DATA CONTRIBUTION GUIDE

Outlined below are the steps that need to be completed to successfully contribute data to INDI. The primary goal of the data preparation process is to anonymize your data prior to transferring them to the INDI team. A member of the INDI team will be available to help with and troubleshoot any difficulties that arise in the process, to provide support and to ensure the process goes smoothly. *If preferred we can develop customized scripts to automate data preparation specifically for your data (that can used be locally at your site), or we can carry out the data preparation process for you (if your IRB/ethics board approves the transfer of non-anonymized data to us).*

Share that brain!

The INDI Team

1. **Data Requirements**
   a. At a minimum, we require one T1-weighted anatomical scan, and one resting state fMRI scan per participant. However, we gladly accept any additional data you are willing to contribute (e.g., DTI, T2 or ASL scans, phenotypic information).
   b. NIFTI is the default data format for INDI due to the decreased risk of unintended information being present in the header. If you'd prefer to make your contribution in DICOM, our team can provide you with a Nipype-based DICOM anonymizer that will insure HIPAA-compliant anonymization.

2. **Data Anonymization**
   a. To protect participant privacy, datasets on INDI need to be anonymized. This means there cannot be any information in the images or image-headers that would allow others to identify the participant. In practice, images need to be defaced (i.e., remove the face from any high-resolution images such as an MPRAGE). In addition, any identifying information, such as date of birth, date and time of scan, and the name of the participant, needs to be stripped from the headers of all contributed images (including command history for AFNI users). Finally, each subject's ID has to be changed to an anonymized numeric that we will provide to you.
   b. The following script for data anonymization is available via INDI (http://fcon_1000.projects.nitrc.org/).
      i. FullAnonymize <input file> <T1: 1 = yes, 0 = no)
         • Purpose: Cleans header and defaces file (if T1)
         • Output: fully anonymized image file
         • Example usage: anat.nii.gz 1
         • Example output: anat_anonymized.nii.gz

Note: FSL must be installed to execute the defacing features within the anonymization script.

As part of our follow-up we will send you a list of numbers you should use to 'rename' your subjects. Hold on to the key you are using to anonymize your data. This will be useful when a question regarding your data comes up (e.g., there is something about the scan of 000001).

3. **Data Organization**

As mentioned, we require at least one anatomical and one resting state fMRI scan per subject you want to contribute. The preferred directory structure and naming convention is as follows (though any organization is acceptable, as the INDI team ensures standardization prior to release):

0000001/session_1/anat_1/anat.nii.gz
0000001/session_2/anat_1/anat.nii.gz
0000001/session_1/rest_1/rest.nii.gz
0000001/session_1/rest_2/rest.nii.gz
0000001/session_1/rest_3/rest.nii.gz
0000001/session_2/rest_1/rest.nii.gz
0000001/session_2/rest_2/rest.nii.gz
0000001/session_2/rest_3/rest.nii.gz
etc.

In the event of multiple sessions, it is important that you maintain session order by numbering them chronologically. This will benefit people interested in studying effects of scan order.

4. **Data Reorientation**

To ensure uniformity across datasets provided by INDI, we provide all datasets in RPI orientation. We can accept data in any orientation, though if possible, please reorient all scans to RPI and ensure the header information of the images is set accordingly. Finally, check that the left/right orientation is in accordance with your original data. Following receipt and review of the images received from your site, the INDI team will send you axial images from five randomly selected datasets from your site to verify left-right orientation.

5. **Data Phenotyping**

We require age, sex, and handedness information for all contributed datasets. We gladly accept any additional phenotypic information you would like to share (e.g., questionnaire measures, behavioral and diagnostic information, BMI and weight information etc.).

Organize your data in a spreadsheet using the anonymized INDI subject numbers. *Note: Specific consortium initiatives (e.g., ADHD-200, ABIDE, CoRR) may have additional criteria for phenotyping; for such efforts, the consortium leaders will provide a data key and corresponding spreadsheet template.*

6. **Data Uploading**

   Let us know when your data are ready. We will then send you information about how to upload the data to the upload portal (depending on the initiative, the upload portal may be located at either NITRC or COINS). **Please note that we will double-check the data before they are publicly released.**

7. **Webpage Contents**

   To optimally share your data, we will put together a **webpage** dedicated to your data sample. For this we require:
   ○ A name for your data sample
   ○ A logo for your site (.png is the preferred file format)
   ○ A list of the names and titles of the PI and all senior personnel affiliated with your data sample
   ○ Acknowledgements - you can be as generous as you would like with this list
   ○ A list of grants that you want acknowledged
   ○ A list of any publications using the data you are contributing
   ○ Any behavioral information about the resting state scan.
   ○ Were the subject's eyes open or closed?
   ○ What instructions were given to the subject just before the resting-state scan?
   ○ What visual stimulus was presented during the scan? (e.g., fixation cross, blank screen, 'relax')
   ○ Scan Parameters
      A PDF of each scan's sequence protocol (i.e. anatomical, resting state, DTI…) obtained from the scanner console.

7. **Data Downloading**

   We will send you the completed webpage before making it public. After you have approved it, the page will be made public and people can start downloading your data. You will then achieve ultimate happiness and will be remembered as a great, sharing scientist. ;)