Data Archive - ADNI and RADC

ADNI

Downloading New ADNI Data

Visit the ADNI Website: http://adni.loni.usc.edu/ and select the 'Access Data Samples' in the Data and Samples dropdown menu.

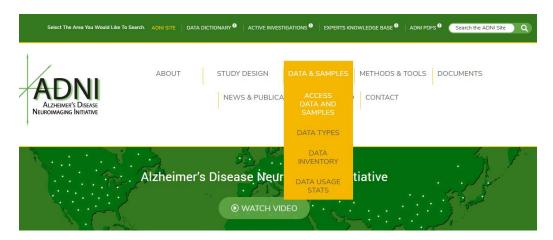


Fig: Accessing ADNI Data Samples

 A login screen will appear. If you do not already have personal access to ADNI data, an application for the same can be made here:

https://ida.loni.usc.edu/collaboration/access/appLicense.jsp;jsessionid=18F0EC9091160BBE877B07EE41C4A8F6

ACCESS DATA AND SAMPLES

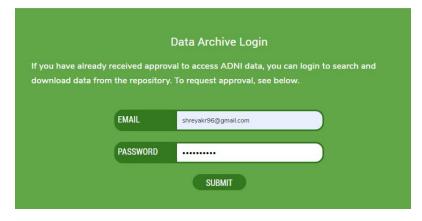


Fig: Logging into the ADNI Database

- This takes you to the ADNI database page. Under the download tab, there are 3
 options:
- **Study Data:** This contains a collection of spreadsheets with different kinds of subject data. This includes test scores, demographic information, diagnosis information etc.
 - Image Collections: This contains all subject scans. This includes structural, functional, diffusion imaging and PET data.
 - Genetic Data: Contains Genome Wide Association Studies (GWAS) and Whole Genome Sequencing (WGS) data for ADNI1, ADNIGO/2 Cohorts.

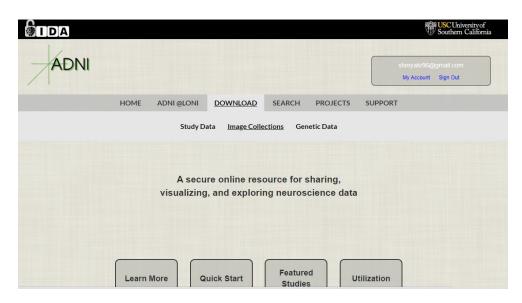


Fig: Navigating the ADNI Data Download Page

• Navigating the Study Data Collection:

- The list in the left column contains the categories under which spreadsheets are arranged.
- The most important table which contains all key ADNI tables merged into one, is present in the **Study Info** Category.
- Select two items Key ADNI tables merged into one (ADNIMERGE table)
 and the corresponding Dictionary and press the Download button on the top
 right corner.
- The former ADNIMERGE table contains a compilation of several entries demographic information like age, gender etc., Diagnosis - Normal, MCI, AD, Several test scores like MMSE, ADAS etc. for all subjects in all 3 ADNI cohorts.

- The latter contains expansions of the column names present within the ADNIMERGE table.
- If specific test scores or neuropathological entries not included in the ADNIMERGE table are required for analysis, they can be downloaded as separate tables and must be sought out from the categories present in the left column.

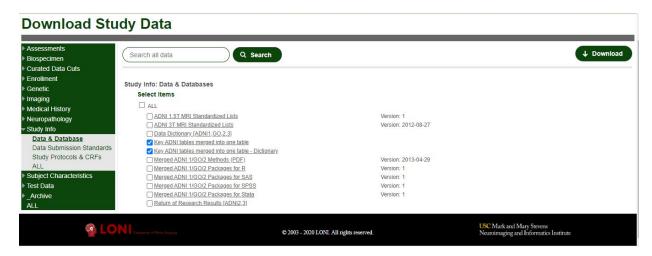


Fig: Downloading the ADNIMERGE table and the dictionary from the Study Data Page

Navigating the Imaging Database:

- Toggle to the 'Advanced Search' Option in Image Collections Window. (See figures below)
- On the left, there are search options that you might want to narrow in on, while downloading subjects' image data.
- Some important fields to check are:
 - Project Phase: This allows you to select the cohorts ADNI1, ADNIGO, ADNI2 and ADNI3 you want to restrict your subject selection to. The factors to be taken into consideration here are the scanner strength for that cohort, availability of certain modalities etc.
 - Subject: This can be used to look for imaging data of only a specific subject if the ID is previously known. It can also be used to look for subjects of a specific age or gender. What this field is most commonly used for is to identify a group of subjects that had the same initial diagnosis (at their baseline visit).
 - Image: This allows you to select the modalities you wish to download from PET, MRI, FMRI and DTI

■ Imaging Protocol: This is primarily used to select a particular scanner manufacturer, as there might be differences in preprocessing steps depending on the manufacturer.

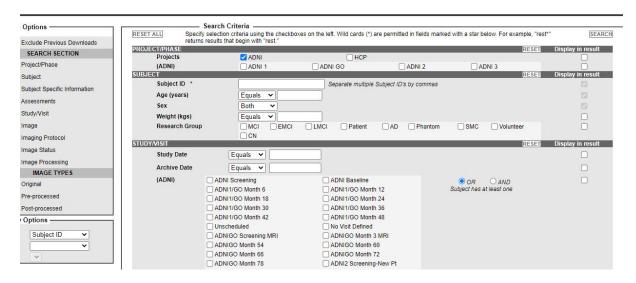


Fig: ADNI Imaging Database Advanced Search Options - Part I

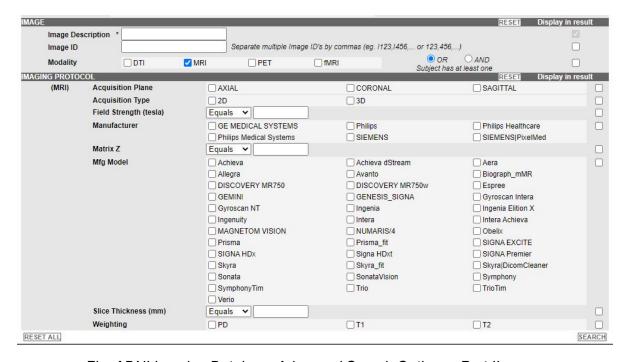


Fig: ADNI Imaging Database Advanced Search Options - Part II

Moving Selected Subjects to Collection:

- Once the search criteria are defined, the search results throw up all images that match the criteria.
- This often includes more than one visit for each subject.

- Select the images you wish to download corresponding to each subject visit as shown in the figure below.
- Select the add to collection on the top right corner. Select an existing collection to add the images to, or create a new collection.



Fig: Selecting Images from the Advanced Search Results



Fig: Adding Images to Collection

Downloading the Selected Images:

- Navigate to the Data Collections Tab.
- Select the Collection you have stored the data in, which is 'Tutorial' in our example (see figure below).
- Select all the scans you wish to download from this collection. Select the 'As
 Archived' option to download in the DICOM format, if the fMRI data is to be
 preprocessed using fMRIprep.
- Click on the '1-Click Download' option, and download the Zip file (see figure below)



Fig: Selecting Images from the Collection to download

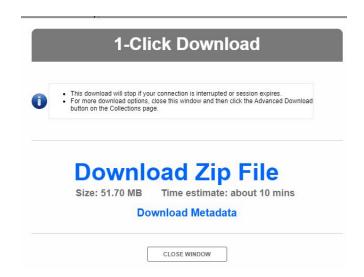


Fig: 1-Click Download Option

Accessing Already Downloaded, Preprocessed and Parcellated ADNI Data:

- Downloaded ADNI Raw Data can be found in Z Drive under the following folder:
 Z:\zSturf\Data\Raw\ADNI. This folder contains subfolders which contain MCIc and MCIs subject scans.
 - MCIc corresponds to subjects who were in MCI and then converted to AD, as seen from their longitudinal data. These scans are taken from visits when these subjects were still in the MCI state.
 - The MCIc subjects are further categorized MCIc-12, MCIc 12-24, and MCIc-24 in these subfolders, which indicates how many months prior to conversion to AD those scans were taken at. This detail might or might not be relevant to your analysis.

 There are also MCIs folders which contain subjects who have stayed in MCI state for at least two years, and never converted to AD until the last known scan.

%% This Data has been preprocessed using the Old Lab pipeline, with a different set of steps, and a different order than what we plan to do in the future with dcm2bids + fmrirpep.

- Preprocessed fMRI data: Data which has been preprocessed using the old lab fMRI preprocessing pipeline can be found here: Z:\zSturf\Data\PreProcessed\ADNI under MCIc and MCIs subfolders.
 - This pipeline did not include susceptibility distortion correction, and there might be errors in slice timing correction. All other standard steps have been applied.
- Parcellated fMRI data: Data which has been parcellated following the above preprocessing pipelines can be found here:
 - Willard 90 Parcellated Data is present in the hard drive labelled Hard Drive 6
 in the following folder F:\Parcellation\Will_90\ADNI containing subfolders
 MCIc and MCIs.
 - Power 264, Willard 14 present in RADC Hard drive (will add links when I have access to it)
 - Parcellation Scripts are present Hard Drive 6: F:\Parcellation\Parcellation

RADC Data

All RADC DICOM and NIFTI information, and their demographic information and visit wise scores have been uploaded to a G-Suite Shared Drive. If you require access, please contact Dr. Sridhar:

https://drive.google.com/drive/folders/16QaAzwSfWCa0aH4rB8K5kp4bX3O10w5X?usp =sharing

Other Backup Locations:

- RADC Raw Data is stored in two hard drives. One labelled RADC and the other labelled Hard Drive 6.
- In the RADC hardrive, the path to the **zipped raw data** is: **E:\MRI_DEVARAJAR**
- In hard drive 6, the path to unzipped RADC raw data is:
 F:\MRI_DEVARAJAR\Unzipped
- Within each of these folders, there are sub-folders MG, BCK and UC for 3 different centres at which scans are collected.
- Some subjects from the MG and BCK center had missing .bvec and .bval files that were separately requested later. These can be found in Hard Drive 6 in this path:
 F:\MRI_DEVARAJAR\New_Scores\bfiles
- Subject Information is stored in: F:\MRI_DEVARAJAR\New_Scores in Hard Drive 6

Subject Information:

- Subject information is present in the following two spreadsheets in the path attached,
 in Hard Drive 6 F:\MRI_DEVARAJAR\New_Scores\:
 - Dataset_611_basic_04-23-2019: This contains information for all subjects included in the study, that stay constant across multiple visits of the same subject. This includes factors like age, gender, race, education level, genetic information like tau tangles, age of death etc.
 - Dataset_611_long_04-23-2019: This sheet contains factors that vary across visits. This includes diagnosis, several cognitive scores, physiological factors like hemoglobin levels, and mental health metrics.
 - RADC_codebook_data_set_611_04-23-2019: This codebook contains
 expansions of the column names in both the spreadsheets above. It also
 contains details of the cognitive tests performed whose scores are reported
 in the above sheets.

RADC Raw Data - Important Details

- As RADC does not share raw data in the DICOM format, some important parameters required for preprocessing which are not present in NIfTI headers can be found in the scanner protocols.
- These protocols can be found on their official website at the attached link: https://www.radc.rush.edu/docs/var/scannerProtocols.htm
- The parameters that have to determined from the above links include Time of Echo
 (TE), slice acquisition order, and the two times of echo for GRE field mapping.
- The protocols for each scanner at each of the different locations is different.
- We have primarily used only the MG centre's data so far because BCK center uses a 1.5T Tesla scanner, and UC does not contain MCI data.
- The UC Center does not have fMRI or Fieldmap Data.
- All fMRI data follows a sequential interleaved slice acquisition.
- MG Center data has GRE fieldmaps 2 Magnitude images and 1 phase difference images.

MG Center Raw Data Example:

Name	Date modified	Туре	Size
DIFFDTI45directions.nii	8/20/2014 5:40 AM	NII File	73,256 KB
DIFFDTI45directionsA.nii	8/20/2014 5:40 AM	NII File	1,593 KB
DIFFDTI45directionsB.nii	8/20/2014 5:40 AM	NII File	1,593 KB
DIFFDTI45directionsC.nii	8/20/2014 5:40 AM	NII File	1,593 KB
DIFFDTI45directionsD.nii	8/20/2014 5:40 AM	NII File	2,390 KB
ep2dfidbasicbold.nii	8/20/2014 5:40 AM	NII File	45,601 KB
FLAIRAxialTE150.nii	8/20/2014 5:40 AM	NII File	4,481 KB
FLAIRAxialTE150A.nii	8/20/2014 5:40 AM	NII File	4,481 KB
grefieldmapping.nii	8/20/2014 5:40 AM	NII File	4,161 KB
grefieldmappingA.nii	8/20/2014 5:40 AM	NII File	2,081 KB
t1mprnssagpat2iso.nii	8/20/2014 5:40 AM	NII File	22,529 KB
t1mprnssagpat2isoA.nii	8/20/2014 5:40 AM	NII File	22,529 KB
t2fl3dtraNoSWI.nii	8/20/2014 5:40 AM	NII File	61,441 KB
t2fl3dtraNoSWIA.nii	8/20/2014 5:40 AM	NII File	61,441 KB
T2mapping.nii	8/20/2014 5:40 AM	NII File	102,001 KB

- fMRI file: ep2dfidbasicbold.nii
- Structural File: t1mprnssagpat2iso.nii
- GRE fieldmap Magnitude File: grefieldmapping.nii (contains 2 volumes each is one magnitude image at a different TE)
- GRE Fieldmap Phase Difference Image: grefieldmappingA.nii

Preprocessed and Parcellated RADC Data Locations:

%% The following data has been preprocessed with the old lab pipeline with a different set and order of steps than dcm2bids + fmriprep

Preprocessed RADC Data:

- It is stored in Z drive in the following location: Z:\zSturf\Data\PreProcessed\RADC
- This folder contains subfolders which contain MCIc and MCIs subject scans.
 - MCIc corresponds to subjects who were in MCI and then converted to AD, as seen from their longitudinal data. These scans are taken from visits when these subjects were still in the MCI state.
 - The MCIc subjects are further categorized MCIc-12, MCIc 12-24, and MCIc-24 in these subfolders, which indicates how many months prior to conversion to AD those scans were taken at. This detail might or might not be relevant to your analysis.

- There are also MCIs folders which contain subjects who have stayed in MCI state for at least two years, and never converted to AD until the last known scan.
- This data has been preprocessed using the old lab fMRI pipeline
- This pipeline did not include susceptibility distortion correction, and there might be errors in slice timing correction. All other standard steps have been applied.

Parcellated RADC Data:

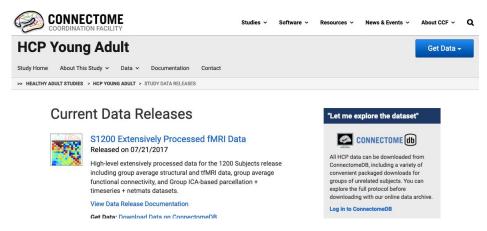
- Parcellation Scripts are present Hard Drive 6: F:\Parcellation\Parcellation
- Parcellated data locations to be added.

Human Connectome Project (HCP) Data

Registering on the ConnectomeDB System and Amazon Web Services

 Visit the Connectome Coordination Facility that will direct you to the Connectome DB database management system.

https://www.humanconnectome.org/study/hcp-young-adult/data-releases



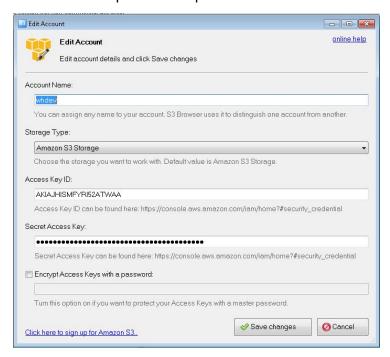
- Click on the "Get Data" button on the top right and choose "Download from Connectome DB".
- Create your own account by registering.
- Once you login, you will see the following page with all the datasets such as the following:

Public Connectome Data

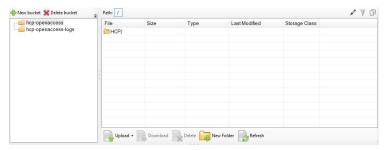
Updated April 2018: All 7T functional MR data (184 subjects) re-preprocessed to remove unwarping bug. FIX-cleaned retinotopy fMRI data now available (183 subjects). NIH Toolbox Cognition composite scores added.



- Accept the Data Use Terms by clicking on the tab circled in blue shown above.
- Next, you must get access to the Amazon S3 browser. For this you need AWS
 credentials. Click on the "Data Available on Amazon S3" tab circled in red shown
 above and follow the steps provided here to create your credentials. Please
 remember to safely save your username, access key ID and secret access key for
 future reference.
- Download the S3 browser for Windows <u>here</u>.
 - Once you launch the S3 browser, you should be promyed to create an account. Enter your access key ID and secret access key that you obtain with your credentials in the previous step.



Once done, you should be able to see the HCP datasets as shown below:



- For Linux, using the AWS Command Line Interface is an easier approach. You can
 follow the instructions here to install the awscli package for your system based on
 your OS and version of Python.
 - Once installed, open a new terminal window and type

\$ aws configure

- You will be prompted to enter your access key ID and secret access key that you obtain with your credentials in the previous step. Once you enter these, you're good to go. Close the terminal
- To test if awscli has been installed correctly, type in a new terminal window

\$ aws s3 ls

• You should be able to see the HCP database as follows:

```
[varsha@dti:~$ aws s3 ls
2018-11-05 16:53:55 hcp-openaccess
2018-07-26 02:49:34 hcp-openaccess-logfiles
2018-04-20 21:18:28 hcp-openaccess-logs-temp
2018-06-08 21:45:19 hcp-openaccess-logstorage-temp
2018-06-29 18:47:03 hcp-openaccess-test
2018-04-20 21:13:24 hcp-openaccess-trail-temp
```

 Detailed instructions can be found here and other information can be found here.

Downloading HCP data from using Amazon S3

Windows

- Launch the S3 browser, enter the folder called HCP/
- You will see a list of all the subjects whose data is available. Enter the folder for a subject and select the file to download. Click the download button at the bottom. You will be prompted to choose a download location. Navigate to your folder of choice and commence download.

Linux

- Open a new terminal.
- To view the all the subjects' data available type

\$ aws s3 Is s3://hcp-openaccess/HCP_1200/

You should see a list something like this:

```
varsha@dti:~$ aws s3 ls s3://hcp-openaccess/HCP_1200/
                            PRE 100206/
                            PRE 100307/
                            PRE 100408/
                            PRE 100610/
                            PRE 101006/
                            PRE 101107/
                            PRE
                                101309/
                                101410/
                            PRE
                            PRE 101915/
                            PRE 102008/
                            PRE 102109/
                            PRE 102311/
                            PRE 102513/
                            PRE 102614/
                            PRE
                                102715/
                            PRE 102816/
                            PRE 103010/
```

Next, to view all the data for a given subject, say 100206, type

\$ aws s3 Is s3://hcp-openaccess/HCP_1200/100206/

```
varsha@dti:~$ aws s3 ls s3://hcp-openaccess/HCP_1200/100206/
PRE .xdlm/
PRE MNINonLinear/
PRE T1w/
PRE release-notes/
PRE unprocessed/
varsha@dti:~$
```

 To download a file, say for example the T1-weighted image which is in the T1w folder, type

\$ aws s3 cp s3://hcp-openaccess/HCP_1200/100206/foldername/... filename path_to_destination

o To download the entire folder, type

\$ aws s3 cp s3://hcp-openaccess/HCP_1200/100206/foldername/...
path_to_destination --recursive

Where to find minimally preprocessed DTI data

What you will need:

- 1. T1-weighted image: T1w_acpc_dc_restore_brain.nii.gz (brain extracted image)
- 2. DTI data: data.nii.gz, bvals, bvecs
- 3. Optional: nodiff_brain_mask.nii.gz (brain mask), grad_dev.nii.gz
- 4. Freesurfer data: parcellations and segmentations

Where will you find it:

- 1. s3://hcp-openaccess/HCP_1200/subject_id/T1w/
- 2. s3://hcp-openaccess/HCP_1200/subject_id/T1w/Diffusion/
- 3. s3://hcp-openaccess/HCP 1200/subject id/T1w/Diffusion/
- 4. s3://hcp-openaccess/HCP_1200/subject_id/T1w/subject_id/

It is recommended that you download the entire "Diffusion" folder. This will cover #2-3. For #4, there must be subfolders by the name **label/**, **mri/**, **surf/**, **stats/**, **touch/**. It is advised that you download all the subfolders.

Where to find minimally preprocessed fMRI data

What you will need:

- 1. T1-weighted image: T1w acpc dc restore brain.nii.gz (brain extracted image)
- fMRI data: files for 2 rest scans and 7 task scans, each in L>>R and R>>L phase-encoding (total of 18 datasets)
- 3. Movement regressors and derivatives (optional): corresponding to each fMRI dataset

What will you find it:

- 1. s3://hcp-openaccess/HCP_1200/subject_id/T1w/
- s3://hcp-openaccess/HCP_1200/100206/MNINonLinear/Results/rfMRI_REST1_LR/rf MRI_REST1_LR.nii.gz (first run of rest fMRI, say). There is one folder corresponding to each of the 18 datasets.
- s3://hcp-openaccess/HCP_1200/100206/MNINonLinear/Results/rfMRI_REST1_LR/M ovement_Regressors.txt, Movement_Regressors_dt.txt.