

An Introduction to BIDS: Data Organization for Transparency and Reproducibility in Neuroimaging Research

Amanda Bischoff-Grethe

14 November 2018

Motivation

- Imaging experiments are complicated and can be organized in many different ways
- Many files/file types per subject
- Members of the same lab may use different ways to arrange data
- Difficulty sharing data across large scale projects



"Think this is bad? You should see the inside of my head."

Benefits of Brain Imaging Data Structure (BIDS)

- Standardized organization and format makes it easy to share data within/across labs
- Focus is on **raw data** (minimally processed), but source (e.g., DICOM) and derived data (e.g., a contrast map, post processed images) can be included as well
- Validation tools are available to check data integrity
- There is movement towards BIDS-ifying processed data as well

BIDS Community



[Home](#) [Features](#) [Get Started](#) [FAQ](#) [About LORIS](#) [Contact](#)

LORIS

(Longitudinal Online Research and Imaging System) is a web-based **data** and **project management** software for **neuroimaging research** studies. It is an **OPEN SOURCE** framework for storing and processing **behavioural, clinical, neuroimaging and genetic data**. LORIS also makes it **easy to manage** large datasets acquired over time in a longitudinal study, or at different locations in a large **multi-site** study.

NIMH Data Archive

Home Query Harmonization Tools Cloud Contribute Request Access Policy Tutorials About FAQ Tools login

Quick Navigation

- Query
 - Data from Labs
 - Data from Papers
 - By Measure/Element
 - By Concept
 - By GUID
- Share
 - Harmonization Standards
 - Prepare and Submit
 - Compute
 - Resources
 - GUID Tool
 - HTML5 Validation Tool **Beta**
 - Validation and Upload Tool
 - Download Manager
 - Data Dictionary
 - Contact Us
 - Request Account

NDAR News

- NDAR Account Dashboard and NIMH Data Archive in Nature
- NIAAA Releases Request for New NDA Studies Link Public
- List All News

Learn About NDAR

- For Participants
- For Scientists
- About NDAR
- Join Mailing List

The **NIMH Data Archive (NDA)** is an infrastructure for sharing de-identified human subjects data with qualified researchers. Data have been collected by hundreds of research projects across many scientific domains and harmonized to a common standard. The NDA also provides tools, methods, and analysis capabilities enabling collaborative science and discovery. Summary data is available to all.

The primary point of entry to the NDA is currently through the National Database for Autism Research (NDAR) website, which serves the autism research community. All NDA repositories can be accessed through this website for data contribution and querying.

Search NIMH Data Archive

OpenNEURO BETA

A free and open platform for analyzing and sharing neuroimaging data

Sign in with Google

Browse Public Datasets

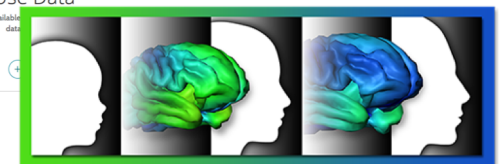
Get Data
Browse and download datasets from contributors all over the world.

Share Data
Upload your data and collaborate with your colleagues or share it with users around the world.

Use Data
Use our available data.

COINS

Collaborative Informatics and Neuroimaging Suite



Adolescent Brain Cognitive Development®
Teen Brains. Today's Science. Brighter Future.

Software Supporting BIDS

- FMRIPREP
- Automatic Analysis (an fMRI processing toolbox)
- Brainstorm (MEG/EEG)
- BIDS-Apps (<http://bids-apps.neuroimaging.io/apps/>)
 - FreeSurfer, afni_proc, tracula, HCPPIipelines, SPM, others

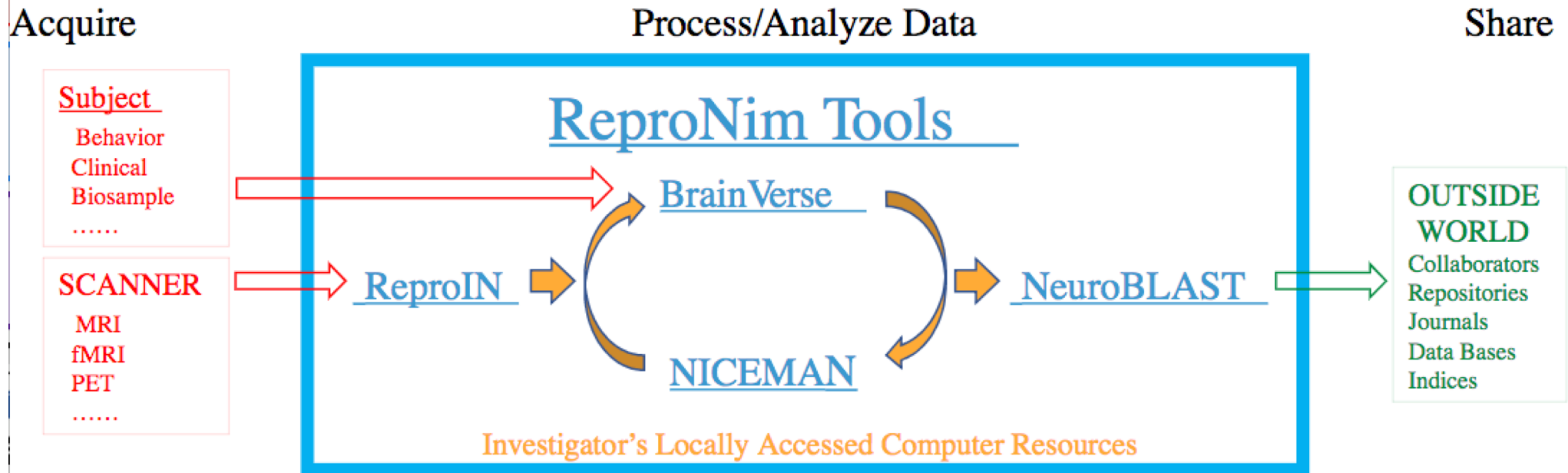
Available BIDS Apps

BIDS-Apps/example	version 0.0.7	open bug issues 0	build passing
BIDS-Apps/freesurfer	version	open bug issues 0	build passing
BIDS-Apps/ndmg	version v0.1.0	open bug issues 0	build project not found
BIDS-Apps/BROCCOLI	version v1.0.1	open bug issues 1	build passing
BIDS-Apps/FibreDensityAndCrosssection	version v0.0.1	open bug issues 0	build passing
BIDS-Apps/SPM	version v0.0.14	open bug issues 1	build project not found
poldracklab/mriqc	version 0.14.2	open bug issues 14	build passing
BIDS-Apps/QAP	Image not found	open bug issues 0	build passing
BIDS-Apps/CPAC	version v1.0.1a_22	open bug issues 0	build passing
BIDS-Apps/hyperalignment	Image not found	open bug issues 0	build passing
BIDS-Apps/mindboggle	version 0.0.4-1	open bug issues 2	build passing
BIDS-Apps/MRtrix3_connectome	version 0.3.0	open bug issues 0	build passing
BIDS-Apps/rs_signalExtract	version 0.1	open bug issues 0	build passing
BIDS-Apps/aa	version v0.1.0	open bug issues 0	build project not found
BIDS-Apps/niak	version latest	open bug issues 1	build passing
BIDS-Apps/oppni	version v0.7.0-1	open bug issues 1	build passing
poldracklab/fmriprep	version 1.2.1	open bug issues 22	build passing
BIDS-Apps/brainiak-srm	version latest	open bug issues 0	build failed
BIDS-Apps/nipypelines	version 0.3.0	open bug issues 0	build passing
BIDS-Apps/HCPPIipelines	version v3.17.0-18	open bug issues 0	build project not found
BIDS-Apps/MAGeTbrain	Image not found	open bug issues 0	build passing
BIDS-Apps/tracula	version v6.0.0-4	open bug issues 0	build passing
BIDS-Apps/baracus	Image not found	open bug issues 0	build passing
BIDS-Apps/antsCorticalThickness	version v2.2.0-1	open bug issues 0	build passing
BIDS-Apps/DPARSF	version v4.3.12	open bug issues 0	build passing
BIDS-Apps/afni_proc	Image not found	open bug issues 0	build passing
BIDS-Apps/rsHRF	version 1.0.0	open bug issues 0	build invalid response data

Transparency and Reproducibility



Data Acquisition/Lab-Centered Experiment Flow



BIDS File Formats

- Imaging
 - NIfTI format (.nii.gz)
 - Converted using dcm2nii or dcm2niix
- Tabular files
 - Tab delimited values (.tsv)
- Metadata files
 - Sidecar (JSON)
 - Provides additional information about the imaging file (e.g., dataset name, task description, scanner used)
 - Data dictionary for tabular files (e.g., column name, description, possible values)
 - Key-value pairs

Directory structure

```
<studyDir>/  
  sub-<participant_label>[/sess-<session-label>]  
    /<data_type>/  
      [can be one of func, dwi, anat, fmap]  
    [/code/]  
    [/derivatives/]  
    [/stimuli/]  
    [/sourcedata/]  
    [participants.tsv]  
    dataset_description.json  
    README  
    CHANGES
```


Imaging Data

```
sub-<participant_label>/[ses-<session_label>]
  anat/
    sub-<participant_label>[_ses-<session_label>][_acq-<label>][_ce-<label>]
    [_rec-<label>][_run-<index>][_mod-<label>]_<modality_label>.nii.gz]
  func/
    sub-<participant_label>[_ses-<session_label>]_task-<task_label>
    [_acq-<label>][_rec-<label>][_run-<index>][_echo-<index>]_bold.nii.gz]
  dwi/
    sub-<participant_label>[_ses-<session_label>][_acq-<label>]
    [_run-<index>]_dwi.nii.gz]
```

Creating BIDS compliant from DICOM

```
dcm2niix -o ${studyDir}/sub-${subj}/anat -z n -f sub-  
${subj}_T1w ${dcmdir}/${fullID}/${examNum}/${sDir}  
gzip *.nii
```

Outputs:

```
sub-6185_T1w.nii  
sub-6185_T1w.json
```

- Creates both nii and json files
- dcm2niix defaults to 'pigz' for gzip compression
- Currently I'm using our gzip (see -z n), may switch to pigz in future
- NOTE: dcm2niix_afni is defaulted to -z n, but dcm2niix is defaulted to -z y

Example T1w JSON for Temperament study

```
[agrethe@induction anat]$ more sub-6185_T1w.json
```

```
{  
  "Manufacturer": "GE",  
  "ManufacturersModelName": "DISCOVERY_MR750",  
  "ScanningSequence": "RM",  
  "SequenceVariant": "NONE",  
  "SeriesDescription": "MPRAGE_PROMO_HCP_.8",  
  "ProtocolName": "MPRAGE_PROMO_HCP_.8",  
  "ImageType": [  
    "ORIGINAL",  
    "PRIMARY",  
    "OTHER"  
  ],  
  "AcquisitionDateTime": "2018-02-14T08:54:36.000000",  
  "MagneticFieldStrength": 3,  
  "FlipAngle": 8,  
  "EchoTime": 0.003656,  
  "RepetitionTime": 0.008776,  
  "InversionTime": 1.06,  
  "ConversionSoftware": "dcm2niix",  
  "ConversionSoftwareVersion": "v1.0.20170411 GCC4.3.2",  
  "ReceiveCoilName": "Nova32ch",  
  "SoftwareVersions": "DV25.0_R02_1549.b"  
}
```

Task, diffusion imaging data

- Single band
 - dcm2niix to create NIfTI, JSON files
- Multiband
 - dcm2niix to jumpstart creating the JSON file, since it will at least populate basic fields of interest correctly
 - Use NIfTI file created by center

Creating/Modifying JSON files

```
jo -p "Name"="Temperament_fMRI - Multiband Imaging Dataset"  
    "BIDSVersion"="1.0.2" >>  
    Temperament_fMRI/dataset_description.json
```

```
jq '. |= . +  
{ "ReceiveCoilName": "'Nova32ch'", "SoftwareVersions": "'1.0.2'" }'  
sub-6185_T1w.json > tmp.json  
mv tmp.json sub-6185_T1w.json
```

- <https://github.com/jpmens/jo>
- <https://stedolan.github.io/jq/>

Checking compliance

- <https://github.com/INCF/bids-validator>
- bids-validator checks whether the data are compliant with the BIDS format
- Available on fmri2
- Can also try <http://incf.github.io/bids-validator>
 - Note: only works with chrome, firefox

```
bids-validator /data/EDresearch/Temperament_fmRI
```

Keck: cfmri2bids.m



```
[dshin@fmrcomet ~/data] bids-validator 170808_bids_bids/
1: You should define 'PhaseEncodingDirection' for this file. If you don't provide this information field map correction will not be possible. (code: 7 - PHASE_ENCODING_DIRECTION_NOT_DEFINED)
/sub-01/dwi/sub-01_dwi.nii.gz
/sub-01/func/sub-01_task-restingstate_bold.nii.gz
/sub-01/func/sub-01_task-stroop_run-01_bold.nii.gz
/sub-01/func/sub-01_task-stroop_run-02_bold.nii.gz

2: You should define 'EffectiveEchoSpacing' for this file. If you don't provide this information field map correction will not be possible. (code: 8 - EFFECTIVE_ECHO_SPACING_NOT_DEFINED)
/sub-01/dwi/sub-01_dwi.nii.gz
/sub-01/func/sub-01_task-restingstate_bold.nii.gz
/sub-01/func/sub-01_task-stroop_run-01_bold.nii.gz
/sub-01/func/sub-01_task-stroop_run-02_bold.nii.gz

3: You should define 'TotalReadoutTime' for this file. If you don't provide this information field map correction using TOPUP might not be possible. (code: 9 - TOTAL_READOUT_TIME_NOT_DEFINED)
/sub-01/dwi/sub-01_dwi.nii.gz

4: Task scans should have a corresponding events.tsv file. If this is a resting state scan you can ignore this warning or rename the task to include the word "rest". (code: 25 - EVENTS_TSV_MISSING)
/sub-01/func/sub-01_task-stroop_run-01_bold.nii.gz
/sub-01/func/sub-01_task-stroop_run-02_bold.nii.gz

5: Not all subjects/sessions/runs have the same scanning parameters. (code: 39 - INCONSISTENT_PARAMETERS)
/sub-01/func/sub-01_dir-fwd_run-01_epi.nii.gz
/sub-01/func/sub-01_dir-fwd_run-02_epi.nii.gz
/sub-01/func/sub-01_dir-rvs_run-02_epi.nii.gz
/sub-01/func/sub-01_dir-rvs_run-03_epi.nii.gz
/sub-01/func/sub-01_task-stroop_run-02_bold.nii.gz

Summary:
27 Files, 189.72MB
1 - Subject
1 - Session

Available Tasks:
restingstate
stroop

Available Modalities:
T1w
T2w
dwi
bold
fieldmap
```

```
function cfmri2bids(dirpath, skipmodal)
% Converts 3T scan data directory to BIDS Format
% David Shin, 2017.08.01

[dshin@fmrcomet ~/data]$ tree -L 3 170808_bids_bids/
170808_bids_bids/
|-- dataset_description.json
'-- sub-01
    |-- anat
    |   |-- sub-01_T1w.json
    |   |-- sub-01_T1w.nii.gz
    |   |-- sub-01_T2w.json
    |   '-- sub-01_T2w.nii.gz
    |-- dwi
    |   |-- sub-01_dwi.bval
    |   |-- sub-01_dwi.bvec
    |   |-- sub-01_dwi.json
    |   '-- sub-01_dwi.nii.gz
    |-- fmap
    |   |-- sub-01_dir-fwd_run-01_epi.json
    |   |-- sub-01_dir-fwd_run-01_epi.nii.gz
    |   |-- sub-01_dir-fwd_run-02_epi.json
    |   |-- sub-01_dir-fwd_run-02_epi.nii.gz
    |   |-- sub-01_dir-fwd_run-03_epi.json
    |   |-- sub-01_dir-fwd_run-03_epi.nii.gz
    |   |-- sub-01_dir-rvs_run-01_epi.json
    |   |-- sub-01_dir-rvs_run-01_epi.nii.gz
    |   |-- sub-01_dir-rvs_run-02_epi.json
    |   |-- sub-01_dir-rvs_run-02_epi.nii.gz
    |   |-- sub-01_dir-rvs_run-03_epi.json
    |   '-- sub-01_dir-rvs_run-03_epi.nii.gz
    '-- func
        |-- sub-01_task-restingstate_bold.json
        |-- sub-01_task-restingstate_bold.nii.gz
        |-- sub-01_task-stroop_run-01_bold.json
        |-- sub-01_task-stroop_run-01_bold.nii.gz
        |-- sub-01_task-stroop_run-02_bold.json
        '-- sub-01_task-stroop_run-02_bold.nii.gz

% Field Maps
Fieldmaps = {'DTI_cal_fwd','DTI_cal_rvs','SE EPI topup fwd','SE EPI topup rvs'};
```

NIBAL: Temperament-01-BIDS.sh

```

echo $anatList
if [ -z "${anatList}" ]; then
    echo "No anatomical scans detected...Skipping this step..."
else
    tmp=(${anatList})
    echo "Setting up anatomy"
    # create the T1w, T2w images
    for direct in $anatList; do
        dcm2niix_afni -o $(niidir)/sub-${subjj}/anat -f ${subjj}_${f}_p ${dcmkdir}/${fullID}/${
$[examNum]}/${direct}
    done

    #Changing directory into the subject folder
    cd $(niidir)/sub-${subjj}/anat

    # Rename T1w anat files
    # BIDS template is:
    # sub-<participant_label>[_ses-<session_label>][_acq-<label>][_ce-<label>]
    #   [_rec-<label>][_run-<index>][_mod-<label>][_modality_label].nii.gz
    # Example filename: 6185-s7833_MPRAGE_PROMO_HCP_8
    # Our BIDS filename: sub-6185[_rec-PURE]_T1w, because we have a set where PURE intensity
correction was applied
    # Capture the number of anat files to change
    anatfiles=$(ls -l *MPRAGE* | wc -l)
    for ((i=1;i<=${anatfiles};i++)); do
        Anat=$(ls *MPRAGE*) #This is to refresh the Anat variable, if this is not in the l
each iteration a new "No such file or directory error", this is because the filename was changed.
        tempanat=$(ls -l $Anat | sed 'lq;d') #Capture new file to change
        tempanatext="${tempanat##*.)}"
        tempanatfile="${tempanat.*}"
        rec=""
        if [[ $tempanat == *"PU"* ]] ; then
            rec="_rec-PURE"
        fi
        mv ${tempanatfile}.${tempanatext} sub-${subjj}${rec}_T1w.${tempanatext}
        echo "${tempanat} changed to sub-${subjj}${rec}_T1w.${tempanatext}"

        # if the JSON file, need to add some fields
        # we are adding head coil information, as well as the following fields for NDAR pu
        # acquisition_matrix: we get this from the NIFTI header [dim1, dim2, dim3] with fs
        # mri_field_of_view_pd: Can calculate from acquisition matrix and image resolution
        # photomet_interpret: from DICOM 0028 0004
        # transformation_type: dcm2niix
        # image_num_dimensions: fslhld dim0 field
        # image_unit1, image_unit2, image_unit3: fslhld vox_units
        # image_resolution1, image_resolution2, image_resolution3: fslhld pixdim1, pixdim2,
        # image_resolution4: pixdim4 (if > 1, as this is temporal)
        # image_slice_thickness: related to orientation and pixdim
        # image_orientation: DICOM 0019 1018: L or R = sagittal; A or P = coronal; I or S
        if [[ $tempanatext == "json" ]] ; then
            sDir=$(tempanat:S:S) # this gets the s-directory for our data
            # figure out our software release, headcoil to add to the json file later
sh Tab Width: 8 Ln 13, Col 48

```

```

1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31 32 33 34 35 36 37 38 39 40 41 42 43 44 45 46 47 48 49 50 51 52 53 54 55 56 57 58 59 60 61 62 63 64 65 66 67 68 69 70 71 72 73 74 75 76 77 78 79 80 81 82 83 84 85 86 87 88 89 90 91 92 93 94 95 96 97 98 99 100
aggrethe@induction:/data/EDResearch/TEMPERAMENT_PBU/code/100
| | | -- sub-6251_task-MID_run-8_bold.json
| | | -- sub-6251_task-rest_bold.json
| | -- sub-6251_scans.tsv
| -- sub-6253
| -- anat
| | | -- sub-6253_T1w.json
| | | -- sub-6253_T1w.nii.gz
| | | -- sub-6253_T2w.json
| | | -- sub-6253_T2w.nii.gz
| | | -- sub-6253_rec-PURE_T1w.json
| | | -- sub-6253_rec-PURE_T1w.nii.gz
| | | -- sub-6253_rec-PURE_T2w.json
| | | -- sub-6253_rec-PURE_T2w.nii.gz
| -- dwi
| | | -- sub-6253_acq-pe0_dwi.json
| | | -- sub-6253_acq-pe1_dwi.json
| -- fmap
| | | 6253_s14778_SE_EPI_topup_fwd.json
| | | 6253_s14778_SE_EPI_topup_fwd.nii
| | | 6253_s14779_SE_EPI_topup_rvs.json
| | | 6253_s14779_SE_EPI_topup_rvs.nii
| | | 6253_s14788_SE_EPI_topup_fwd.json
| | | 6253_s14788_SE_EPI_topup_fwd.nii
| | | 6253_s14789_SE_EPI_topup_rvs.json
| | | 6253_s14789_SE_EPI_topup_rvs.nii
| | | 6253_s14793_MB_Resting_topup_fwd.json
| | | 6253_s14793_MB_Resting_topup_fwd.nii
| | | 6253_s14794_MB_Resting_topup_rvs.json
| | | 6253_s14794_MB_Resting_topup_rvs.nii
| -- func
| | | -- sub-6253_task-MID_run-01_events.tsv
| | | -- sub-6253_task-MID_run-02_events.tsv
| | | -- sub-6253_task-MID_run-03_events.tsv
| | | -- sub-6253_task-MID_run-04_events.tsv
| | | -- sub-6253_task-MID_run-05_events.tsv
| | | -- sub-6253_task-MID_run-06_events.tsv
| | | -- sub-6253_task-MID_run-07_events.tsv
| | | -- sub-6253_task-MID_run-08_events.tsv
| | | -- sub-6253_task-MID_run-1_bold.json
| | | -- sub-6253_task-MID_run-2_bold.json
| | | -- sub-6253_task-MID_run-3_bold.json
| | | -- sub-6253_task-MID_run-4_bold.json
| | | -- sub-6253_task-MID_run-5_bold.json
| | | -- sub-6253_task-MID_run-6_bold.json
| | | -- sub-6253_task-MID_run-7_bold.json
| | | -- sub-6253_task-MID_run-8_bold.json

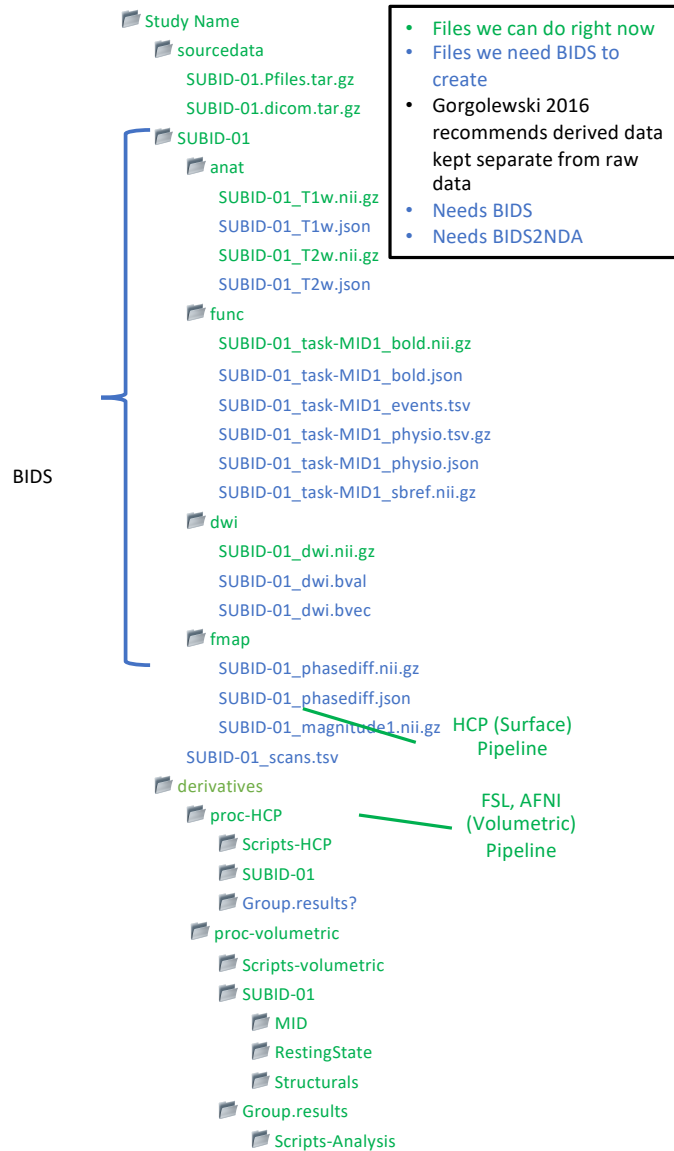
```


Additional resources

- <http://bids.neuroimaging.io>
 - Documentation, list of BIDS-compliant software, etc
- <http://reproducibility.stanford.edu>
 - Select blog for 3 different approaches on creating BIDS compliant datasets
- <https://github.com/INCF/BIDS2NDA>
 - Tool to convert your BIDS data to NDA compliant data to upload to NIMH
 - Seems to lag behind NDA requirements
- <http://www.reproducibleimaging.org>
 - ReproNim study flow tools
 - Heudiconv tool to convert Siemens acquired DICOM to BIDS



Temperament Directory Structure



Example task file

sub-control01/

func/

sub-control01_task-stopsignal_events.tsv

onset	duration	trial_type	response_time	stim_file
1.2	0.6	go	1.435	images/red_square.jpg
5.6	0.6	stop	1.739	images/blue_square.jpg

Participant file

- Describe participant properties like age, group, sex, etc
- The only required argument is `participant_id`, which should match the `sub-<participant_label>` in your study
- Note this file, `participants.tsv` [and its json counterpart] is optional