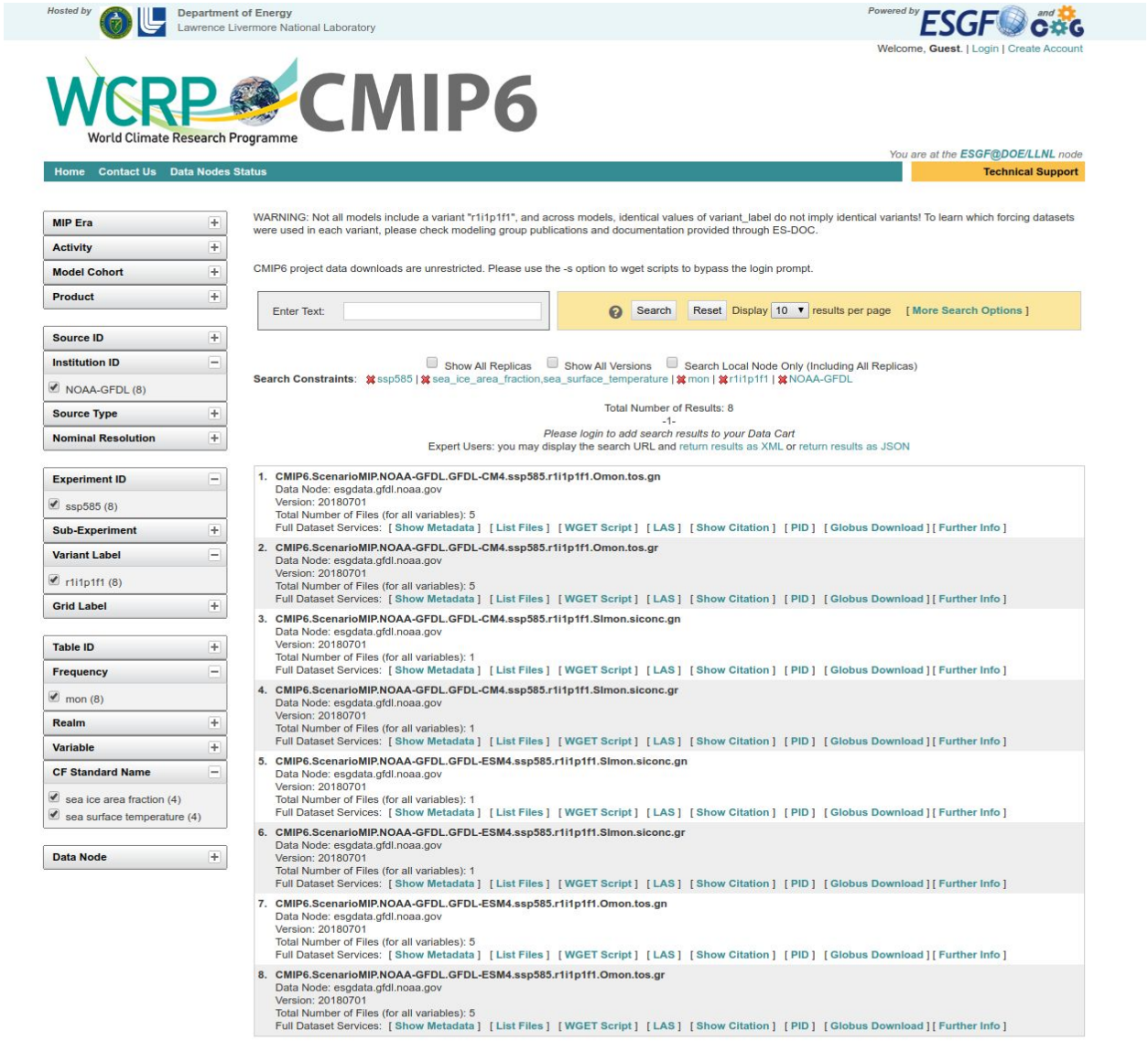
CMIP6 Download Instructions

This is a walkthrough designed to help you download CMIP6 from ESGF. The data is currently open-access, so you won't need to sign up for an account.

1. Go to https://esgf-node.llnl.gov/search/cmip6/



2. Expand the filters and check boxes for variables/experiment/time resolution/model institution you want to download, then hit search. In the example above I've searched for sea surface temperature and sea ice area fraction for the GFDL model's SSP585 experiment at a monthly resolution. The variant label label 'r1i1p1f1' is the *ripf* label, the

***r**ealization, **i**ndex, **p**hysics, **f**orcing* indices. R1i1p1f1 is recommended, selecting it helps to narrow down the number of results (often there will be hundreds).

3. You may click ‘Show Metadata’ to view more details of the data.
4. When you have decided which data you want to download, click ‘WGET script’ and a .sh file will download.

5. Go into your jupyter notebook and create a folder to store the data. Drag the wget scripts from your local machine’s Downloads folder or from your browser’s downloads bar into the new folder. Click the Upload button.
6. Now open up a terminal and use the cd command to change to the directory where the wget scripts are located:

   cd other_data/
7. For a single wget script, or if you want to run them one at a time you may just enter

   sh <wget_name.sh> -s
8. If you want to run more than one wget script at once you can type python into your terminal and follow each of the following lines (paste into the terminal window), pressing enter twice after os.system('sh '+ i+' -s') , the tab before os.system('sh '+ i+' -s') is important!!!

```
from glob import glob
import os
f = glob(os.getcwd()+'/*.sh')
for a,i in enumerate(f):
    os.system('sh '+ i+' -s')
```

```
(base) [twemmen@jupyter other_data]$ python
Python 3.7.6 (default, Jan  8 2020, 19:59:22)
[GCC 7.3.0] :: Anaconda, Inc. on linux
Type "help", "copyright", "credits" or "license" for more information.
>>> from glob import glob
>>> import os
>>> f = glob(os.getcwd()+'/*.sh')
>>> for a,i in enumerate(f):
...     os.system('sh '+ i+' -s')
...
######## This is a bash script! ##############
Change the execution bit 'chmod u+x /home/twemmen/other_data/wget-20200422171113.sh' or start with 'bash /home/twemmen/other_da
ta/wget-20200422171113.sh' instead of sh.
Trying to recover automatically...
********************************************************************************
*                                                                              *
* Note that new functionality to allow authentication without the need for     *
* certificates is available with this version of the wget script.  To enable,  *
* use the "-H" option and enter your OpenID and password when prompted:        *
*                                                                              *
* $ wget-20200422171113.sh -H [options...]                                     *
*                                                                              *
* For a full description of the available options use the help option:          *
*                                                                              *
* $ wget-20200422171113.sh -h                                                  *
*                                                                              *
********************************************************************************
Running wget-20200422171113.sh version: 1.3.2
Use wget-20200422171113.sh -h for help.

Script created for 1 file(s)
(The count won't match if you manually edit this file!)
```

Your terminal should look like this after pressing enter twice

9. Now the netCDF files will be in your current directory. You can use the data_skeleton scripts to adapt your code to make plots of the data!
10. If you wish to interpolate the data (model comparison) or if you wish to combine multiple files into a single .nc file, you will need to copy the regrid_skeleton file and follow the instructions there.