The download CMIP6 data using the RESTful API

Author: A. Novoselsky

Last update: July 5, 2021

The CMIP6 data can be downloaded from a repository using Web API: e.g., from the Lawrence Livermore National Laboratory (LLNL) repository via the Web page <https://esgf-node.llnl.gov/search/cmip6/>. This way has definite restrictions:

* To be downloaded, the search result should be added first to the Data Card by portions no more than 100 files,
* The number of download files should not exceed 1000.

The mentioned restrictions are rather cumbersome for a download of big data set. In this case, using the RESTful API can be a convenient and fast alternative.

The description of the RESTful API queries can be found on the Web page <https://esgf.github.io/esgf-user-support/user_guide.html#restful-api> or in other online manuals.

We’ll learn this approach step by step in the following example.

1. Define the parameters of the needed data set.

MIP Era CMIP6

Experiment ID amip

Variant Label r1i1p1f1

Table ID day

Frequency day

Realm atmos

Variable ua, va

2. Build the RESTful API query for this data set.

<https://esgf-node.llnl.gov/esg-search/wget?mip_era=CMIP6&experiment_id=amip&member_id=r1i1p1f1&table_id=day&frequency=day&realm=atmos&variable=ua,va&distrib=true&replica=true&limit=9999>

Such query contains parameters as name=value pairs separated by the ampersand sign (&).

3. Copy this query to the browser search bar.

4. Save the resulting WGET script (e.g., wget-20210704200058.sh) on your computer.

5. Create the folder for downloading, e.g., *bulk\_amip\_day*.

6. Copy the resulting WGET script to this folder.

7. Run the resulting WGET script in this folder using the PuTTY or other client: e.g.

wget-20210704200058.sh –v –s

The WGET script will download the NetCDF files of the requested data set. The downloading time depends on the number and size of NetCDF files: it can take minutes, hours, or even days.