Instructions on how to download imaging data from AWS for the ADHD200 study

There are two options for downloading neuroimaging data from the ADHD200 study. Option 1 uses Cyberduck, and Option 2 is based on a bash script. While option 2 is faster, option 1 is helpful if you are uncomfortable 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 Cyberduck

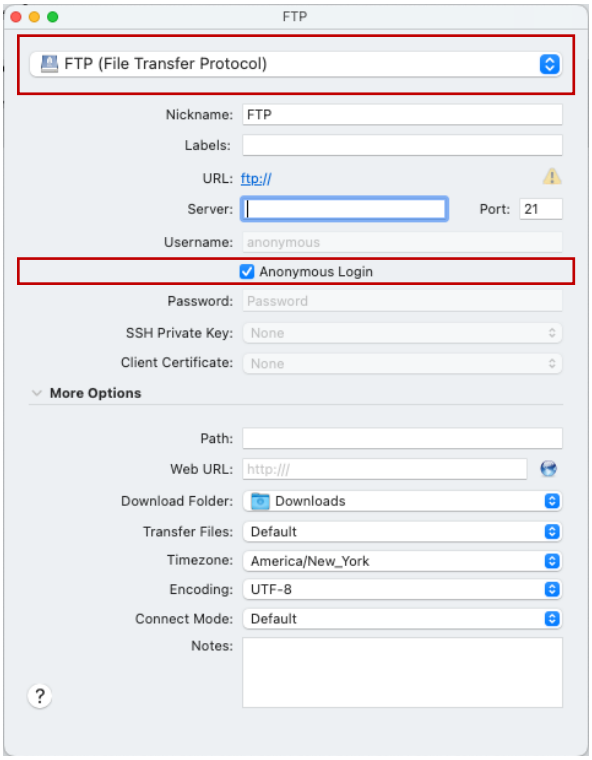
There are file transfer programs that can handle S3 natively and will allow you to navigate through the data using a file browser. Cyberduck is one such program that works with Windows and Mac OS X. Cyberduck also has a command line version that works with Windows, Mac OS X, and Linux. Step-by-step instructions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

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 the figure below:



5. Now, change the application protocol to *Amazon S3* in the dropdown menu. This should keep the *Anonymous Login* checkbox checked.

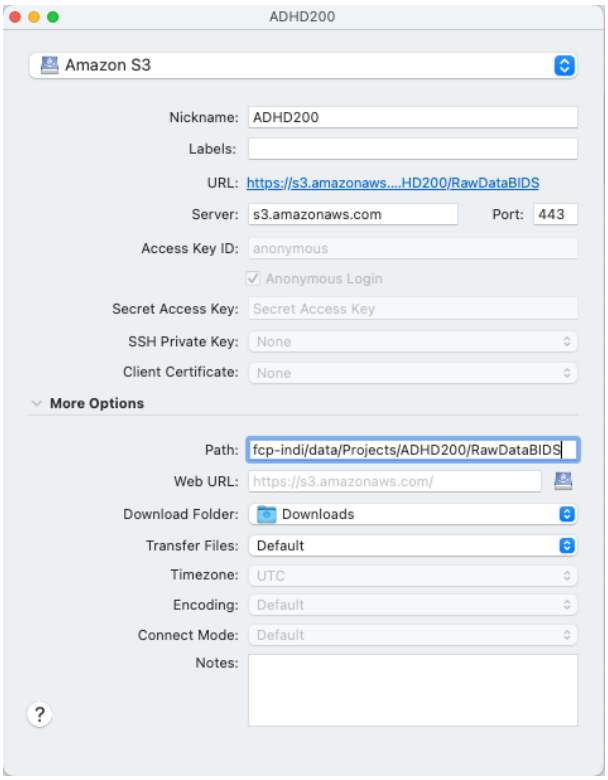
6. Leave the server as *s3.amazonaws.com*.

7. Change the *Nickname* to ADHD200 (or to any other name).

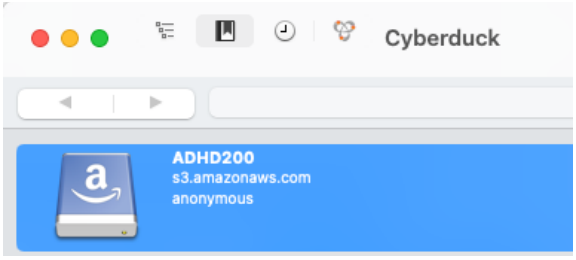
8. Expand the *More Options* tab and set the *path* to:

```
fcp-indi/data/Projects/ADHD200/RawDataBIDS
```

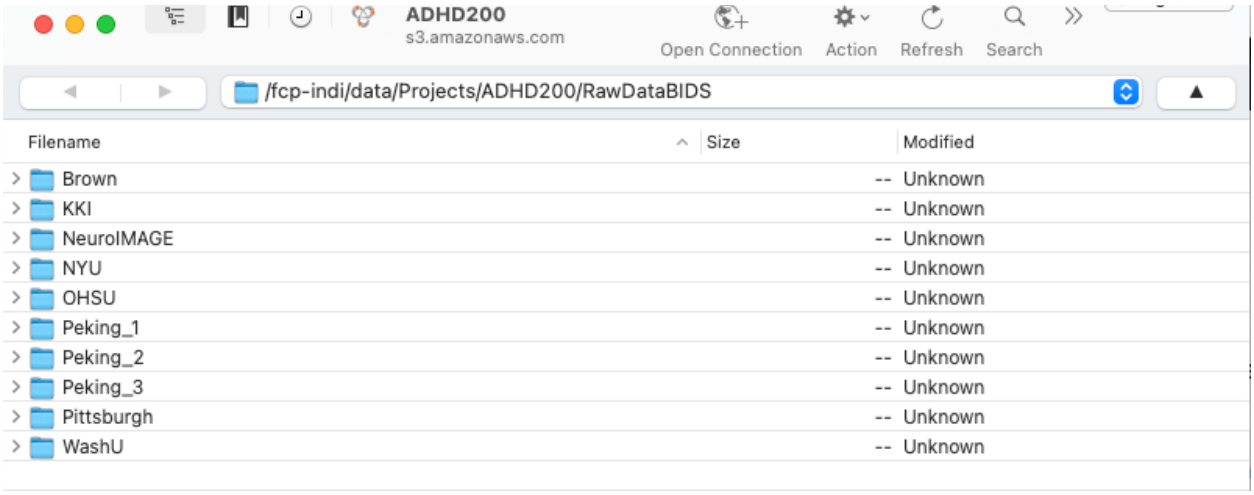
9. The window should look like this:



10. You can close the configuration window and the new *ADHD200* bookmark should appear in the bookmark tab.



11. Double-click on the *ADHD200* bookmark to connect. You can now navigate through the files in Cyberduck and download the data to your local computer.



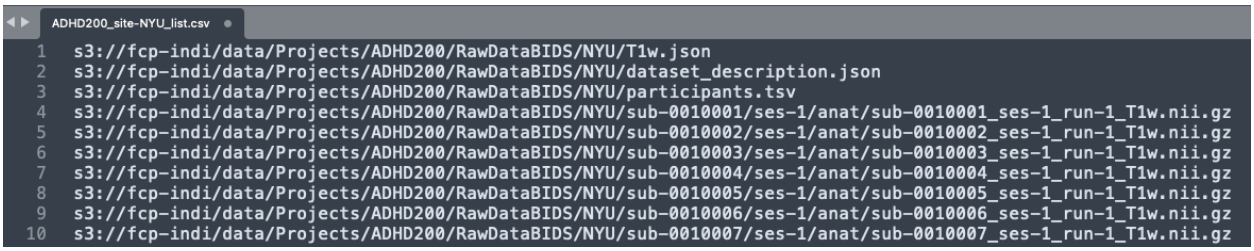
Option 2: Download using a bash script

1. You will need to install the latest 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 "ADHD200_download_links.sh" script. You can download it here.

3. Now, you will need to download the list of links to the files on AWS. There is one file per study site, and you can download each one below:
   - Bradley Hospital / Brown University
   - Kennedy Krieger Institute
   - NeuroIMAGE
   - New York University – Child Study Center
   - Oregon Health & Science University
   - Peking University
   - University of Pittsburgh
   - Washington University in St. Louis

4. If necessary, you can modify the rows of the downloaded csv file and leave only files that you would like to download. For example, if you want the anatomical data and not the functional data, 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 "ADHD200_download_links.sh" script. Next, run the following command:

```
bash ADHD200_download_links.sh -i file_list.csv -o download_folder
```

Note 1: make sure that the list with the files (file_list.csv) is in the same location as the script or you will need to 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 39, 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-signal-request
--no_verify_ssl
```

Now, while downloading the files, you will receive the following warning:

```
urllib3/connectionpool.py:1013: InsecureRequestWarning: Unverified HTTPS request is being made to
 host 'fcp-indi.s3.us-east-1.amazonaws.com'. Adding certificate verification is strongly advised.
 See: https://urllib3.readthedocs.io/en/1.26.x/advanced-usage.html#ssl-warnings
```

You can ignore these warnings.

Questions:

If there are any questions about accessing these data, you can contact Nathalia Esper (nathalia.esper@childmind.org) or Alexandre Franco (alexandre.franco@childmind.org).