Brain Image Data Structure (BIDS): what, why, and how?

Suzanne T. Witt, Ph.D.
3 October 2022

What is BIDS?

- BIDS is a system for naming and organizing ‘raw’ neuroimaging files
- BIDS is not a new file format

Stuff

What is BIDS?

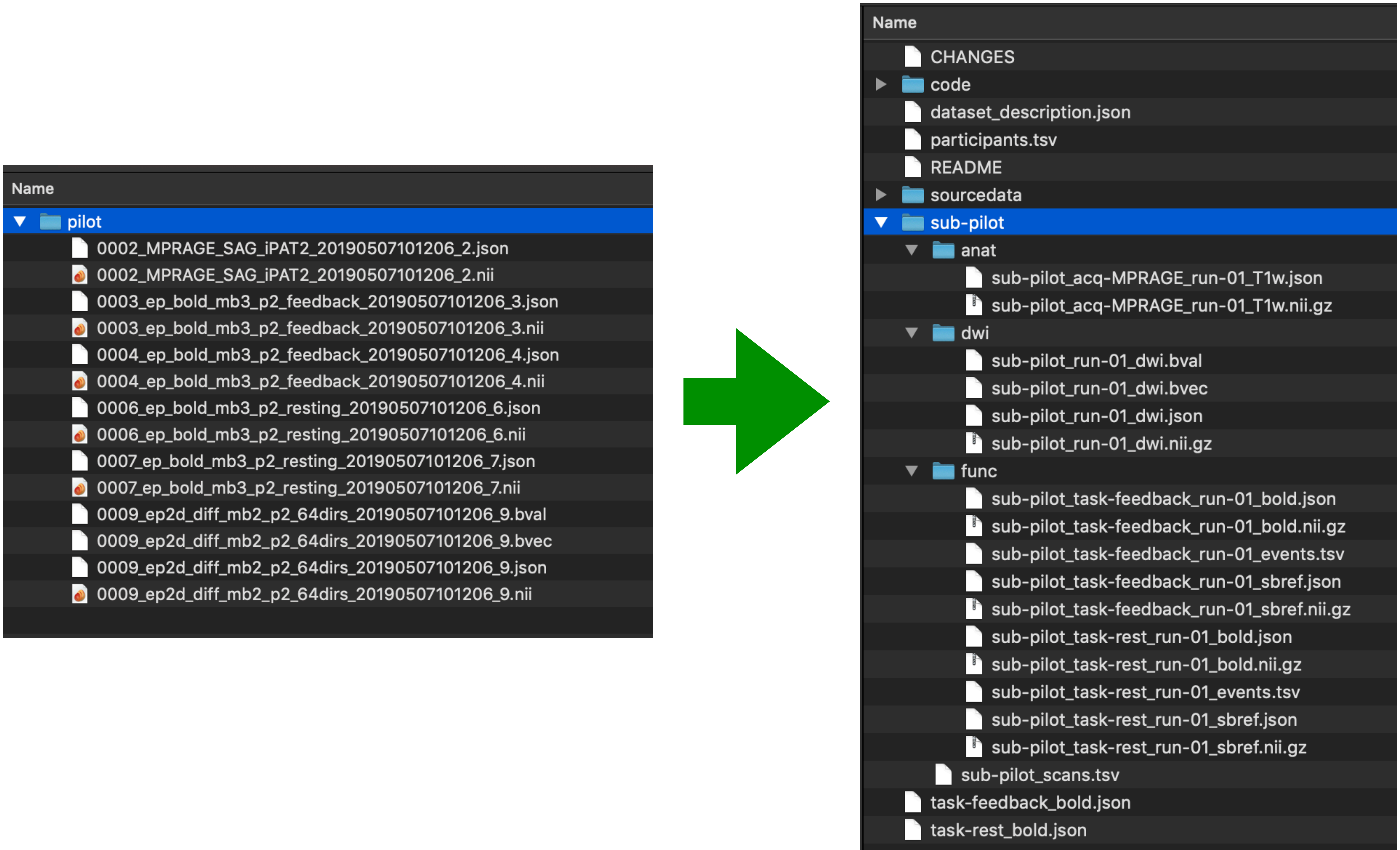
- BIDS is a system for naming and organizing ‘raw’ neuroimaging files
- BIDS is not a new file format

Animals

Household Items

Fruits and Vegetables

In practice for MRI data



EEG-BIDS

- Recommended to use European data format (.edf) or BrainVision data formats (.vhdr, .vmrk, .eeg)

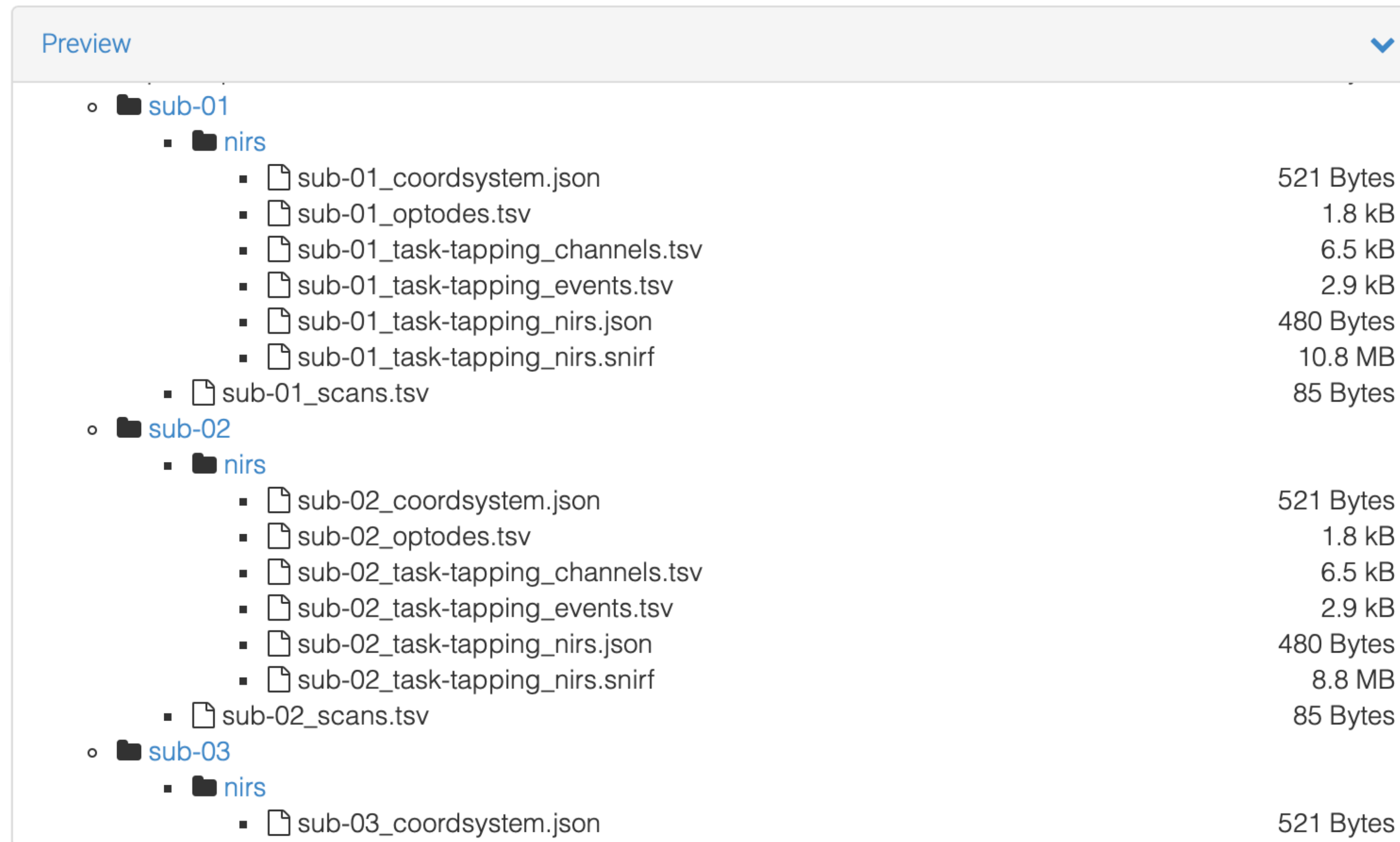
```
/
├── README
├── CHANGES
├── dataset_description.json
├── participants.tsv
├── participants.json
├── task-TASKNAME_events.json
├── stimuli
│   ├── stim1.png
│   └── ⋮
├── sourcedata
│   ├── sub-01
│   │   └── eeg
│   │       └── sub-01_task-TASKNAME_eeg.xdf
│   └── ⋮
├── sub-01
│   ├── anat
│   │   ├── sub-01_T1w.nii.gz
│   │   └── sub-01_T1w.json
│   └── eeg
│       ├── sub-01_task-TASKNAME_eeg.edf
│       ├── sub-01_task-TASKNAME_eeg.json
│       ├── sub-01_task-TASKNAME_events.tsv
│       ├── sub-01_task-TASKNAME_channels.tsv
│       ├── sub-01_task-TASKNAME_electrodes.tsv
│       └── sub-01_task-TASKNAME_coordsystem.json
└── ⋮
```

MEG-BIDS

- Unprocessed MEG data MUST be stored in the native file format of the MEG instrument used to collect the data

fNIRS-BIDS

- The Shared Infrared File Format (SNIRF) has been chosen as the “official” BIDS file format



The contents of a BIDS directory

General format for BIDS filenames

- MRI anatomy
  sub-<label>_ses-<label>_acq-<label>_run-<index>_T1w.nii.gz

- fMRI
  sub-<label>_ses-<label>_task-<label>_run-<index>_bold.nii.gz

- EEG
  sub-<label>_ses-<label>_task-<label>_run-<index>_eeg.edf

- MEG
  sub-<label>_ses-<label>_task-<label>_run-<index>_meg.ds

- NIRS
  sub-<label>_ses-<label>_task-<label>_run-<index>_nirs.snirf

Why should I adopt the BIDS standard?

- BIDS directory contains all necessary information about a study and its data in one place
- Make use of a growing number of BIDS apps, such as fmriprep, HCP pipelines, C-PAC, etc.
- Required format for many data repositories
- Warm, fuzzy feeling from actively supporting the open neuroscience/reproducibility movement

How do I get my data into BIDS format?

- Growing number of tools available to convert MRI, EEG, and MEG data into BIDS format
- Most popular for MRI/fMRI is HeuDiConv
- MNE-BIDS is a free option for converting both EEG and MEG data

- AFNI BIDS-tools
- BIDS2ISATab
- BIDSto3col
- BIDS2NDA
- bidskit
- Data2Bids
- Dcm2Bids
- DCM2NIIx
- DICM2NII
- HeuDiConv
- OpenfMRI2BIDS
- ReproIn (HeuDiConv-based turnkey solution)
- bids2xar (for XNAT import)
- XNAT2BIDS
- Horos (Osirix) export plugin
- BIDS2NIDM
- BIDScoin
- MNE-BIDS (MEG/EEG/iEEG)
- EEGLAB with plugin (MEG/EEG analysis package)

Special case: autobids

What is autobids?

- autobids is a tool developed by the Khan Lab to automatically download DICOM files from the CFMM server and convert them to nifti format using BIDS standard filenames

Who can make use of autobids?

- Any researcher collecting MRI data using either the CFMM 3T or 7T scanner

How it works

1. You register your study for autobids conversion

2. We create the necessary configuration and heuristic files to ensure your scans are correctly converted into BIDS format

3. You follow instructions for setting up Globus

4. You receive an email with a summary of the BIDS validator, indicating that your BIDS-ed data is ready to be downloaded

What you need to do

1. Register your study using the new autobids website
2. Add ‘bidsDump’ is an authorized user for your PI’s scans on the CFMM DICOM server, if not already
3. Decide on a single scheme for naming your subjects and scans on the CFMM DICOM server

Your responsibilities as an autobids user

1. Adopt a single, consistent naming scheme for subject IDs, session IDs, task names, etc.

2. Use the following pattern for patient IDs in the DICOM server: <date>_<otherInfo>_<subjectID>_<sessionID>

   20190923_Study3T_s101_001

3. Include task information in fMRI series descriptions:

   ep_bold_mb3_p2_FingerTapping

4. Let me know about any changes to the protocol

Further information

- BIDS documentation: https://bids.neuroimaging.io/
- Heudiconv tutorial: http://reproducibility.stanford.edu/bids-tutorial-series-part-2a/
- Online BIDS validator: https://bids-standard.github.io/bids-validator/
- MNE BIDS tool: https://mne.tools/mne-bids/index.html
- BrainsCAN Neuroimaging wiki autobids page: https://osf.io/k89fh/wiki/autobids/?view_only=8281493cc75d429285735c98c1267261
- BIDS-MEG: https://www.nature.com/articles/sdata2018110
- BIDS-EEG: https://www.nature.com/articles/s41597-019-0104-8
- BIDS: https://www.nature.com/articles/sdata201644

HeuDiConv

General strategy

- ‘Look-up table’ matching DICOM sequences to BIDS filenames

- ‘Look-up table’ is user-specified

- Only DICOM sequence-BIDS filename pairs in the ‘look-up table’ will be converted dicominfo

example_dcm_file	series_id	dcm_dir_name	unspecified2	unspecified3	dim1	dim2	dim3	dim4	TR	TE	protocol_name
example_dcm_file	series_id	dcm_dir_name	unspecified2	unspecified3	dim1	dim2	dim3	dim4	TR	TE	protocol_name
2019_05_07_PILOT.MR.BODELL^PEERSTUDY.00	1-localizer	1	-	-	384	384	15	1	0.0086	4.0	localizer
2019_05_07_PILOT.MR.BODELL^PEERSTUDY.00	2-MPRAGE SAG iPAT2	2	-	-	256	256	192	1	2.3	2.98	MPRAGE SAG iPAT2
2019_05_07_PILOT.MR.BODELL^PEERSTUDY.00	3-ep_bold_mb3_p2_feedback	3	-	-	84	84	48	1	1.0	30.0	ep_bold_mb3_p2_feedback
2019_05_07_PILOT.MR.BODELL^PEERSTUDY.00	4-ep_bold_mb3_p2_feedback	4	-	-	84	84	48	766	1.0	30.0	ep_bold_mb3_p2_feedback
2019_05_07_PILOT.MR.BODELL^PEERSTUDY.00	6-ep_bold_mb3_p2_resting	6	-	-	84	84	48	1	1.0	30.0	ep_bold_mb3_p2_resting
2019_05_07_PILOT.MR.BODELL^PEERSTUDY.00	7-ep_bold_mb3_p2_resting	7	-	-	84	84	48	360	1.0	30.0	ep_bold_mb3_p2_resting
2019_05_07_PILOT.MR.BODELL^PEERSTUDY.00	9-ep2d_diff_mb2_p2_64dirs	9	-	-	98	98	78	68	3.5	55.0	ep2d_diff_mb2_p2_64dirs
2019_05_07_PILOT.MR.BODELL^PEERSTUDY.00	10-ep2d_diff_mb2_p2_64dirs	10	-	-	98	98	156	1	3.5	55.0	ep2d_diff_mb2_p2_64dirs
2019_05_07_PILOT.MR.BODELL^PEERSTUDY.00	12-ep2d_diff_mb2_p2_64dirs_PA	12	-	-	98	98	78	1	3.5	55.0	ep2d_diff_mb2_p2_64dirs_PA

. . .

dicominfo-1

patient_id	study_description	referring_physician_name	series_description	sequence_name	image_type
patient_id	study_description	referring_physician_name	series_description	sequence_name	image_type
2019_05_07_pilot	Bodell^PeerStudy		localizer	*fl2d1	('ORIGINAL', 'PRIMARY', 'M', 'ND', 'NORM')
2019_05_07_pilot	Bodell^PeerStudy		MPRAGE SAG iPAT2	*tfl3d1_16ns	('ORIGINAL', 'PRIMARY', 'M', 'ND', 'NORM')
2019_05_07_pilot	Bodell^PeerStudy		ep_bold_mb3_p2_feedback_SBRef	epfid2d1_84	('ORIGINAL', 'PRIMARY', 'M', 'ND', 'NORM', 'MOSAIC')
2019_05_07_pilot	Bodell^PeerStudy		ep_bold_mb3_p2_feedback	epfid2d1_84	('ORIGINAL', 'PRIMARY', 'M', 'MB', 'ND', 'NORM', 'MOSAIC')
2019_05_07_pilot	Bodell^PeerStudy		ep_bold_mb3_p2_resting_SBRef	epfid2d1_84	('ORIGINAL', 'PRIMARY', 'M', 'ND', 'NORM', 'MOSAIC')
2019_05_07_pilot	Bodell^PeerStudy		ep_bold_mb3_p2_resting	epfid2d1_84	('ORIGINAL', 'PRIMARY', 'M', 'MB', 'ND', 'NORM', 'MOSAIC')
2019_05_07_pilot	Bodell^PeerStudy		ep2d_diff_mb2_p2_64dirs	ep_b995#55	('ORIGINAL', 'PRIMARY', 'DIFFUSION', 'NONE', 'MB', 'ND', 'NORM', 'MOSAIC')
2019_05_07_pilot	Bodell^PeerStudy		ep2d_diff_mb2_p2_64dirs_TRACEW	ep_b1000t	('DERIVED', 'PRIMARY', 'DIFFUSION', 'TRACEW', 'MB', 'ND', 'NORM')
2019_05_07_pilot	Bodell^PeerStudy		ep2d_diff_mb2_p2_64dirs_PA	ep_b0	('ORIGINAL', 'PRIMARY', 'DIFFUSION', 'NONE', 'MB', 'ND', 'NORM')

```python
import os
import numpy
from cfmm_base import infotodict as cfmminfodict
from cfmm_base import create_key

def infotodict(seqinfo):
    """Heuristic evaluator for determining which runs belong where
    allowed template fields - follow python string module:
    item: index within category
    subject: participant id
    seqitem: run number during scanning
    subindex: sub index within group
    """

    # call cfmm for general labelling and get dictionary
    info = cfmminfodict(seqinfo)

    rest = create_key('{bids_subject_session_dir}/func/{
        bids_subject_session_prefix}_task-rest_run-{item:02d}_bold')
    rest_sbref = create_key('{bids_subject_session_dir}/func/{
        bids_subject_session_prefix}_task-rest_run-{item:02d}_sbref')

    feedback = create_key('{bids_subject_session_dir}/func/{
        bids_subject_session_prefix}_task-feedback_run-{item:02d}_bold')
    feedback_sbref = create_key('{bids_subject_session_dir}/func/{
        bids_subject_session_prefix}_task-feedback_run-{item:02d}_sbref')

    info[rest]=[]
    info[rest_sbref]=[]
    info[feedback]=[]
    info[feedback_sbref]=[]

    for idx, s in enumerate(seqinfo):

        if ('resting' in (s.series_description).strip()):
            if (s.dim4==1 and  'SBRef' in (s.series_description).strip()):
                    info[rest_sbref].append({'item': s.series_id})
            elif (s.dim4==360):
                    info[rest].append({'item': s.series_id})

        elif ('feedback' in (s.series_description).strip()):
            if (s.dim4==1 and  'SBRef' in (s.series_description).strip()):
                    info[feedback_sbref].append({'item': s.series_id})
            elif (s.dim4>700):
                    info[feedback].append({'item': s.series_id})

    return info
```