Instructions on how to download imaging data from AWS for the EEG/FMRI Naturalistic Viewing Study from the Nathan Kline Institute (NKI)

There are two options on how to download neuroimaging data from the EEG/FMRI Naturalistic Viewing Study. Option 1 is based on a *bash* script and Option two is using Cyberduck. While option 1 is a faster to use, option 2 is useful if you are not comfortable using a command line and cannot install AWS's command line interface (CLI) on your computer/server. We recommend reading through both methods before selecting which option is best for you.

Option 1: Download using a bash script:

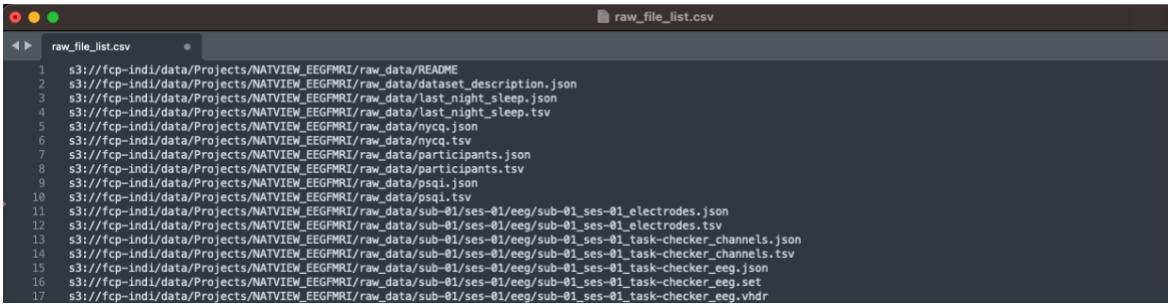
1. You will need to install the lasted version of AWS's Command Line Interface
   You can follow step-by-step instructions on AWS's website here.

   Note: for using our download script you do not need to perform a setup, such as setting up the AWS Access Key ID or AWS Secret Access Key.

2. Next, you will need to download our "NAT_VIEW_download_links.sh" script. You can download it here.

3. Download the full list of links to the files on AWS. You can download the *raw data* csv file "raw_file_list.csv" file from here, or download the list with the preprocessed EEG and FMRI data "processed_file_list.csv" from here.

4. If you downloaded a csv file, if you link, you can modify the rows and leave only files that you would like to download (for example, only want the EEG data, not the FMRI). For example, you will need a file that looks like this:



5. Now you are ready to download the files. You will need to open a terminal and go to the location where you saved the "NAT_VIEW_download_links.sh" script. Next, run the following command:

   bash NAT_VIEW_download_links.sh -i raw_file_list.csv -o ./DOWNLOADFOLDER/

Note 1: make sure that the list with the "raw_file_list.csv" is in the same location of the script or you can also include the full path. Also, if you would like to download the data to another location, make sure to include the full path.

Note 2: If you are running into firewall issues at your institution, you will need to modify the script. In line 36 you will need to include an additional flag in the aws command. The new line will need to look like this:
aws s3 cp ${filepath} ${output}/${exactPath} --no-sign-request --no-verify-ssl

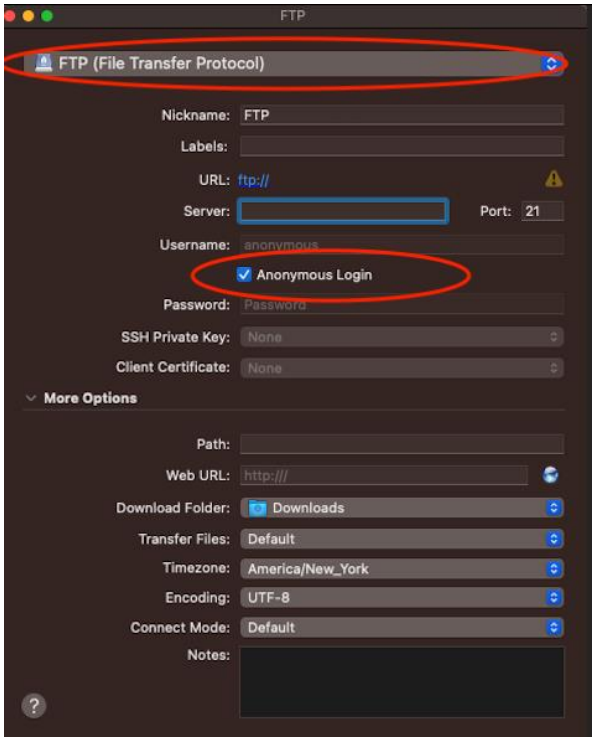
Now while downloading the files, you will receive the following warning:
" Unverified HTTPS request is being made to host 'fcp-indi.s3.amazonaws.com'. Adding certificate verification is strongly advised. See:
https://urllib3.readthedocs.io/en/latest/advanced-usage.html#ssl-warnings"

You can ignore these warnings.

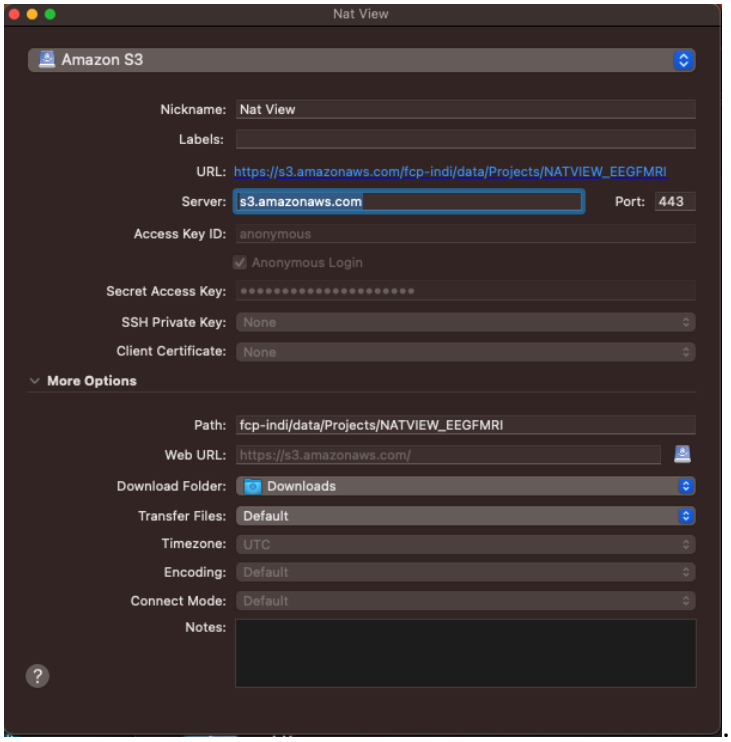
Option 2: Download using Cyberduck:

There are file transfer programs that can handle S3 natively and will allow you to navigate through the data using a file browser. Cyberduck is one such program that works with Windows and Mac OS X. Cyberduck also has a command line version that works with Windows, Mac OS X, and Linux. Step-by-step instructions for using Cyberduck are show below. Due to some issues of connecting to S3 bucket on Cyberduck using an anonymous login on S3 buckets, you will need to take the following steps for a workaround (this might change in more recent version):

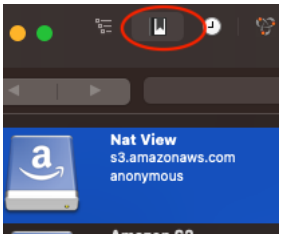
1. Install Cyberduck following the instructions provided on their website.
2. Open Cyberduck and click on *Bookmarks* (in the menu bar) then *New Bookmark*.
3. Set the application protocol in the dropdown menu to *FTP (File Transfer Protocol)*.
4. Check the "Anonymous Login" checkbox. See figure below:



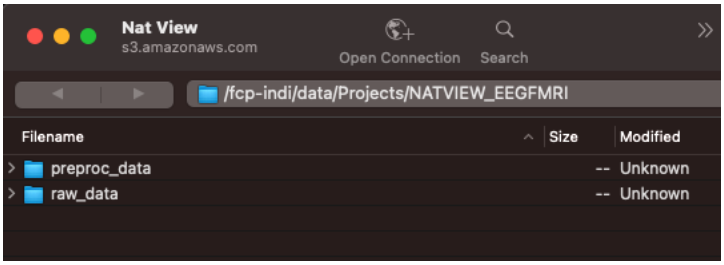
5. Now change the application protocol to "*Amazon S3"* in the dropdown menu. This should keep the "Anonymous Login" checkbox still checked.
6. Leave the server as "s3.amazonaws.com"
7. Change the Nickname to "Nat View" (or to any other name)
8. Expand the More Options tab and SET Path to:
   fcp-indi/data/Projects/NATVIEW_EEGFMRI
9. The window should look like this:



10. You can close the configuration window and the new “Nat View” bookmark should appear in the bookmark tab



11. Double click on the “Nat View” bookmark to connect. You can now navigate through the files in Cyberduck and download the data to your local computer



Questions:

If there are any questions on accessing these data, you can contact Alex Franco (alexandre.franco@nki.rfmh.org) or Qawi Telesford (qtelesford@gmail.com)