



MRC Cognition  
and Brain  
Sciences Unit



UNIVERSITY OF  
CAMBRIDGE

# Functional Magnetic Resonance Imaging



**GitHub** [https://github.com/dcdace/fMRI\\_training](https://github.com/dcdace/fMRI_training)

Dace [datza] Apšvalka  
February 2025

# 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<sup>1,2</sup> & Richard N. Henson<sup>2</sup>

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/>.

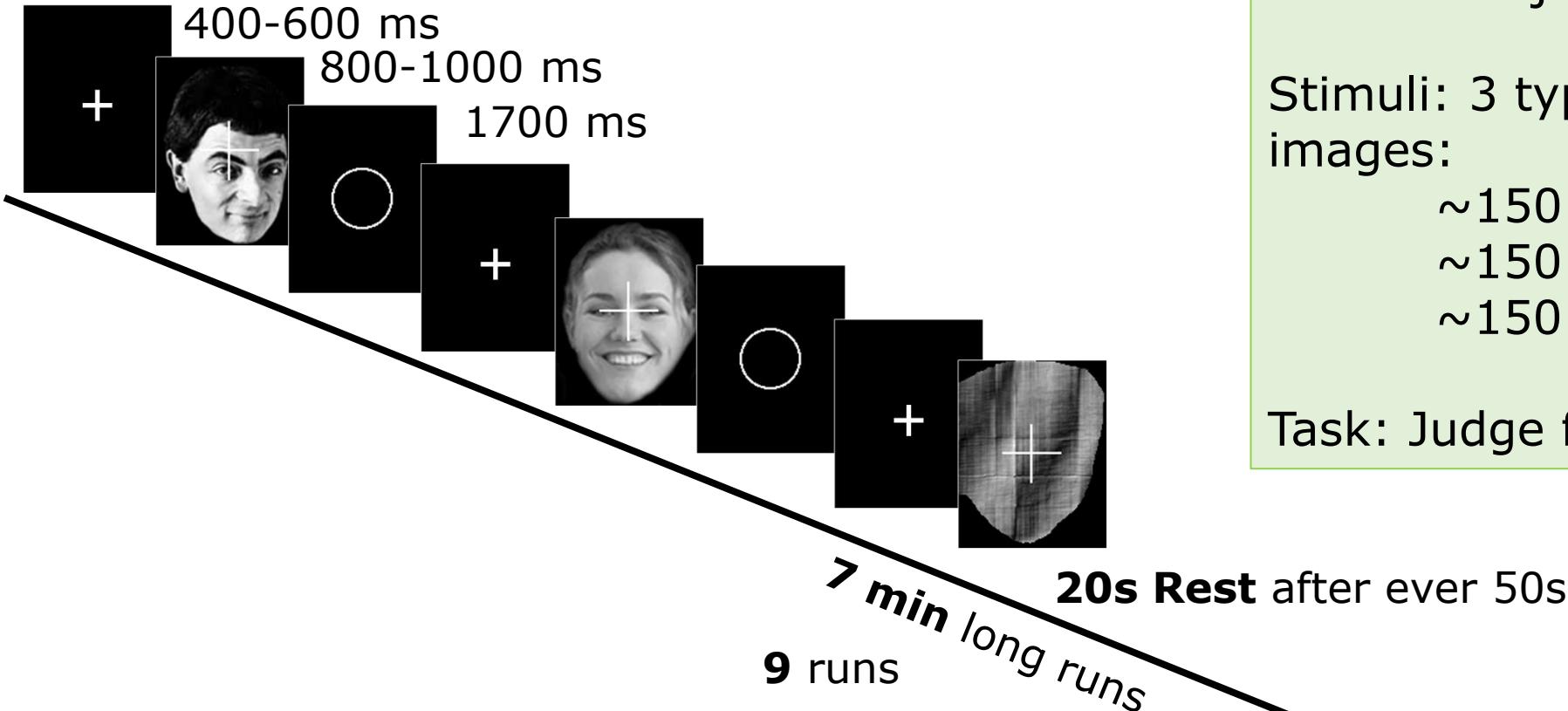
**Wakeman & Henson (2015), Scientific Data, <http://www.nature.com/articles/sdata20151>**

# Example Dataset

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

# 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

CONDA

## 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.

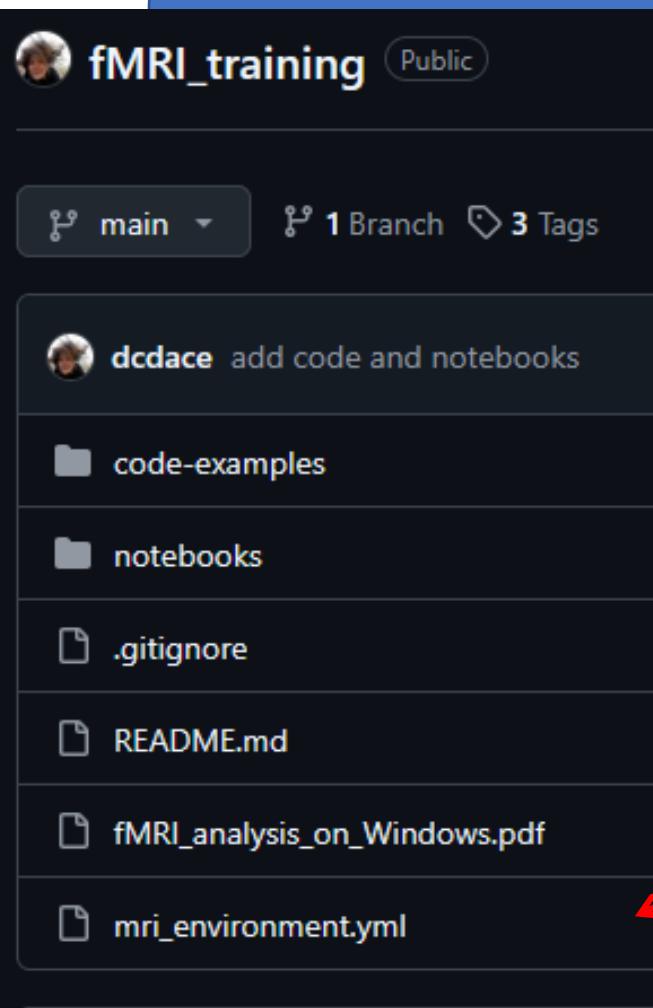
You can, for example, create your **fMRI analysis environment** that includes packages needed for your analysis work.

G LANGUAGES

ASH  
&  
I Scripts

thon™

ATLAB®



```
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 visualization
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 **fMRI 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



The screenshot shows the Visual Studio Code interface with several windows open:

- File Explorer:** Shows a hierarchical view of files and folders, including notebooks like `nb06_Group-Level-Analysis.ipynb`, Python scripts like `first_level_script.py`, and shell scripts like `step08_first_level_analysis.sh`.
- Code Editors:** Two main code editors are visible:
  - Python Script:** `first_level_script.py` containing Python code for fMRI analysis using NiLearn.
  - Shell Script:** `step08_first_level_analysis.sh` containing a bash script for running the Python script.
- Terminal:** A terminal window at the bottom showing the command `squeue -u da05` outputting a list of pending jobs on a cluster.

A large blue 'X' watermark is overlaid across the center of the interface.

# Visual Studio Code

## Introduction to fMRI

Materials for the "Introduction to fMRI" workshop at the MRC CBU

Instructor: Dr Dace Apšvalka

### Slides

- Introduction
- Experimental Design
- Data Management
- Pre-processing
- Statistical Analysis

### Example scripts

Example scripts for the fMRI analysis steps are available in the `code` directory. They include:

- Step 1: Discovering what DICOM series there are in your raw data
- Step 2: Converting DICOM to BIDS with HeuDiConv
- Step 3: Filling in events files with trial types, onsets, duration etc.
- Step 4: Removing dummy scans from the functional volumes
- Step 5: **Step 6:** Data quality assessment with MRIQC
- Step 7: Pre-processing with fMRIPrep
- Step 8: First-level statistical analysis with NiLearn

### Tutorial/Hands-On Notebooks

You can access the hands-on notebooks in two ways:

SSH: login-j05  
python3.1 2024-C...  
-bin/tcsh



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



git

## 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

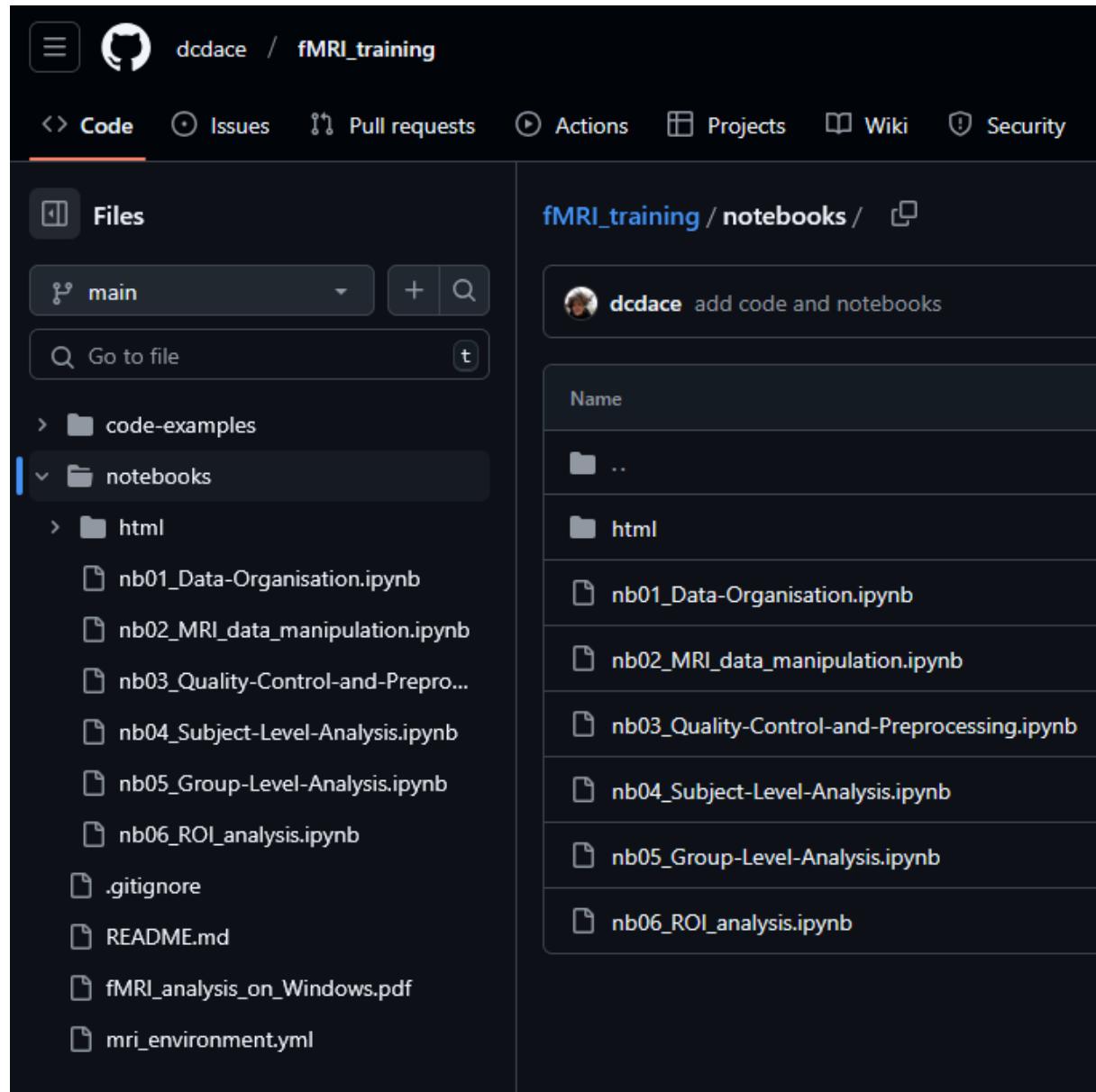


Visual Studio Code

git



# Analysis notebooks



A screenshot of a GitHub repository interface. The repository is named "fMRI\_training". The "Code" tab is selected. The left sidebar shows a file tree with the following structure:

- main
- code-examples
- notebooks (selected)
- html
  - nb01\_Data-Organisation.ipynb
  - nb02\_MRI\_data\_manipulation.ipynb
  - nb03\_Quality-Control-and-Prepro...
  - nb04\_Subject-Level-Analysis.ipynb
  - nb05\_Group-Level-Analysis.ipynb
  - nb06\_ROI\_analysis.ipynb
- .gitignore
- README.md
- fMRI\_analysis\_on\_Windows.pdf
- mri\_environment.yml

The right pane displays the contents of the "notebooks" folder. It shows a list of files with their names and preview icons:

- ..
- html
- 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
- nb06\_ROI\_analysis.ipynb



[https://github.com/dcdace/fMRI\\_training](https://github.com/dcdace/fMRI_training)

**Example data (~6GB):**

<https://cloud.mrc-cbu.cam.ac.uk/index.php/s/gUJZ6Ehl92Sm6X>  
(email me for password)



# Environment

**Data**

Organise & Manage

Pre-process

Analyse

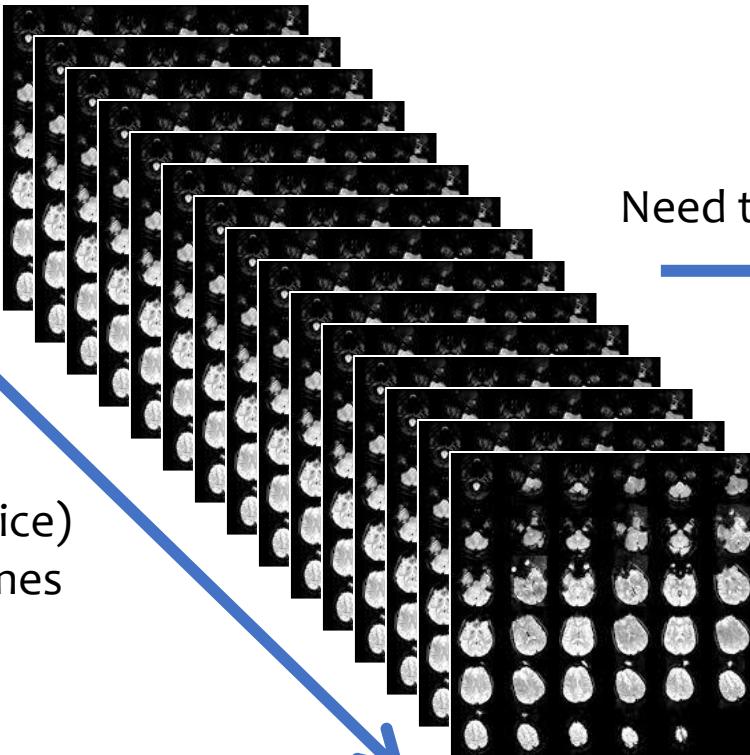
Report

# fMRI file formats

Collect the data



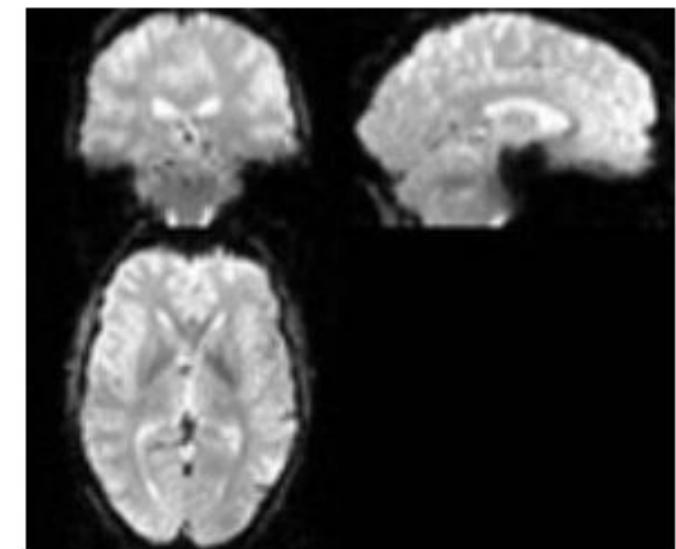
DICOM format



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
<b>Series_002_CBU_MPRAge</b>
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
<b>Series_003_CBU_DWEPI_BOLD210</b>
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  
1.3.12.2.1107.5.2.32.35119.200908260955594691084988.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  
1.3.12.2.1107.5.2.32.35119.2009082609555981609235012.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  
1.3.12.2.1107.5.2.32.35119.20090826095600358337385044.v



ST: 1 SL: -46.243837

RT: 2250 ET: 2.98

FS: 3

MR

LittleEndianImplicit

Images: 55/192

Series: 2

Zoom: 491%

WL: 352 WW: 759

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

CBU090938  
MR09029  
4-September-1983 M  
CBU Neuroimaging  
CBU\_DWEPI\_BOLD210

V  
V

R  
L

ST: 3 SL: -50.170833  
RT: 2000 ET: 30  
FS: 3  
MR  
LittleEndianImplicit  
Images: 1/208  
Series: 3

AS

PI

26-August-2009 9:59:31

MRC-CBU  
TrioTim

DICOM Tags	
Patient Name	CBU090938
PatientID	MR09029
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	9:59:31
Series Description	CBU_DWEPI_BOLD210

Zoom: 327%  
WL: 743 WW: 1555

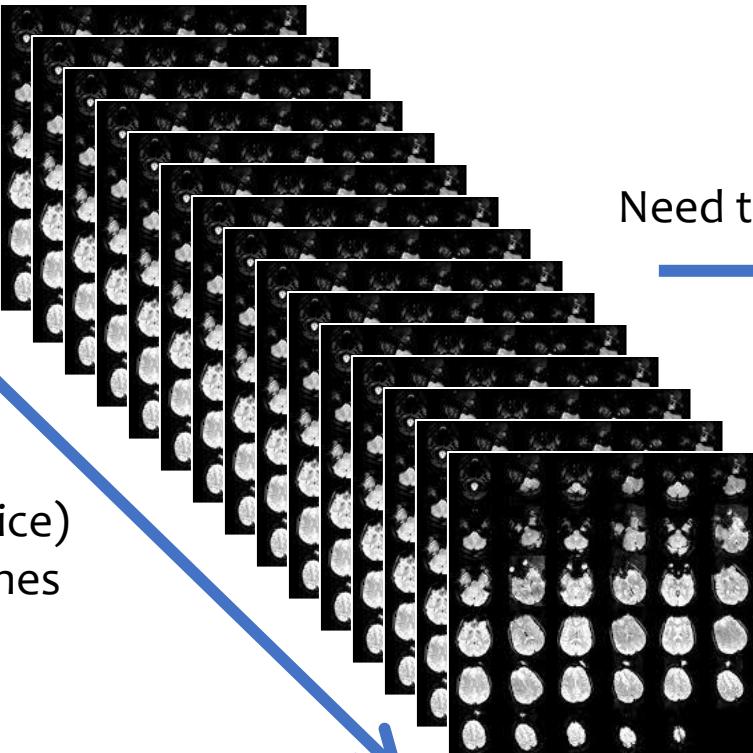
Patient information All Tags Favorite ↑

Collect the data



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

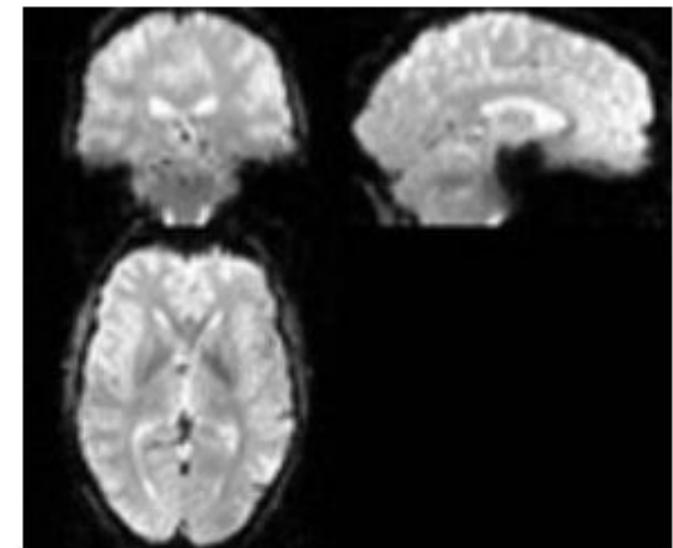
DICOM format



Need to convert to NIfTI

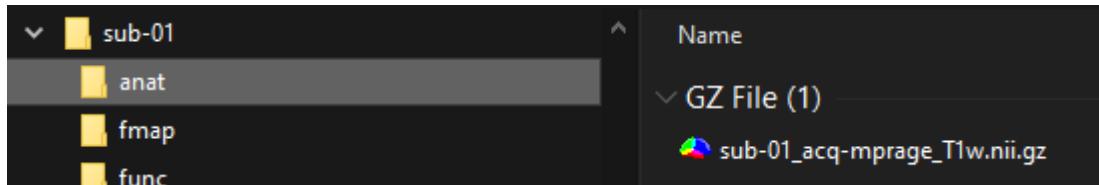
**functional scan**

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



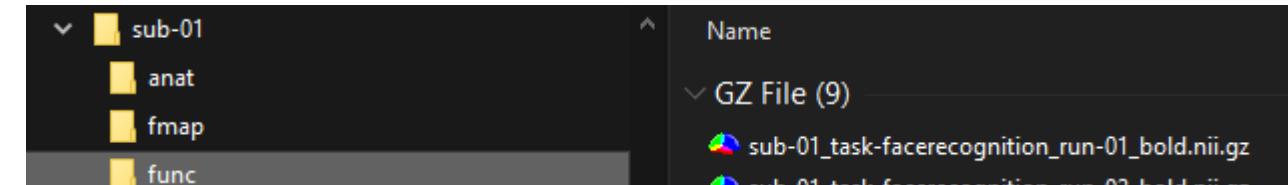
# NIfTI

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

Collect the data



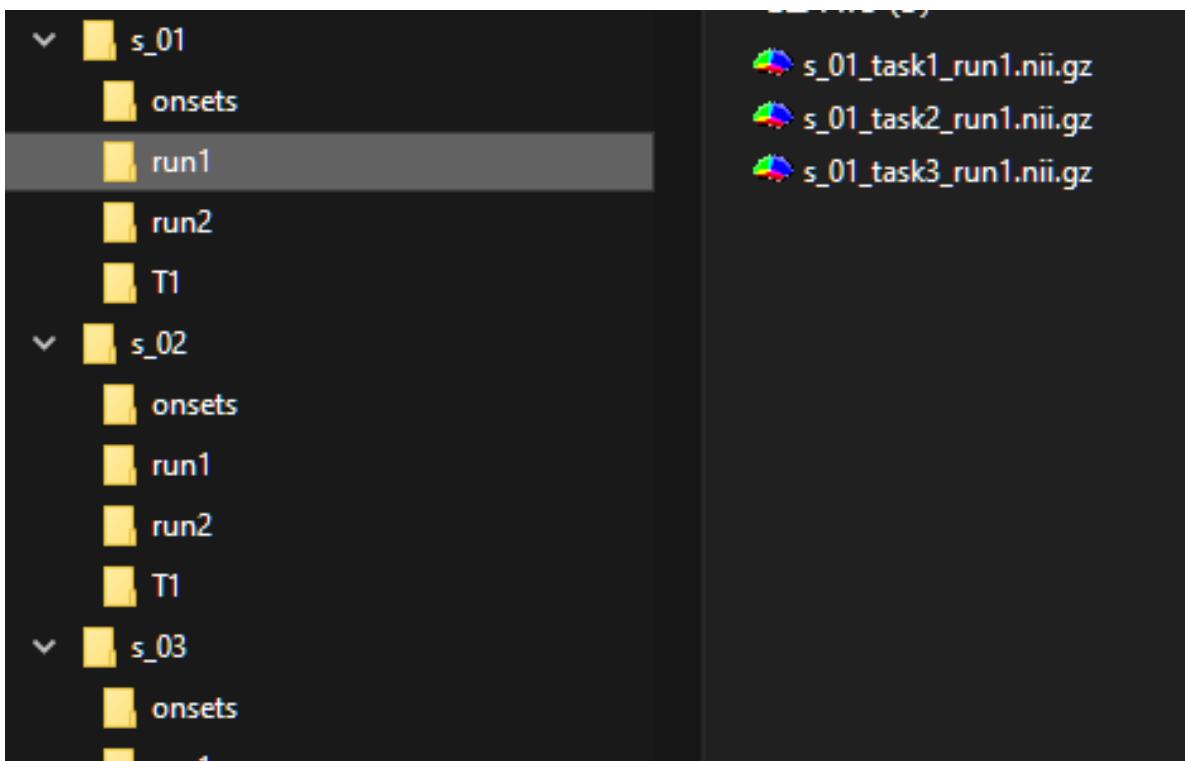
## Anatomical (T1w) image & Functional (T2\*/BOLD) image

- 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)

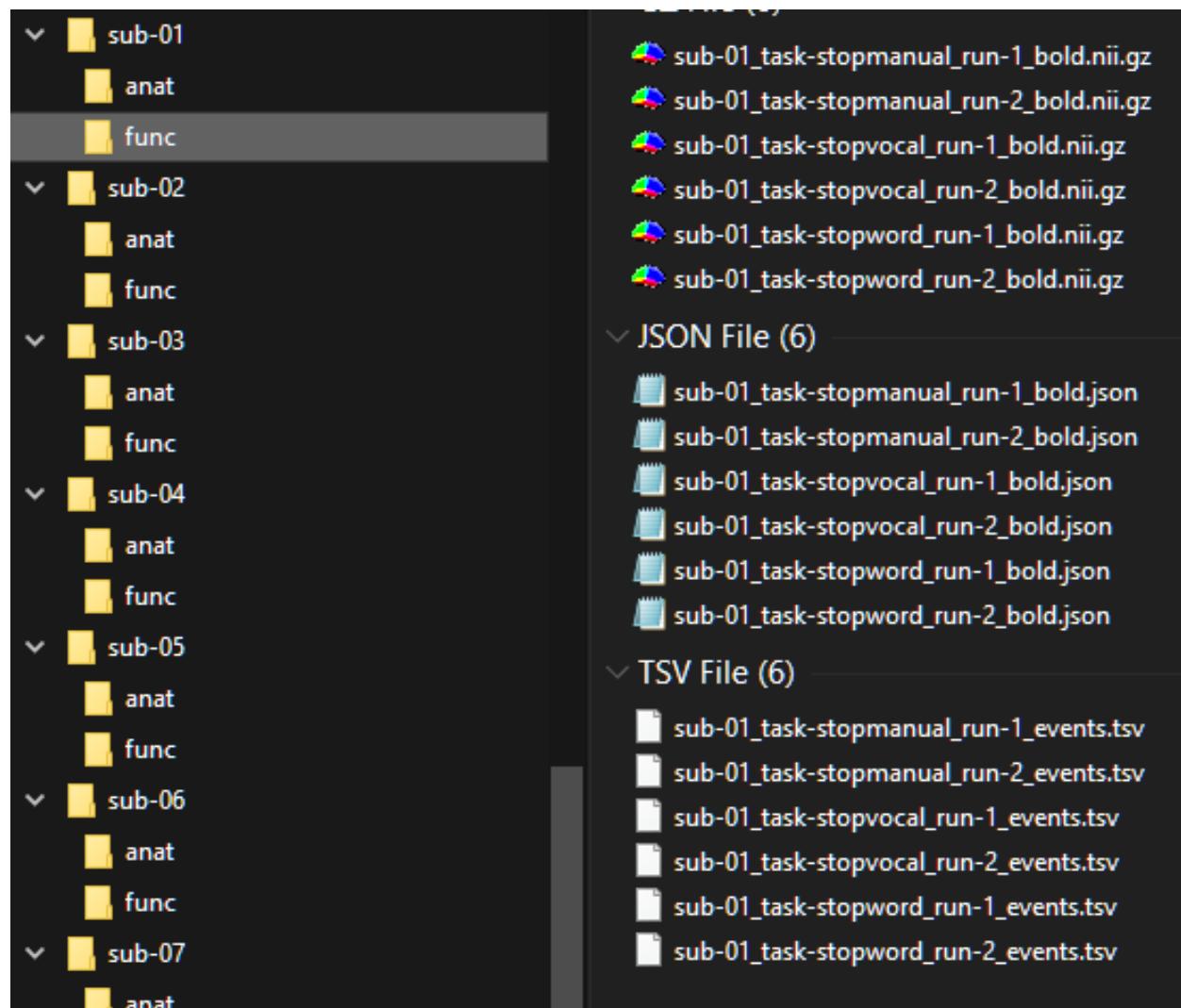
How should we  
organise our files?

# File organisation

## 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



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

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

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



RESEARCH ARTICLE

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

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





- **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



# BIDS

- Contains participant information

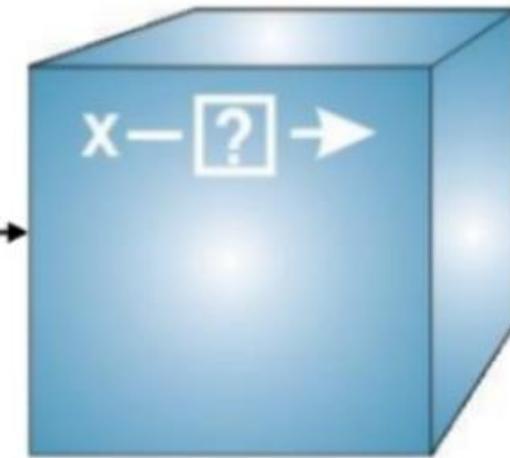
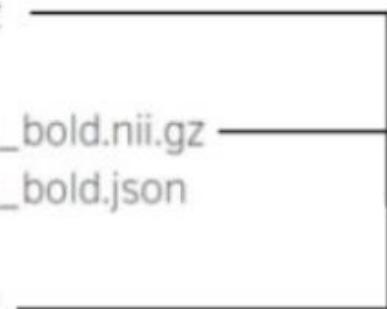


participant_id	age	sex
sub-001	34	M
sub-002	12	F
sub-003	33	F

# BIDS

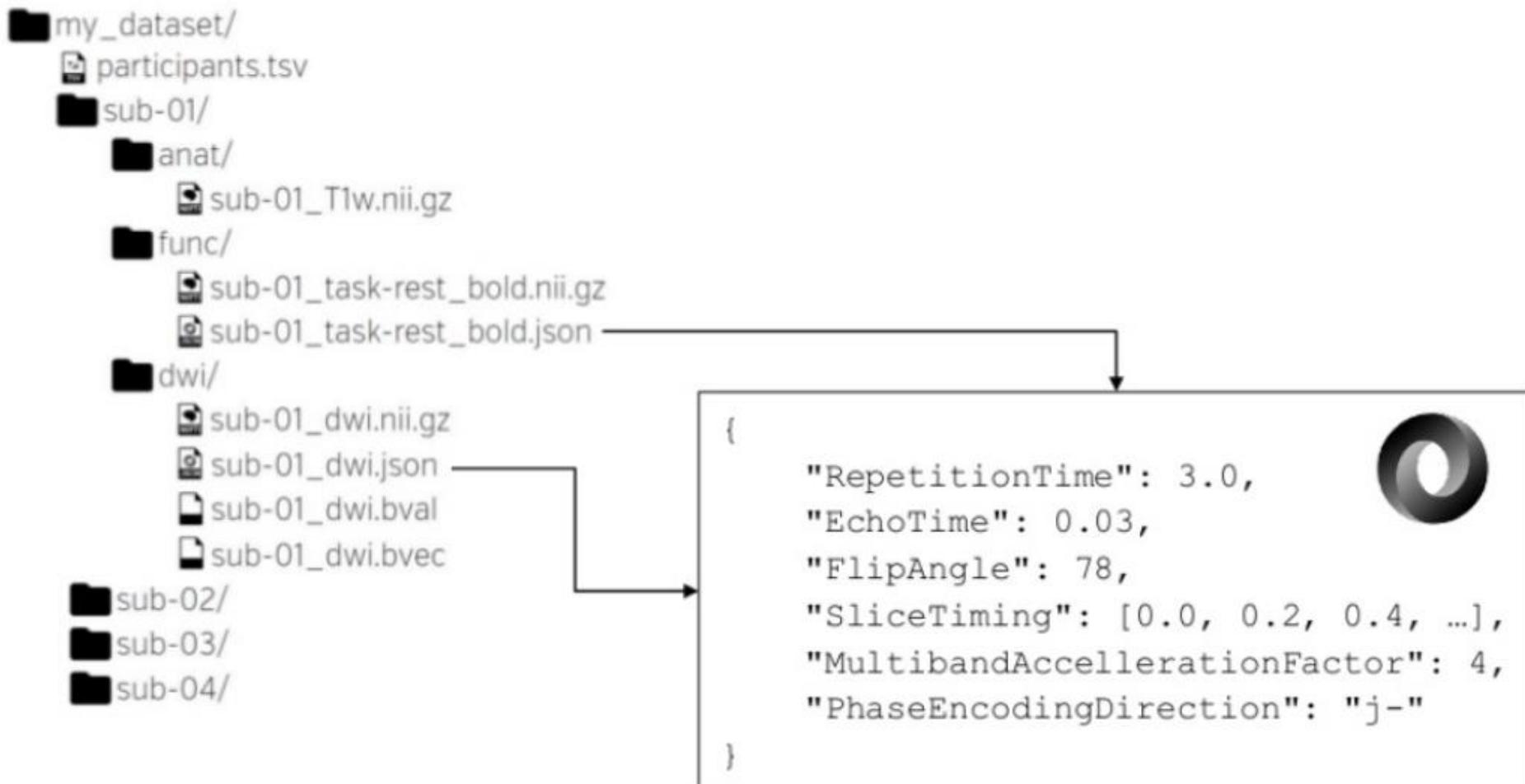
- Contains data files

```
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/
```



# BIDS

- Contains study specific JSON (metadata) files



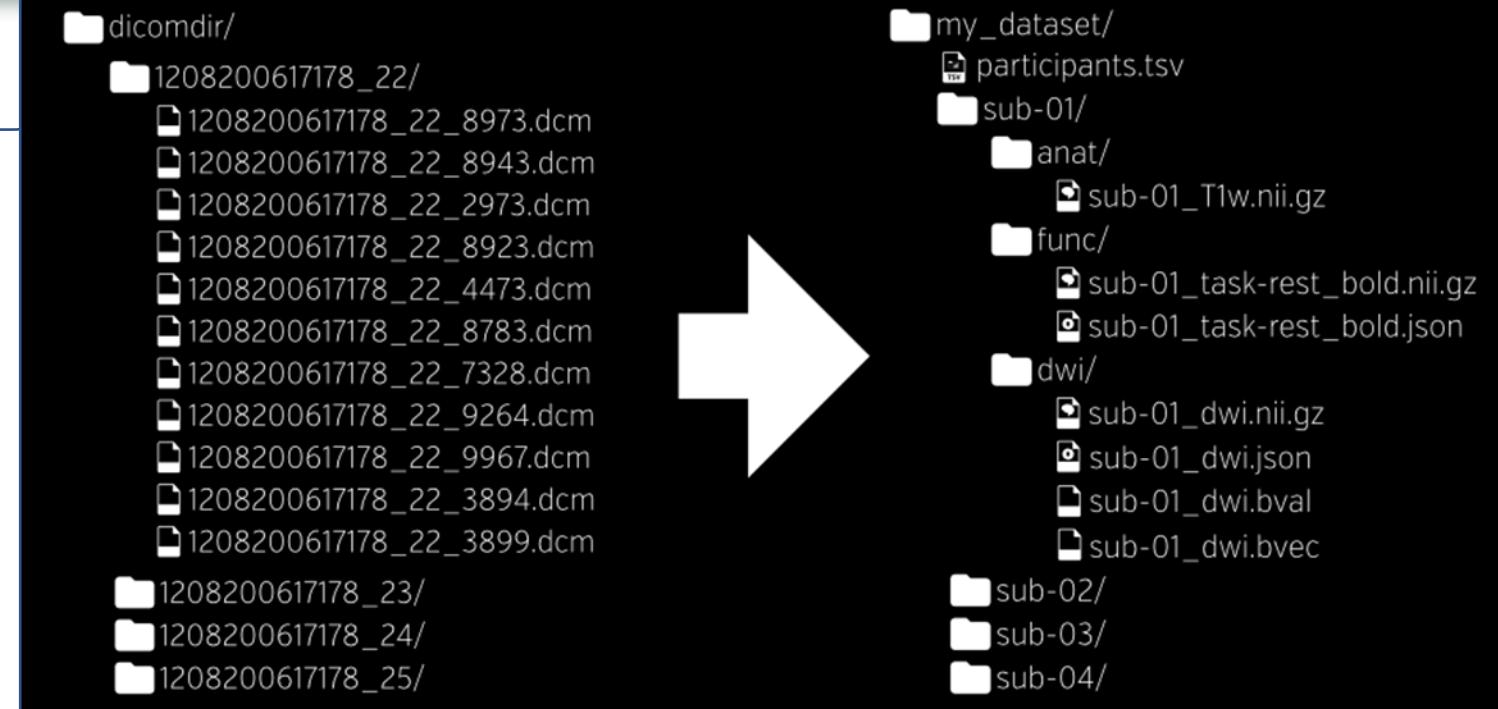
# DICOM → BIDS

Collect the data

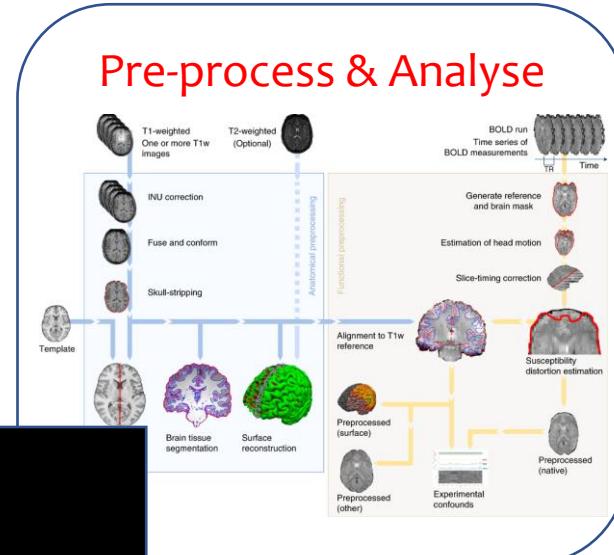


Many BIDS converters available

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



Pre-process & Analyse



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>

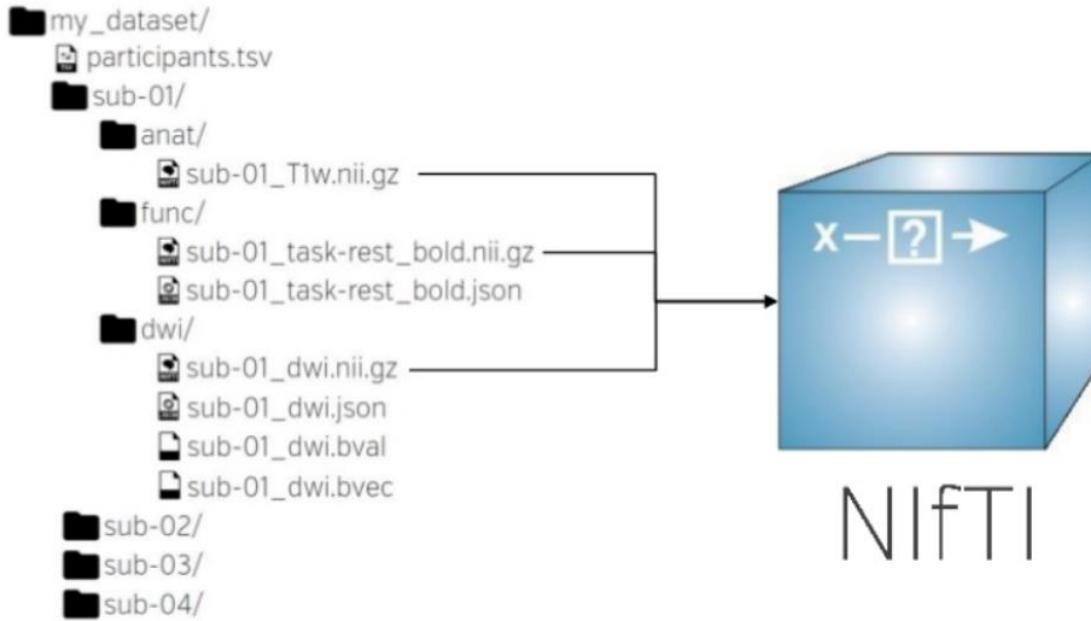
[https://github.com/MRC-CBU/BIDS\\_conversion/tree/main/MRI](https://github.com/MRC-CBU/BIDS_conversion/tree/main/MRI)

# Environment

Data  
Organise & Manage



Notebook: [nb01\\_Data-Organisation.ipynb](#)

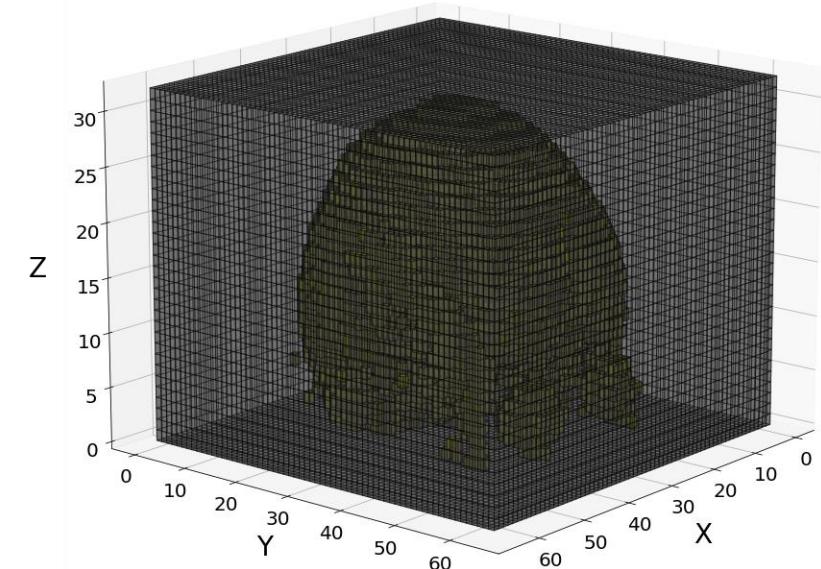
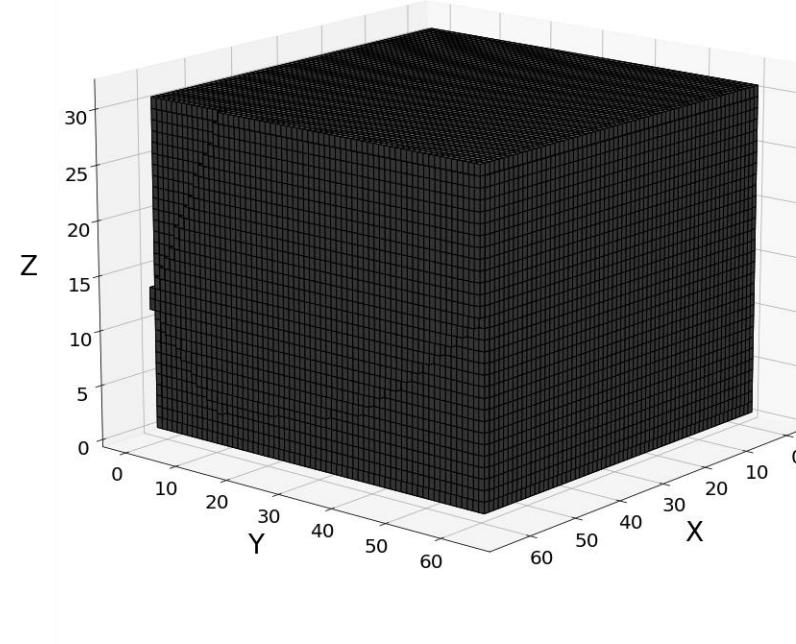


# Imaging data content

# MRI data structure

A 3D or 4D arrays of numbers

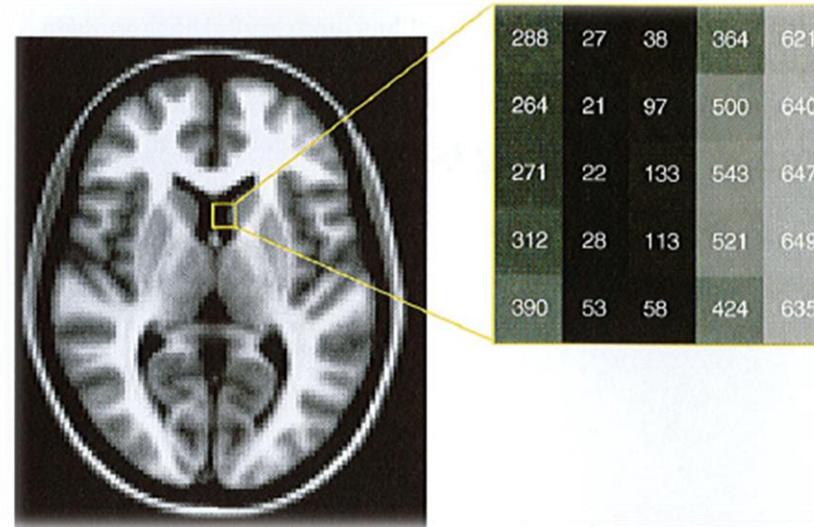
```
([[[ 0., 0., 0., ..., 0., 0., 0.],  
 [ 0., 0., 0., ..., 0., 0., 0.],  
 [ 0., 0., 0., ..., 0., 0., 0.],  
 ...,  
 [ 0., 0., 0., ..., 0., 0., 0.],  
 [ 0., 0., 0., ..., 0., 0., 0.],  
 [ 0., 0., 0., ..., 0., 0., 0.]],  
  
 [[ 0., 0., 0., ..., 0., 0., 0.],  
 [ 0., 25., 23., ..., 23., 32., 0.],  
 [ 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

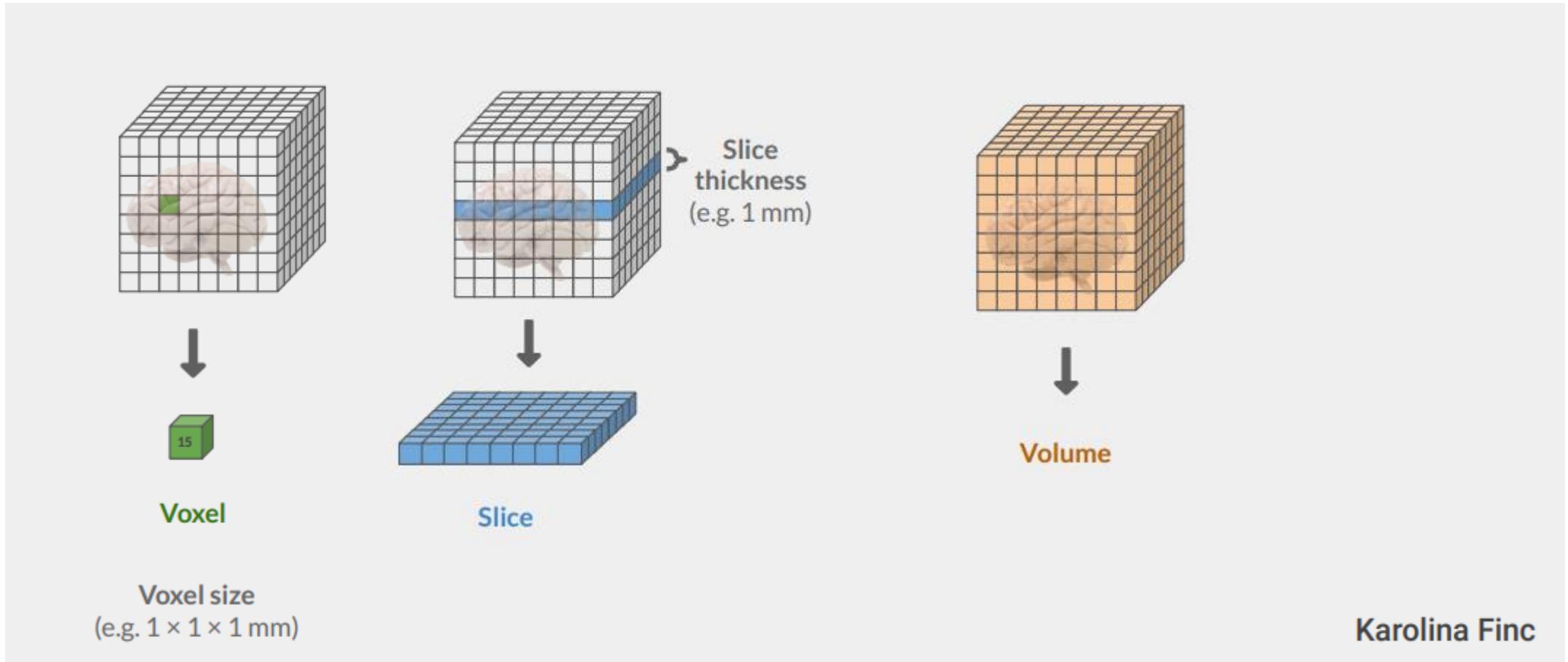
A 3D or 4D arrays of numbers – intensity values

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

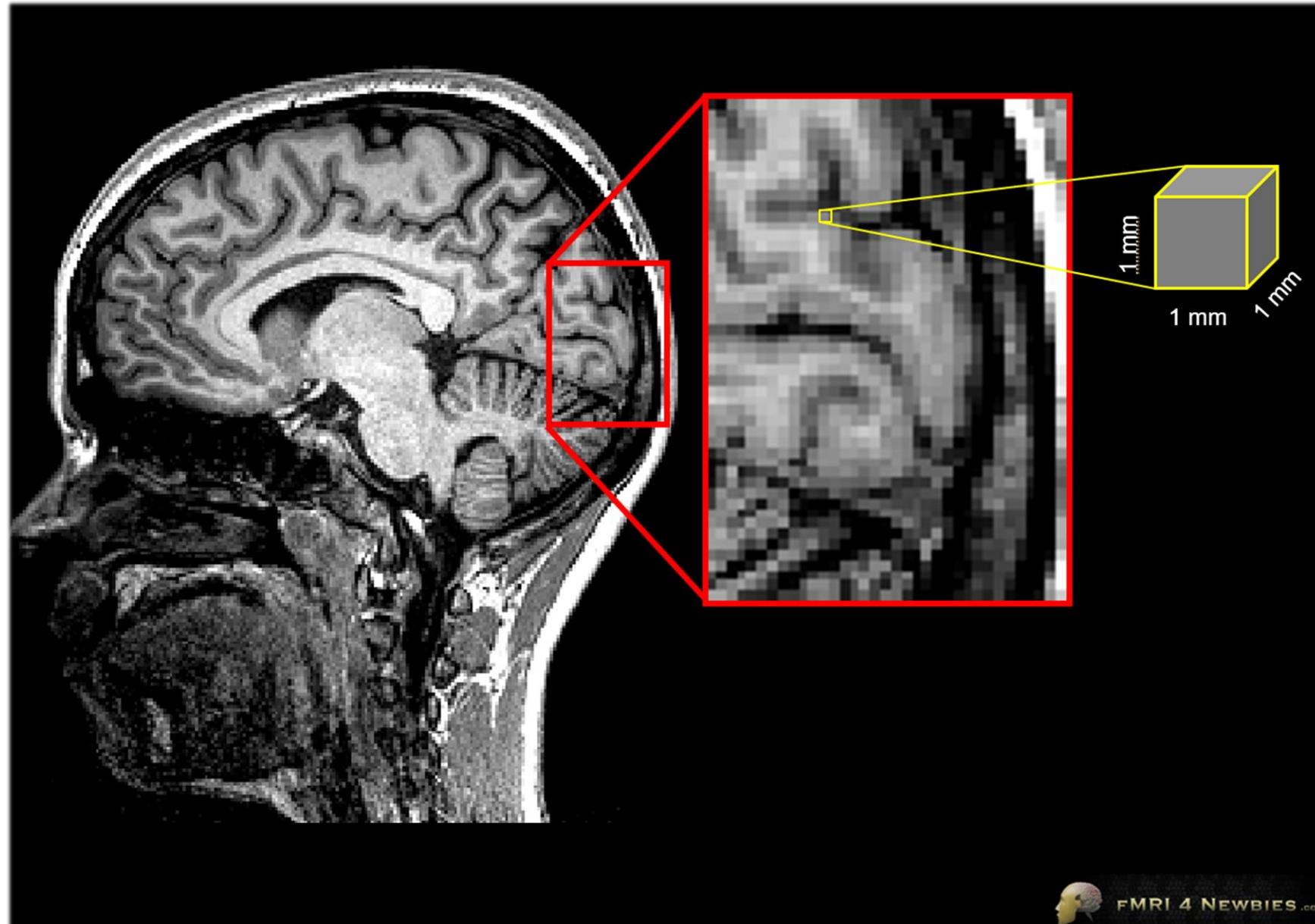


*Image from Poldrack et al., 2011*

# MRI data structure



# MRI data



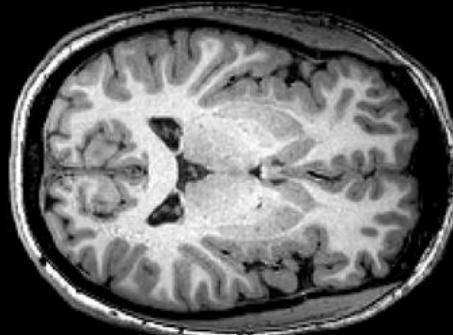
# MRI data

What determines the resolution?

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

high resolution

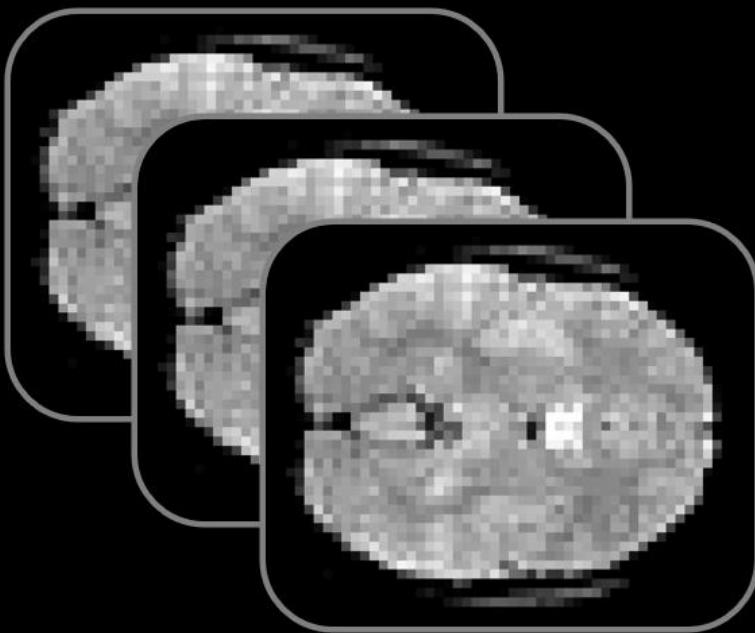
MRI



One 3D volume

fMRI

low resolution

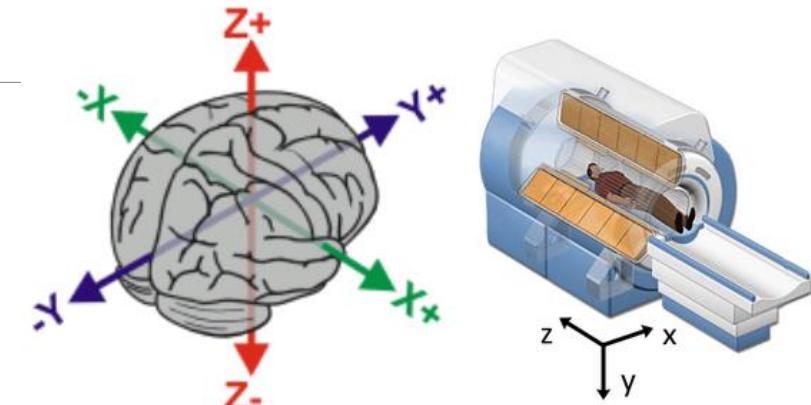
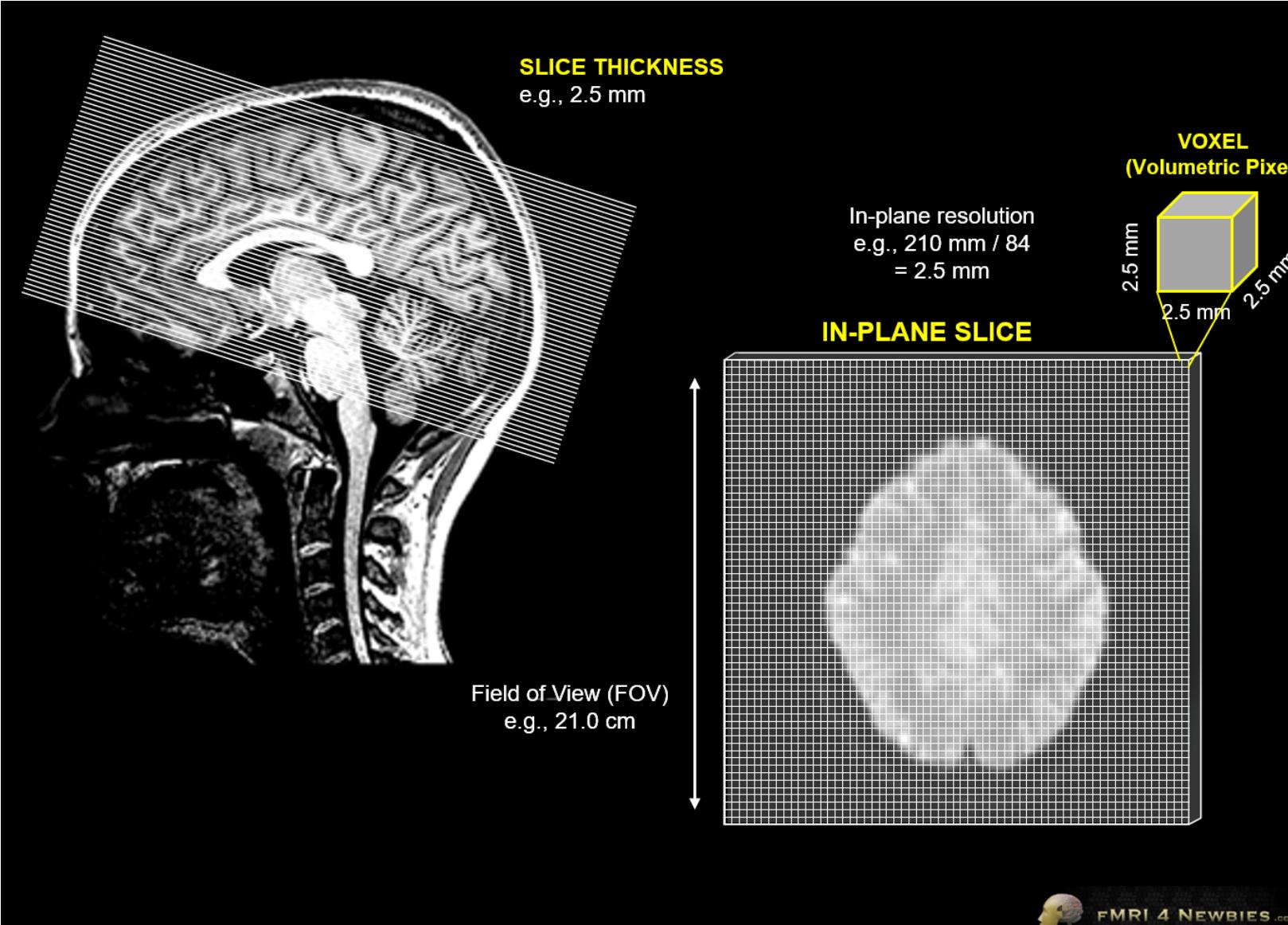


series of 3D volumes (i.e., 4D data)  
(e.g., every 2 sec for 5 mins)



# fMRI data

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

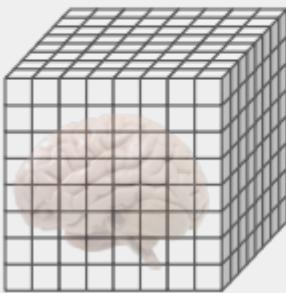


- Temporal resolution (TR), usually 1.5-3s
- Modern sequences allow acquiring multiple slices at the same time (multi-band)
- Typically, 30-50 slices acquired
- More slices = longer TR

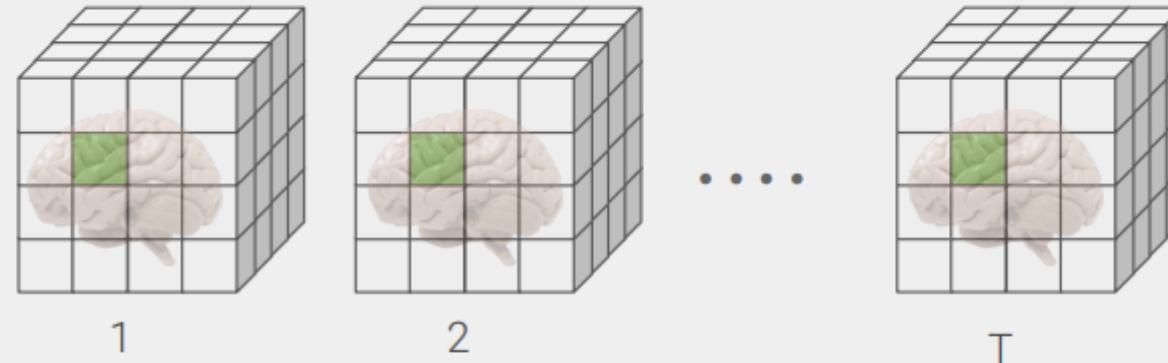


# MRI data structure

Structural data

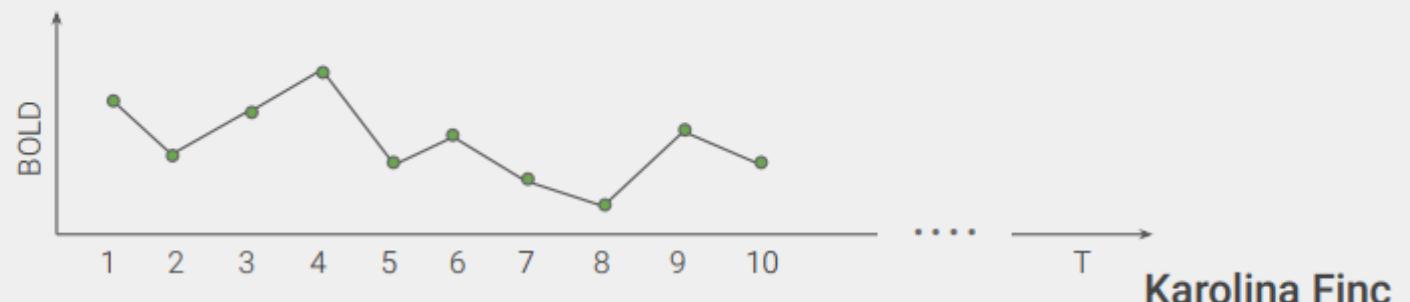


Functional data



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

Every voxel has its own time-series.



Karolina Finc

# 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