RaBIDS

**Rapid analysis pipeline using the Brain Imaging Data Structure**

Manual – v0.1

Christian Paret – 2020

RG Psychobiologie of Selfregulation

Department of Psychosomatic and Psychotherapeutic Medicine

Central Institute of Mental Health

[christian.paret@zi-mannheim.de](mailto:christian.paret@zi-mannheim.de)

# Introduction

Tested with Matlab R2019a and SPM12 v7219

## Vision:

* Meet high-level replicability standards with our research
* Provide tool for ZI researcher to efficiently analyze their data
* Establish common standard at PSM (and maybe ZI) for analysis, quality control and archiving of fMRI data

## Specific aims:

* Import fMRI-neuroimaging data to BIDS compatible format
* Provide generic pipeline for preprocessing and firstlevel analysis for most common data formats in our Department (Siemens MRI-scanner, SPM12-based matlab software, Presentation software)

# Before you start measuring

## Naming MRI sequences

An MRI measurement protocol can consist of several MRI series. For example, you may administer two functional tasks in one session, and the technical assistant (MTA) will start an echoplanar imaging (EPI) sequence for each task-run by selecting the corresponding tab at the MRI computer. It is recommended to assign unique names to each of these MRI series by changing the name of this tab. Ask the MTA to do this when you set up the sequence. You should also have information ready on the number of images per sequence.Name your MRI sequence and use this name to identify each MRI series

### Define subject ID to make it RaBIDS compatible

The subject ID is a series of characters or numbers used to identify a study participant. The subject ID + session ID allow unambiguous identification of a measurement. For example:

* Subject ID: sub02
* Session ID: ses-01

(The naming format of Subject and Session ID is arbitrary, as can be seen below)

🡪 Unique ID for identification of this measurement: sub-02ses-01

You need to use the unique ID to register the participant at the MRI scanner. The name of your Presentation logfiles needs to begin with the unique ID.

If you have multiple sessions you need to make sure that the session identifier comes as a suffix to the subject code as shown above (I call this the the “session suffix”). E.g. for subject named “sub-02”, session 1, following unique IDs are allowed:

* sub-02ses-01 (“ses-01” is the session suffix)
* sub-0201 (“01” is the session suffix)
* 0201 (“01” is the session suffix)
* Etc.

You should NOT put the session identifier anywhere else than at the end of the unique ID! That is, the following unique IDs CANNOT be processed with the RaBIDS software, because it will incorrectly interpret your input:

* ses01sub02
* 01sub02
* 0102
* Etc.

## Exporting data from the MRI scanner

The person operating the experiment and MRI machine should export data to hobbes (=data exchange server) in the following structure:

1. Presentation logfiles:
   * [Your project folder on hobbes]
     + sourcedata
       - [Subject ID] (e.g. “sub-02”)
         * [Session ID] (e.g. “ses-01”)

*Save logfiles here*

(e.g. “sub02ses01\_task-emo.log”)

1. MRI images:
   * [Your project folder on hobbes]
     + dicomdir
       - [Subject ID] (e.g. “sub-02”)
         * [Session ID] (e.g. “ses-01”)

*Save MRI images here*

The session-level can be skipped in case there is only one session per subject.

# Getting started

## Download additional software

### SPM

… unless you don’t have it yet (installed on flts02)

<https://www.fil.ion.ucl.ac.uk/spm/software/spm12/>

### dicm2ni

… unless you don’t have it yet (installed on flts02)

<https://github.com/xiangruili/dicm2nii>

* For automated production of .json files during dicom import: open dicm2nii GUI and select option ''save json file'. This is needed to receive a BIDS-compatible data set
* Note when downloading and using for first time Matlab will return an error, which is easy fixed (you can write me an email).

## Get the most recent RaBIDS package

* Go to directory [\\zifnas\psm\AG-Austausch\RaBIDS](file:///\\zifnas\psm\AG-Austausch\RaBIDS)
* Go to directory with most recent version (e.g. v0.1)
* Copy the “code”-directory to a directory you call “dataset” within your local project directory.

# Import a data set into BIDS format

## Sourcedata

Sourcedata is all data other than MRI data that is recording in the context of a measurement (Gorgolewski et al., 2016). You need to populate this folder according to the BIDS structure (adhere to the structure given for saving Presentations logfiles to your local project directory, given in 2.1.2). RaBIDS expects your Presentation logfiles to be located here.

## Dicom images

You have two options:

* + Import dicom MRI images directly from data exchange server “hobbes” without copy pasting in your local project directory. This works as long as you use the data structure suggested above for exporting data from the scanner site.
  + Create a directory called “dicomdir” in your project directory. Copy-paste all dicoms to this directory. You do not need to sort them into subdirectories.

## Fill the datasheet

Go to directory dataset\code\RaBIDS templates. Copy the file datasheet.xlsx to the directory dataset\code. The Excel-table has six colomns named “object type, description, user input, minimum images, maximum images, UserNotes”. Your input is required to the fields “user input, minimum images, maximum images”. Never change the “object type” fields. You may enter notes in the UserNotes colomn (ignored by the program).

Don’t forget to save your input everytime before running matlab (e.g. by pressing keys Ctrl+S)!

### Object type: add path

Absolute path to software used by RaBIDS (see 3.1).

### Object type: data\_analysis\_path

Parent directory of dataset-directory (probably your local project directory).

### Object type: data exchange path

Define path to the directory called “dicomdir” containing the dicom MRI images of your study on the data exchange server (=hobbes; see 2.2B).

### Object type: dicoms

Explain how your dicoms are saved: if they are all mixed in one directory enter “allinone”, if they were exported in a BIDS-inspired directory structure enter “BIDS”.

### Object type: subject info

Duplicate the subject info row as many times as you have subjects to analyse. You can also just analyze a subset of your sample. Input the exact name of each subject as it has been registrated in the scanner and in the Presentation logfile. Note that the file name of dicom images may differ, because lower case letters are ignored in Siemens-file names. You need to type the file name to the subject-info field exactly as you registered the subject, that is, including lower case letters (this information is still present in the meta-data and will be used by the RaBIDS program).

### Object type: MRI series

Assign TaskName to MRI series (“rest” for resting state scan, “anat” for anatomical scan, e.g. “emoreg” for emotion regulation task). You may consult the recent BIDS manual to find inspiration for an appropriate label.

For each MRI series you need to define “minimum images” and “maximum images”. The program only use a scan with number of images within the given range (Minimum images ≤ number of images in the scan ≤ maximum images). The program cannot deal with scans with more than one series with the exactly same name (series info, see 4.3.7) and overlapping ranges of number of volumes.

### Object type: series info

This is the name of the MRI series as it has been labeled at the MR computer.

Example: this is an example path to an anatomical sequence at our Siemens-scanner (PrismaFit): \\USER\psm\EFPNFB\Tag1\t1\_mpr\_ns\_sag\_pat2\_iso\_asy.

Highlighted in purple = project directory; yellow = series info.

### Object type: first image

Usually in fMRI BOLD, you start the experiment only after recording a few volumes (I usually start with the fifth scan) to exclude T1-effects in your T2\*-weighted images. If you write “yes” in the “user input” field and declare the first image (e.g. 5) in the “minimum images field”, the initial X-1 (e.g. 4) images will be deleted (they are not needed anymore). The program will check for the volume number in the name of the dicom image, so you don’t need to be afraid that more images than intended will be deleted if you need to run the program again on that subject later. Use this feature to avoid problems in the following automated processing steps.

### Object type: general suffix

The general suffix corresponds to the name of project at the MRI scanner. This object type needs to be set according to the name of the scan, see example below.

Note, that the general suffix can take value underscore (i.e., “\_”). For our Siemens-scanners at ZI, use the value between “MR.” and “.XXXX”, where XXXX=running number of sequence.

Example: this is an example path to an anatomical sequence at our Siemens-scanner (PrismaFit): \\USER\psm\EFPNFB\Tag1\t1\_mpr\_ns\_sag\_pat2\_iso\_asy.

Highlighted in purple = project directory; yellow = series info.

The export-dicom tool uses the first two sub-directories to USER to create a general suffix in the dicom name, i.e. “psm” and “efpnfb”. The first volume of the mprage scan is called:

EFP01.MR.PSM\_EFPNFB.0006.0001.2018.05.14.13.00.47.560079.855171.IMA

Highlighted text in cyan=subject name; purple = general suffix; text in yellow = series number of scan; green = running number of scan in this series.

### Object type: add prefix

For BIDS compatibility each subject name needs to begin with “sub-“. If you have not yet adopted this format, write “yes” in the “user input” field.

### Object type: session ID

RaBIDS assumes that subject IDs list a unique identifier for the subject (e.g. sub01) and are followed by a unique session identifier (e.g. ses02). See 2.1.1 for more information.

Each field “session ID” identifies a session (session = one series of MRI measurement typically comprising localizer and fieldmap scans, anatomical and functional scans). If your experiment has just one session per subject, you can leave the “user input” field empty. The “session info” row below the “session ID” row relates to this session and is explained below. If you have more than one session, you need to add as many “session ID” and “session info” rows as you have sessions per subject. Explain in the “user input” field how this particular session should be named. This input is used for naming the corresponding directory in your data structure. According to BIDS, it needs to begin with “ses-“, and can be followed by e.g. “01” to name the first session (i.e. “ses-01”, or e.g. “pre” for the pre-measurement session of your clinical intervention (i.e. “ses-pre”).

### Object type: session info

Optional/not needed: Input the session suffix of this particular session. This is usually not needed but may be helpful in some cases.

### Object type: write scan protocol

If user input is “yes”, a scan protocol listing series number, series name and number of volumes will be written to the dicom directory when importing dicoms.

### Object type: overwrite import

Write “yes” if you want to overwrite existing files in your data structure and “no” if you don’t want to do it. This is helpful e.g. if you have imported some subject previously for an interim analysis, and you have collected more data meanwhile to import. In this case you can have all subjects listed in your datasheet, and RaBIDS-Import will skip those that were already imported and only work with the new ones.

### Object types: [overwrite sots, overwrite preprocess, overwrite firstlevel]

Same as “overwrite import”, but for steps 2-4 (i.e. RaBIDS-create sots, RaBIDS-preprocess, RaBIDS-firstlevel)

## Import dataset

Run the Matlab program RaBIDS\_1\_Import.m, which you find in the code-directory.

## Define Stimulus Onset Times (SOTS)

To become BIDS compatible, your dataset needs to include “events.tsv”-files that list trial information of your experiment. Although we will not use the events.tsv file for the RaBIDS firstlevel analysis, this file is required for BIDS validation and can be used for other analysis packages (e.g. fmriPrep, (Esteban et al., 2019)).

“RaBIDS\_2\_Create\_SOTS.m” works for basic analysis of fMRI data, based on your input to the “conditions\_TaskName.xlsx”-file. It produces files called “multicond.mat”, which go straightforward to SPM12’s first-level analysis batches. An events.tsv-file is created in addition. Output will be saved to your project directory (e.g.\dataset\sub-02\ses-01\func). Two files generated per task:.

1. sub-02\_ses-01\_func\_task-TaskName\_multicond.mat
2. sub-02\_ses-01\_func\_task-TaskName\_events.tsv

Go to directory code/RaBIDS templates. Copy “conditions\_TaskName.xlsx” to the code-directory. Change the name of the Excel file according to your task:

* conditions\_TaskName.xlsx, where TaskName = the name you assigned previously in the datasheet (i.e. user input to the corresponding MRI series, see 4.4.4; e.g. “conditions\_emoreg.xlsx”).

### Condition

Add a row for each condition that goes into the model. Identify each condition by counting upwards. Other than in the datasheet.xlsx table you are required to edit the items in the first colomn of the condition.xlsx table. That is, the first condition is named “Condition1”, the second “Condition2”, and so on.

#### Name

Assign a name (arbitrary) that will be used to identify this condition in the SPM-model. White spaces not allowed.

#### OnsetID

This is the code to identify onset of the condition written to the Presentation logfile.

#### Duration

You may assign a fixed duration for trials of this condition type. Set duration = 0 for stimulus events that are to be modeled with a stick function (aka “onset transient”). If you don’t want to use the “Duration” feature, leave the user-input field empty. See OffsetID for more information.

#### OffsetID

This is the code to identify the offset of the condition in the Presentation logfile. For example, in case each trial is immediately followed by a “fixation cross” event denoted by code “fixcross” in the logfile, you may assign “fixcross” to OffsetID and leave the “Duration” field empty. If you want to use “Duration”, enter some arbitrary character string to user input (e.g. ‘none’), otherwise RaBIDS\_2\_Create\_SOTS.m will return an error.

### Log ID

Unique ID to identify the logfile corresponding to the task. Wildcards are allowed (i.e. use an asterisk to allow for undefined characters, e.g. \*emoreg\*). User input is required in the “Name” colomn.

### Logfile ID format

* ‘BIDS’: logfile name format is: “sub-SUBJECTNUMBER\_ses-SESSIONNUMBER\_ LogID.log” (see 2.2A)
* ‘free’: logfile name can have any format, as long as the LogID is included. This option will not double-check whether the logfiles have the right subject name. The program searches for files with format “LogID.log” in the corresponding session directory of that subject.

User input is required in the “Name” colomn. Logfiles are saved in the sourcedata directory.

### Contrast

Add a row for each contrast that should be estimated. Identify each contrast by counting upwards. Other than in the datasheet.xlsx table you are required to edit the items in the first colomn of the condition.xlsx table. That is, the first condition is named “Contrast1”, the second “Contrast2”, and so on.

#### Name

Assign name (arbitrary) to identify the contrast in the SPM-model.

#### ContrastType

Only t-contrast is implemented in version v0.1. Input is ignored.

#### ContrastPlus1

Enter the condition that you want to contrast against another condition (i.e. ContrastPlus1 > ContrastMinus1).

#### ContrastMinus1

This condition is contrasted against the condition entered in the ContrastPlus1 field. Leave empty to estimate ContrastPlus1>implicit baseline.

## Create SOTS file

Run matlab program RaBIDS\_2\_Create\_SOTS.m. You are requested to input the TaskName as a character string (i.e., within inverted commas; ‘TaskName’) in Matlab’s Command Window.

# Data analysis

Version v0.1 expects images from standard PSM EPI sequence with TR=2s and 36 slices. This will be more flexible in future. If you have different MRI parameters, you need to adapt RaBIDS functions. All routines have limitations and are not generic to all types of preprocessing pipelines and firstlevel models.

## Preprocessing

Run matlab program RaBIDS\_3\_Preprocess.m. You are requested to input the TaskName as a character string (i.e., within inverted commas; ‘TaskName’) in Matlab’s Command Window.

## Firstlevel

Run matlab program RaBIDS\_4\_Firstlevel.m. You are requested to input the TaskName as a character string (i.e., within inverted commas; ‘TaskName’) in Matlab’s Command Window.