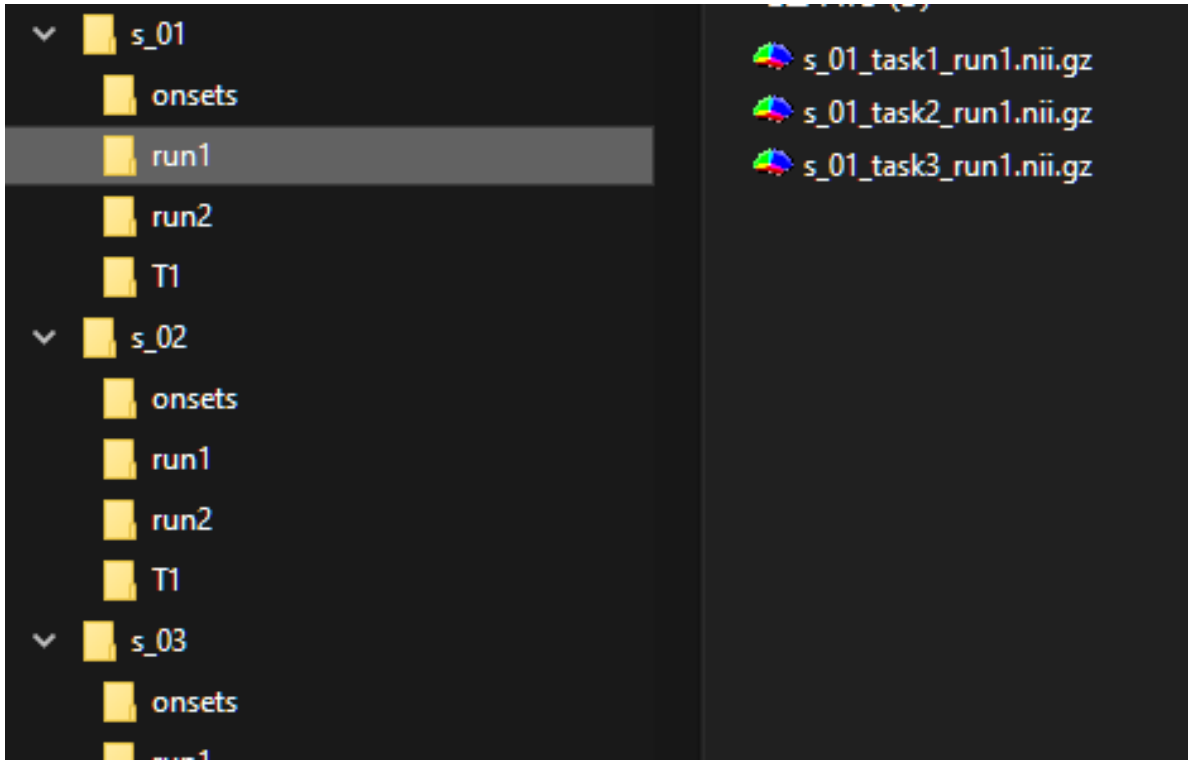
How should we organise our files?

File organisation

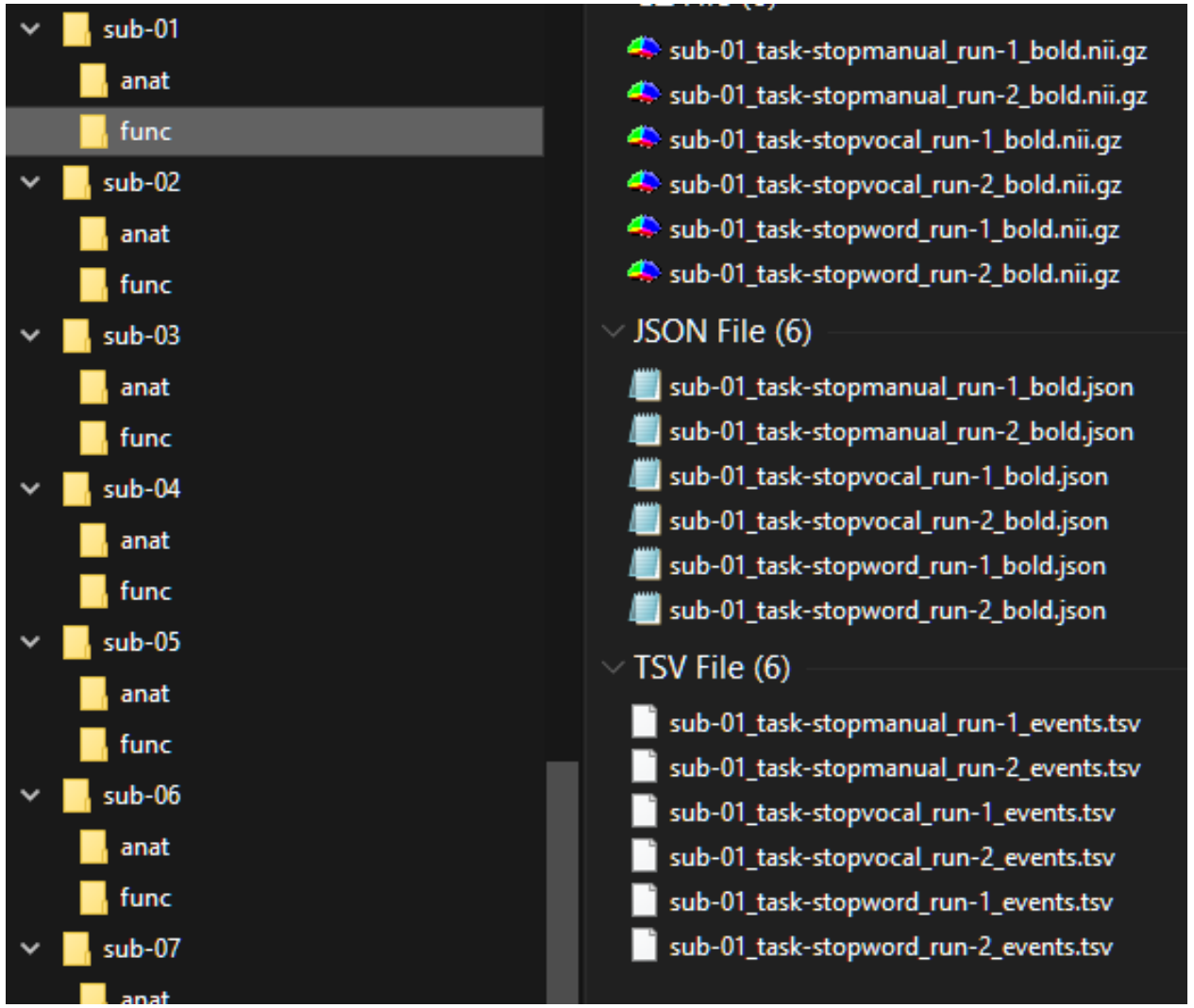
fMRI terminology

- Session
  - The time that the subject enters the scanner until they leave the scanner. This will usually include multiple scanning runs with different pulse sequences, including anatomical, functional, etc. Participant can be invited for a follow up session, next day or even later. That will then be Session 2.
- Run
  - A period of temporally continuous data acquisition using a single pulse sequence. Functional acquisitions are often split into multiple runs (5-10min) with brief breaks in between.
- Volume
  - A single 3D image acquired as part of a run. There is 1 anatomical volume and > 100 functional volumes.
- Condition
  - A set of task features that are created to engage a particular mental state. E.g., look at faces (condition 1), or look at houses (condition 2).
- Trial
  - A temporally isolated period during which a particular condition is presented, or a specific behaviour is observed. E.g., the first occurrence of the 'faces' condition is trial_1, the second occurrence is trial_2.
- Event
  - A trial can consist of multiple subunits. E.g., viewing faces trial may include pressing a button if you saw this face in the previous trials. Or working memory task may contain encoding, delay, retrieval. These subunits are labelled as 'events' and the 'trial' is defined as an overarching task.
- Block (or an 'epoch')
  - A temporarily contiguous period when a subject is presented with a particular condition.

Example 1



Example 2



fMRI data management

- **Problems with heterogeneity in data management**
  - Difficult for others (and you!) to understand your data and keep track of changes
  - Scripts have to be adapted (can't be easily reused)
  - Huge effort to automate workflows and no way to automatically validate data sets
  - Sharing data becomes a hustle

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

fMRI data management

- A standardised way for organising & describing neuroimaging data

Brain Imaging Data Structure - BIDS

BIDS
BRAIN IMAGING DATA STRUCTURE

Stanford | Center for Reproducible Neuroscience

- Documentation: https://bids-specification.readthedocs.io/en/latest/

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]

BIDS

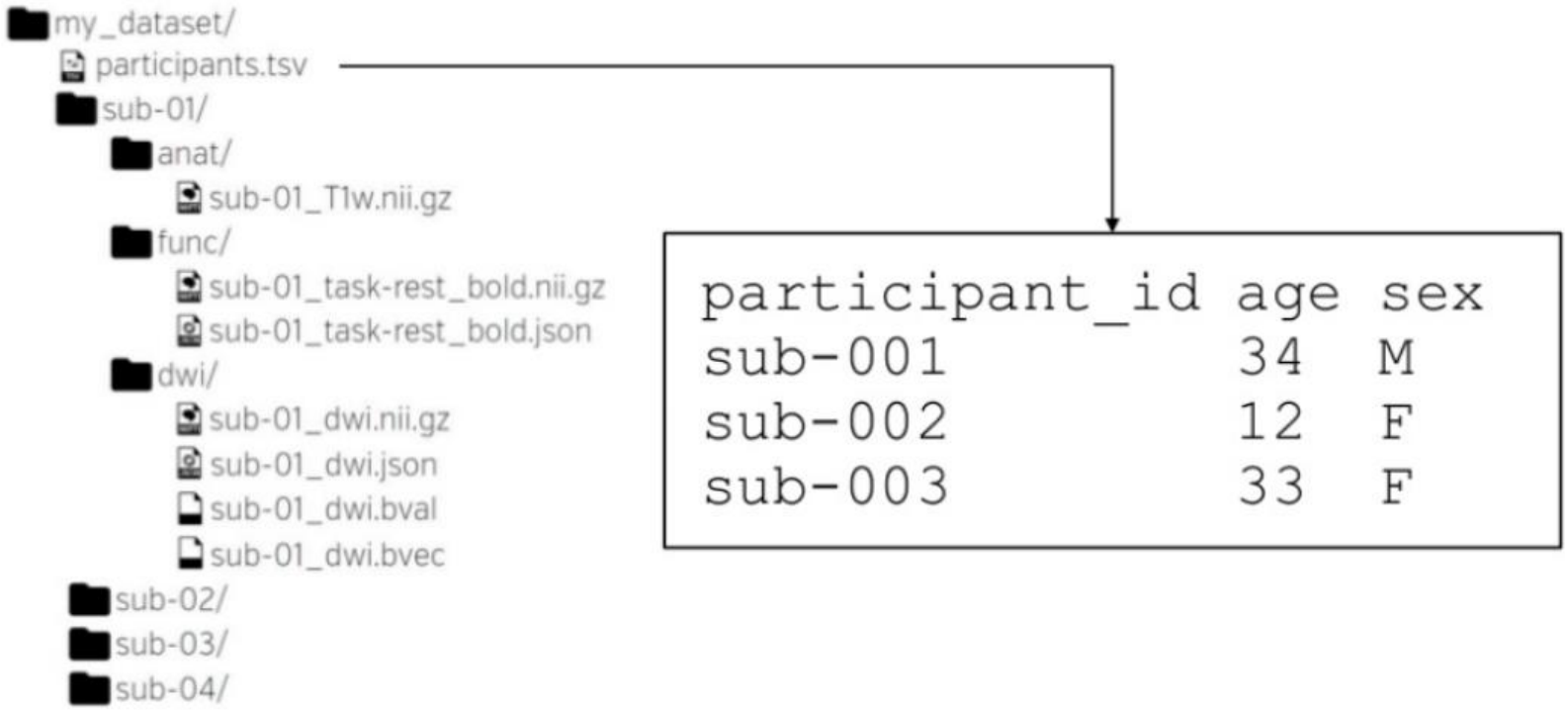
- **Benefits of BIDS**
  - Easy for other people to work on your data (for collaborations or contract changes)
  - Growing number of data analysis software packages that understand BIDS
  - Databases, such as OpenNeuro and LORIS etc., accept and export datasets organised according to BIDS
  - Validation tools that can check your dataset integrity and let you easily spot missing values

OpenNEURO

LORIS

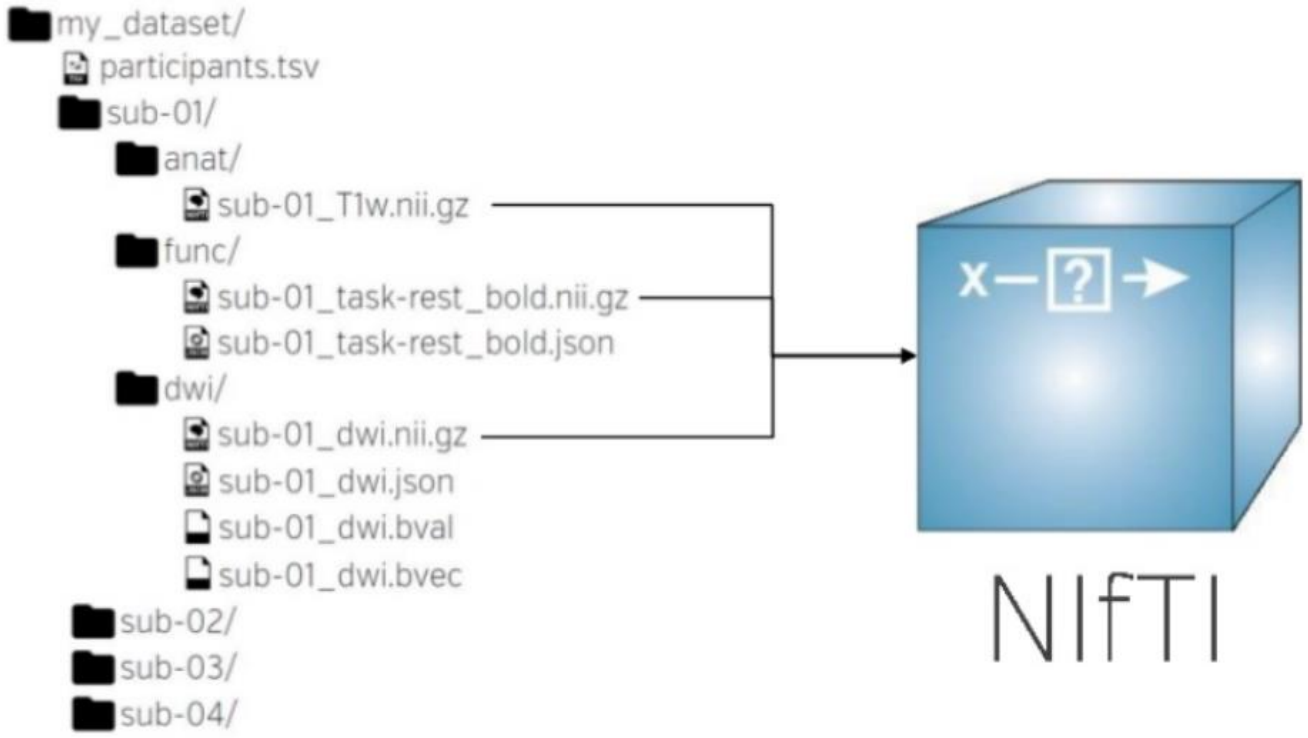
BIDS

- Contains participant information



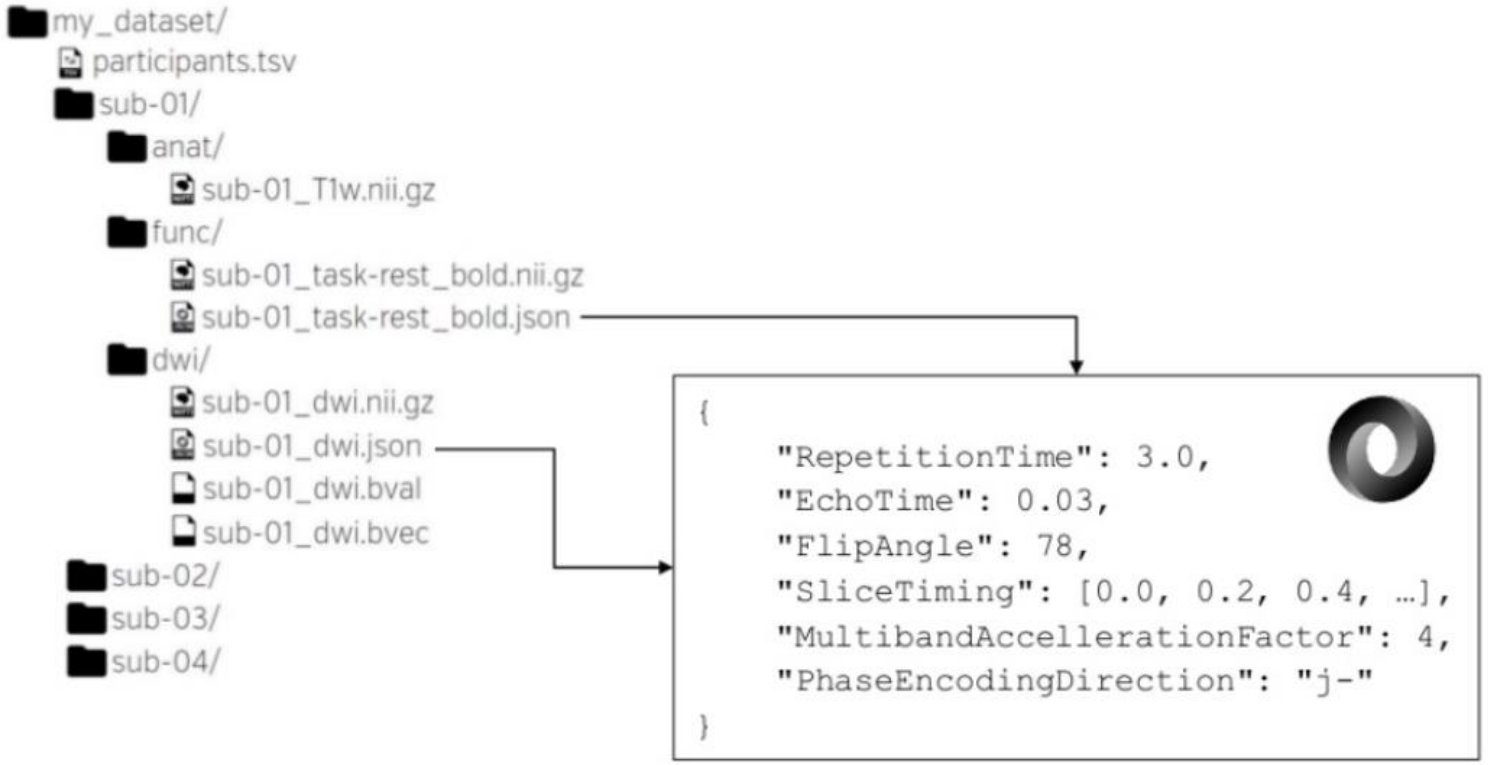
BIDS

- Contains data files



BIDS

- Contains study specific JSON (metadata) files



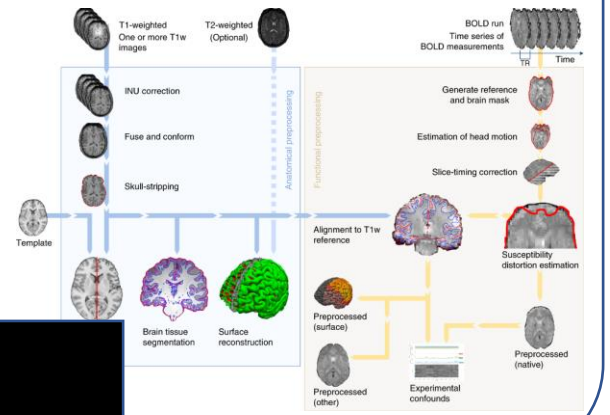
DICOM → BIDS

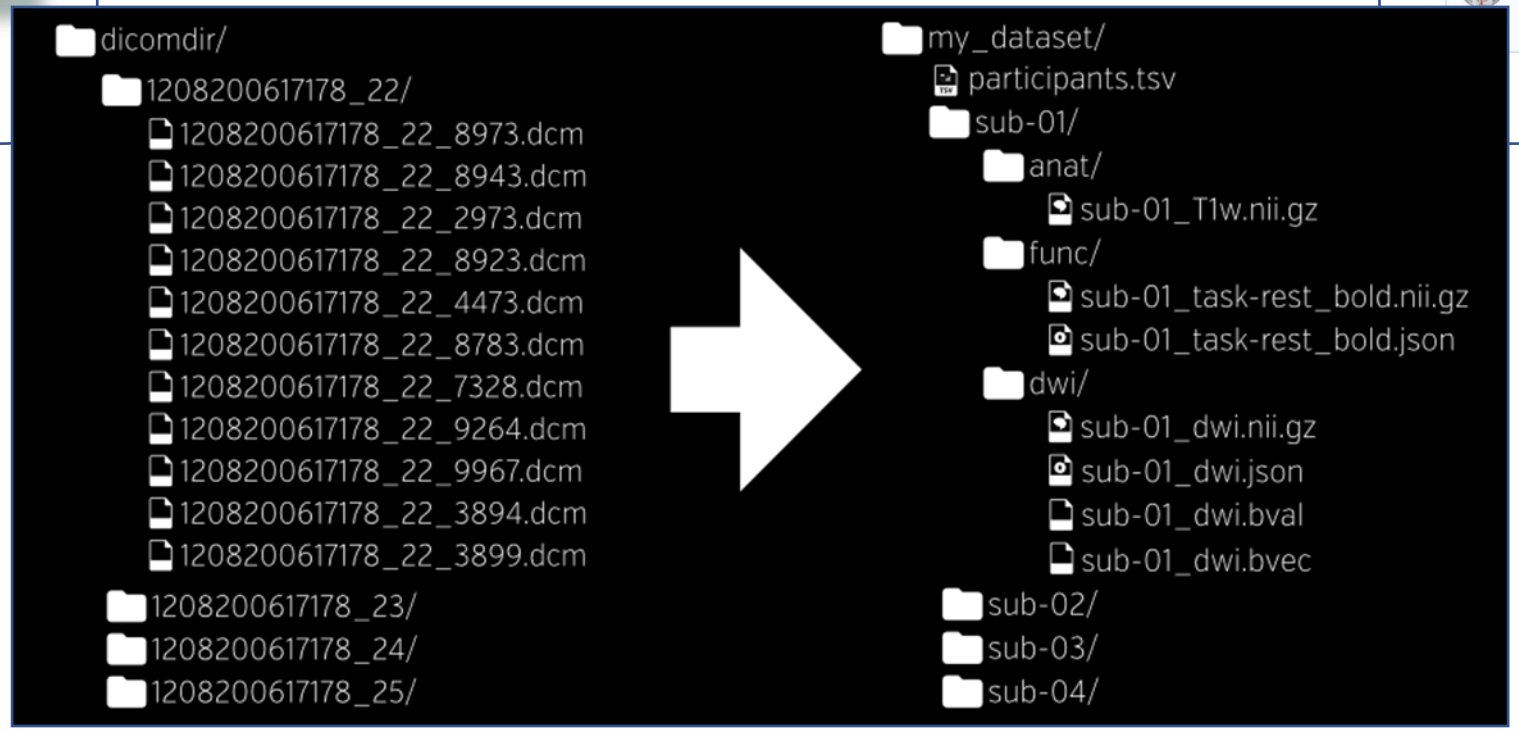
Collect the data

Many BIDS converters available

HeuDiConv (**Heu**ristic **Di**com **Conv**ersion)
provides sophisticated and flexible creation of BIDS datasets.

Pre-process & Analyse





Environment

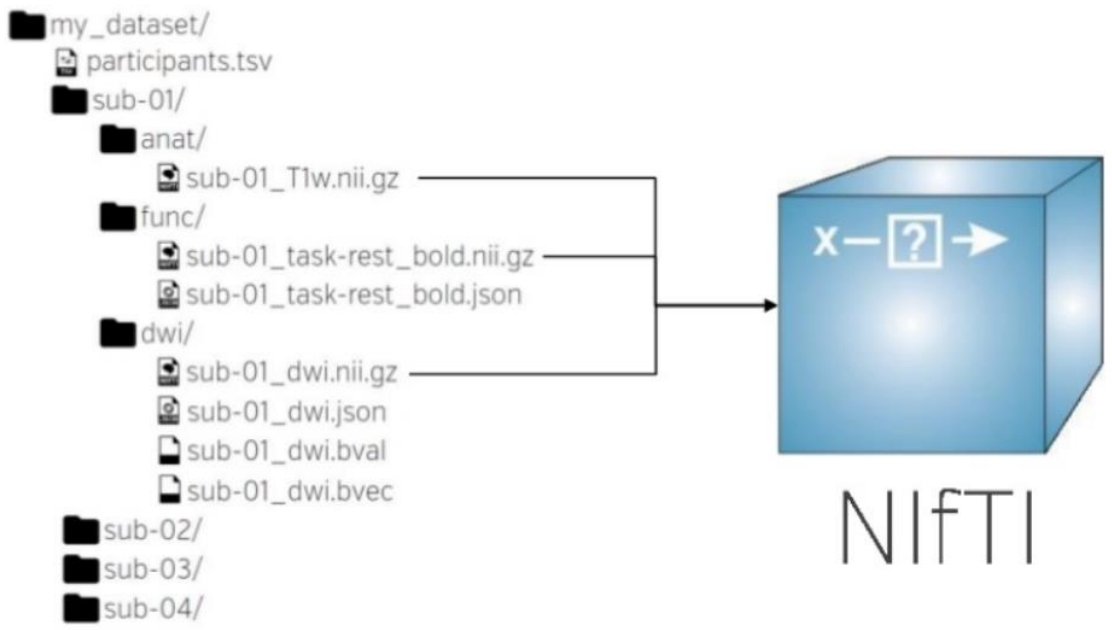
Data
Organise & Manage

Let's see the 02-fMRI_Data_Management.ipynb notebook

Explore BIDS dataset structure with PyBIDS

- Interactive notebook
  - https://github.com/dcdace/fMRI_training

Hands-On_01_PyBIDS.ipynb



Imaging data content

MRI data structure

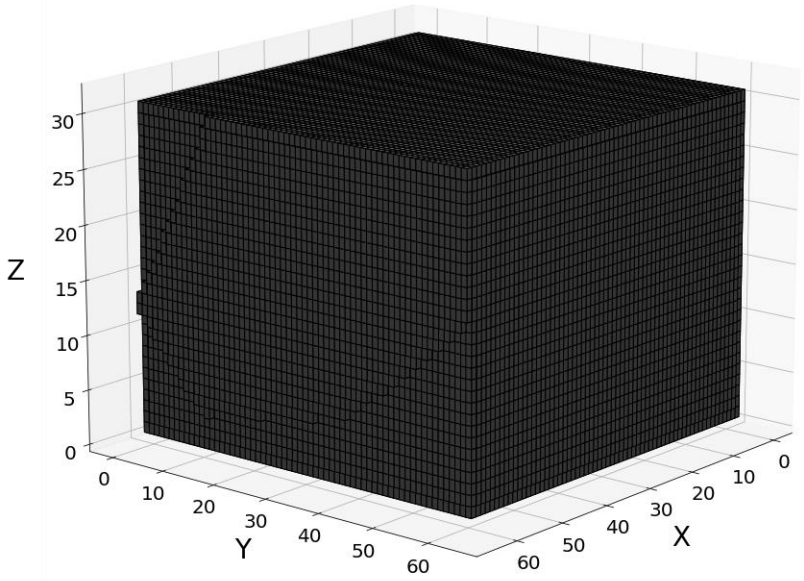
A 3D or 4D arrays of numbers

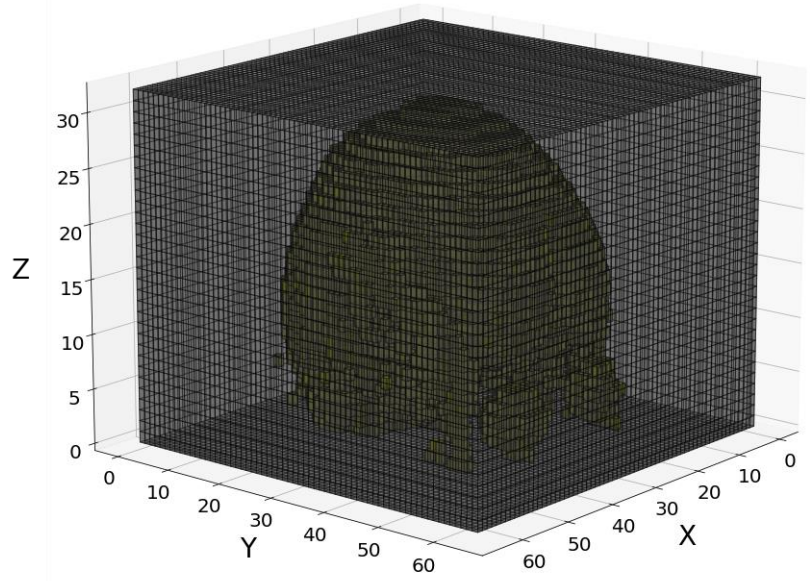
```
([[[ 0.,  0.,  0., ...,  0.,  0.,  0.],
   [ 0.,  0.,  0., ...,  0.,  0.,  0.],
   [ 0.,  0.,  0., ...,  0.,  0.,  0.],
   ...,
   [ 0.,  0.,  0., ...,  0.,  0.,  0.],
   [ 0.,  0.,  0., ...,  0.,  0.,  0.],
   [ 0.,  0.,  0., ...,  0.,  0.,  0.]],

  [[ 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.]],

  ...,
```

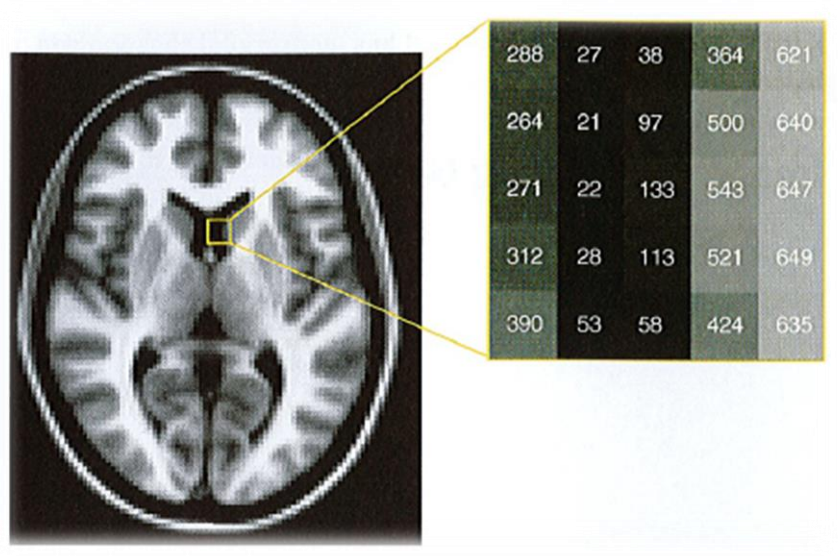
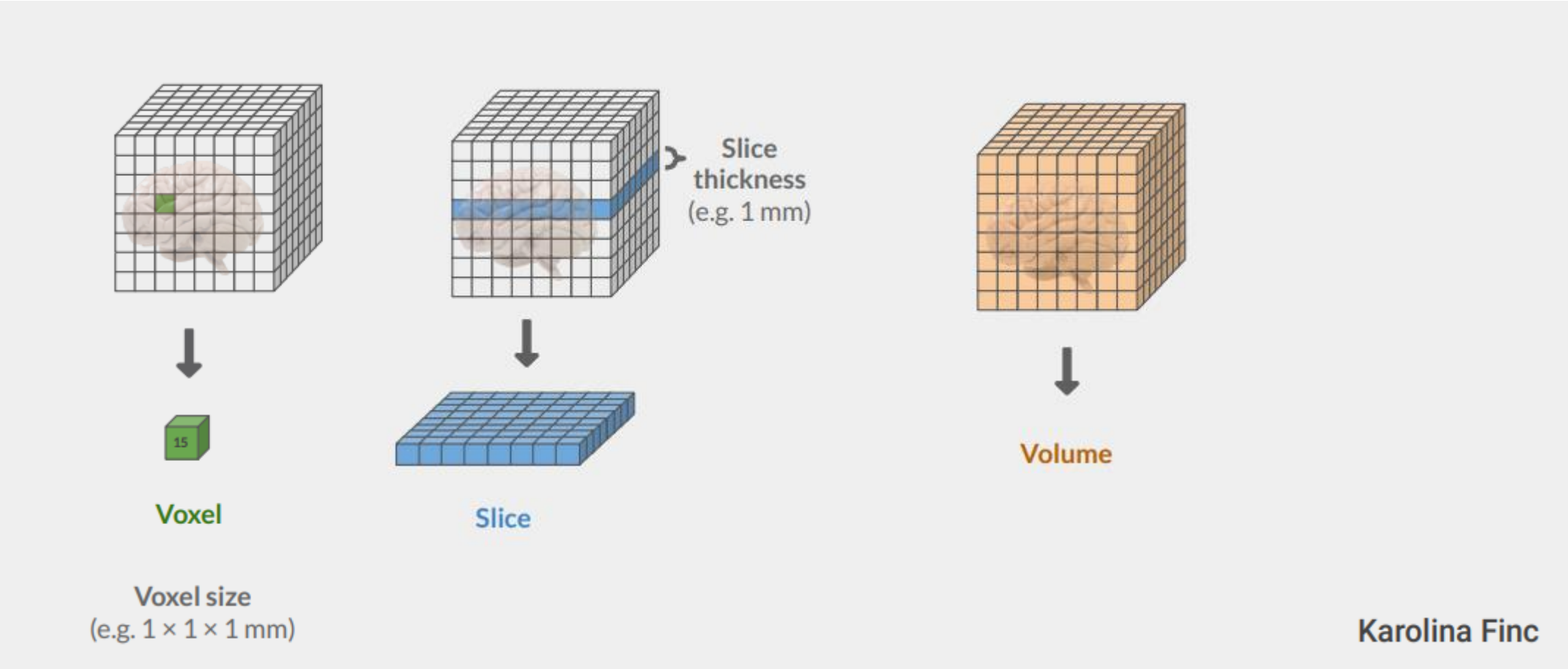


Image from Poldrack et al., 2011

MRI data structure

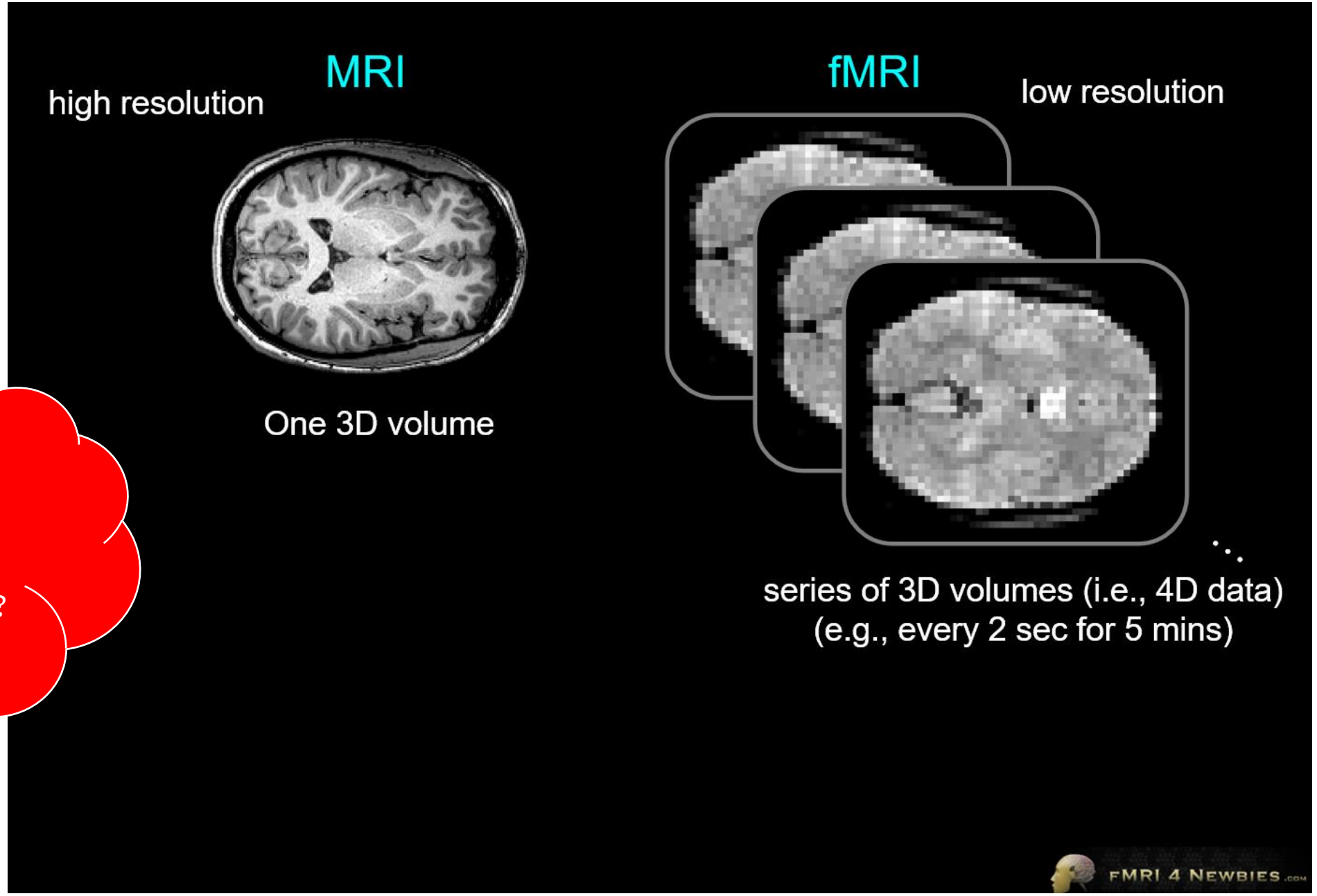


MRI data

MRI data

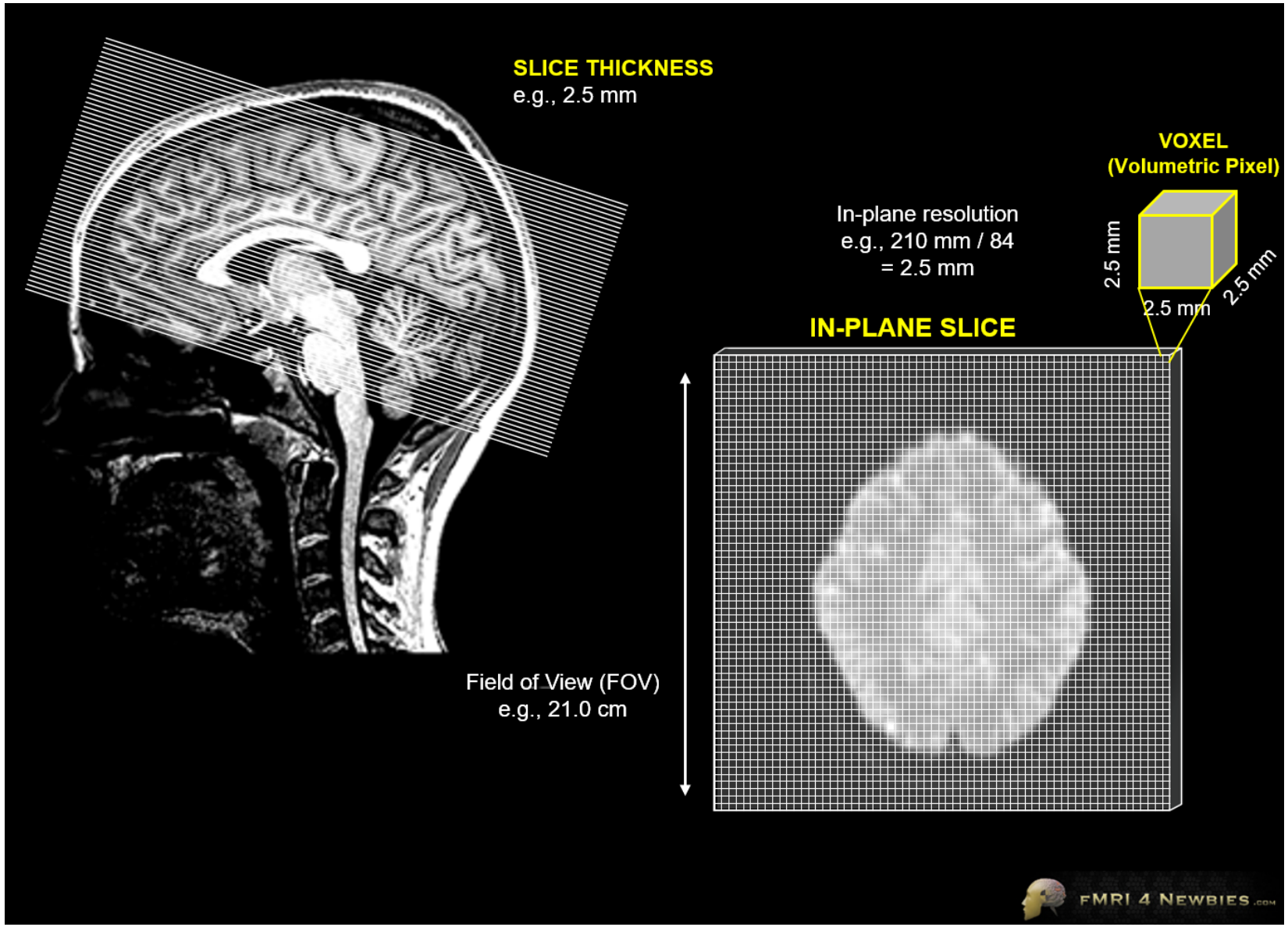
What determines the resolution?

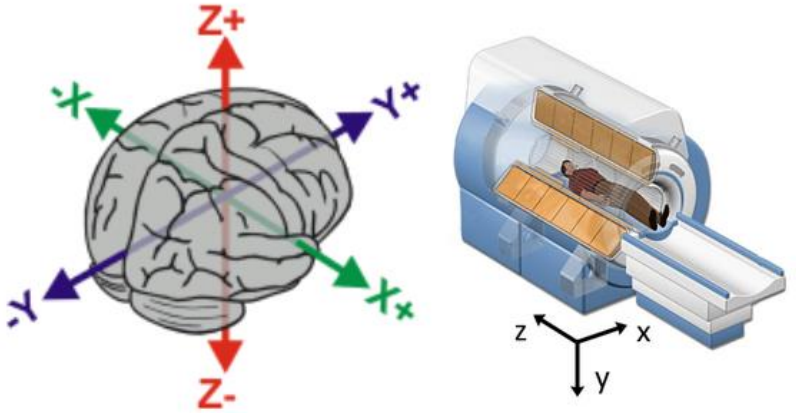
Why can't we acquire the functional images with higher resolution?



fMRI data

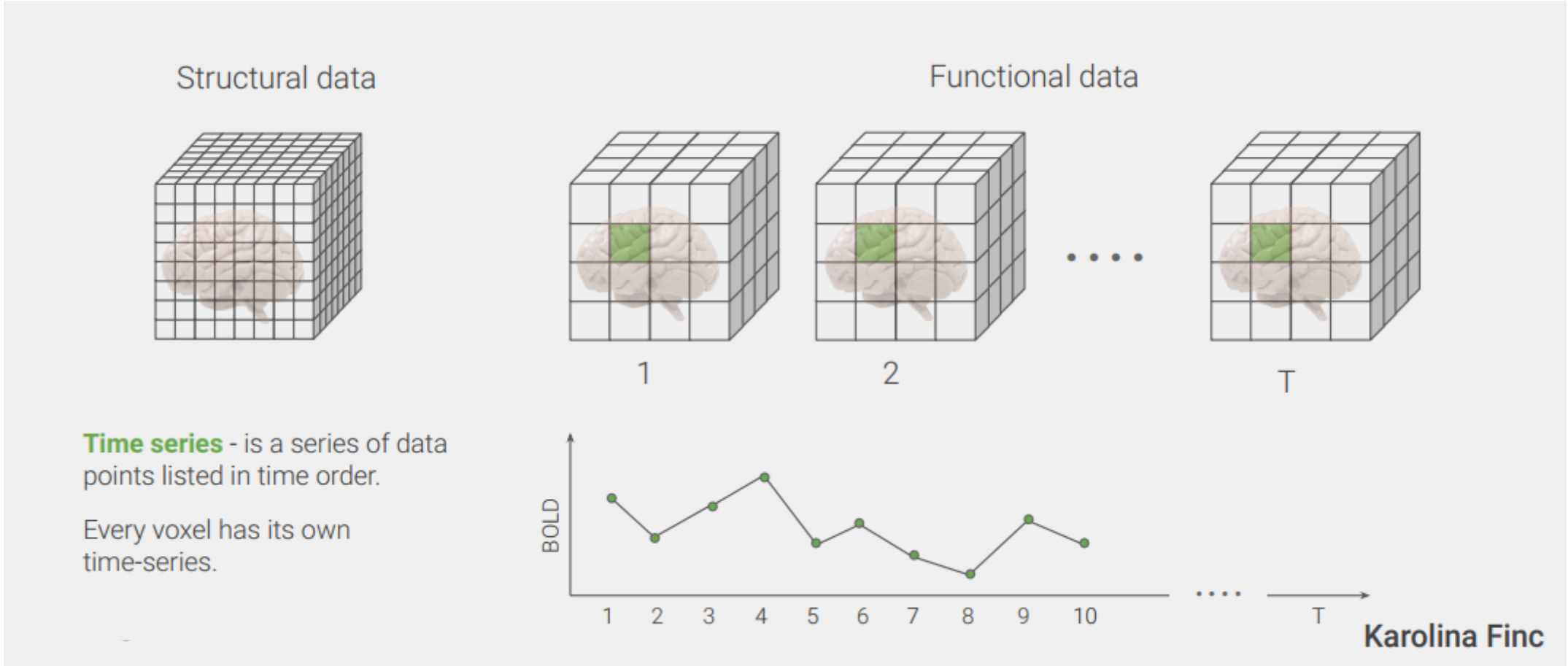
- Acquired in slices (usually axial; z-axis)





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

MRI data structure



MRI data content & manipulation

NiBabel

Access a cacophony of neuro-imaging file formats

Nilearn:

Statistics for NeuroImaging in Python

- Interactive notebook
  - https://github.com/dcdace/fMRI_training

Hands-On_02_Neuroimaging_data_manipulation.ipynb