

# **GISAID covCLI™**

## **Command Line Interface (CLI) for batch uploading consensus sequences and metadata to the EpiCoV™ database**

### **1. Description**

This document describes the installation and usage of the covCLI software for making batch submissions of consensus sequence data and metadata to GISAID from the command line.

Use of this software requires familiarity with the \*nix command line terminal. For a basic overview, please refer to myriad [available web resources on the topic](#).

### **2. Installation**

After downloading the software move it to a preferred destination prior to installation. Add the destination to the \$PATH variable to allow running with covCLI -h. For additional help on adding folders to \$PATH, refer to online resources. If not adding to the \$PATH, use a relative path to run the file. For example, change into the directory containing the file and run ./covCLI -h.

Note to macOS users: You need to execute the following commands in the installation directory before using the application:

```
chmod +x covCLI
```

```
xattr -c covCLI
```

This software has not yet been notarized by Apple. As a result, might issue a warning at the start of installation. To overcome this issue, see <https://support.apple.com/en-us/HT202491#openanyway>

### **3. Usage**

#### **Get help, top level**

After running covCLI -h, the following should be printed to screen.

```
usage: covCLI [-h] ...
```

Version 4 Command Line Interface (CLI) for uploading sequences and metadata to GISAID.

options:

```
-h, --help show this help message and exit
```

Sub-commands help:

```
upload    Upload sequences and metadata.
version   Show version and exit.
```

The above output shows two sub-commands:

- covCLI upload (refer to section 3.1)
- covCLI version (refer to section 3.2)

### 3.1. Upload

Use of a GISAID CLI requires a client-ID as an additional security measure that compliments standard GISAID access credentials. Obtain a client-ID by clicking the “Request client-ID” from where you downloaded this software, complete all required fields and click “Request.” Your client-ID may not be shared with others, and it is solely for use with official GISAID uploader software. To protect the integrity of data released in GISAID, safety measures will invalidate a client-ID used with unauthorized uploader software.

Once the metadata file is prepared in CSV format, and a sequences file in FASTA format have been prepared the upload via CLI can begin.

To get help on the upload sub-command, do `covCLI upload -h` or `covCLI upload --help`. If the program is working correctly, the following will be printed to standard out (i.e., ‘stdout’):

```
usage: covCLI upload [-h] [--username USERNAME] [--password PASSWORD] [--clientid CLIENTID] [--log LOG] --metadata METADATA --fasta FASTA [--proxy PROXY] [--dateformat {YYYYMMDD,YYYYDDMM,DDMMYYYY,MMDDYYYY}] [--debug]
```

Perform upload of sequences and metadata to GISAID's curation zone.

options:

```
-h, --help            show this help message and exit
--username USERNAME   Your GISAID username. (default: None)
--password PASSWORD   Your GISAID password. (default: None)
--clientid CLIENTID   Submitter's client-ID. Follow the steps in the README
                      to request a client-ID. (default: None)
--log LOG             All output logged here. (default: None)
--metadata METADATA   The csv-formatted metadata file. (default: None)
--fasta FASTA         The fasta-formatted nucleotide sequences file.
                      (default: None)

--frameshifts {catch_all,catch_novel,catch_none}
                      'catch_none': catch none of the frameshifts or spike
                      truncations and release immediately; 'catch_all':
                      catch all frameshifts or spike truncations and
                      require email confirmation; 'catch_novel': catch
```

```

novel frameshifts or spike truncations and require
email confirmation.
(default: catch_all)
--proxy PROXY Proxy-configuration for HTTPS-Request in the form:
http://username:password@proxy:port (default: None)
--dateformat {YYYYMMDD,YYYYDDMM,DDMMYYYY,MMDDYYYY}
Specify the date format, with 'Y' for 'year', 'M' for
'month', 'D' for 'day'. Dates will parse correctly
with the following delimiters: '/', '.', '-' or '-'.
(default: YYYYMMDD)
--debug Switch on debugging output. (default: False)

```

During upload, a message similar to the following will be printed to stdout.

```

INFO; Logon status: ok;
SUCCESS; hCoV-19/Anzark/TEST_DONOTRELEASE-13/2023 assigned to accession:
EPI_ISL_18641035
SUCCESS; hCoV-19/Anzark/TEST_DONOTRELEASE-17/2023 assigned to accession:
EPI_ISL_18641036
INFO; Logoff status: ok;

```

Any ERROR or WARNING messages will also be printed to stdout here.

When making larger submissions, it is useful to save stdout messages to file. All stdout messages can be redirected to a new file by adding `--log logfile.txt` to the command. If the file exists, messages will be appended to the file.

Following a successful upload, submissions can be edited inside EpiCoV™.

### 3.2 Version

To retrieve the software version number, enter:

```
covCLI version
```

This user manual concurs with version:

```
covCLI version: 4.1.2
```

## 4. Support

For additional CLI queries contact [clisupport\[at\]gisaid.org](mailto:clisupport@gisaid.org).

## 5. FAQs

### *What does the error “submitter invalid” mean?*

The submitter column in the `metadata.csv` file contains an email address that is not authenticated for use by the submitter. To fix, either: 1) change the email address in the row of the metadata file to be equal to the email address of the submitter performing the

upload; 2) email `clisupport[at]gisaid.org` to request adding the email address(es) to the list of authorized email addresses for the submitter (e.g., if you are uploading on behalf of someone else, aka “Submitter Proxy”, not to be confused with “internet proxy” under the `--proxy` option);

### ***What format should the location be in?***

The following are acceptable: `continent / country / region` and `continent / country / region / sub-region`, but; `country / region / sub-region`, `country / region` and `region / sub-region` are not. Hence, `Europe/Germany/Bavaria/Munich` and `Europe/Germany/Bavaria` are acceptable, but `Mexico / Jalisco / Puerto Vallarta` needs to be corrected to `North America / Mexico / Jalisco / Puerto Vallarta`. Spaces and hyphens in the location are allowed.

### ***How to correct submitted sequence or metadata?***

Update data via My Files within EpiCoV™. Or contact the curation team and provide the EPI\_ISL accession numbers and the information to be updated using a table (.xls, .csv or .tsv).

### ***How to add additional information to the location field?***

Use the `covv_add_location` field to add the extra data (e.g., Lat-Lon, suburb, postal code etc.).

### ***What does “date unplausable” mean in the error messages?***

This error can be caused by having a date value younger than the submission date (UTC). Adjust the problematic date in the `metadata.csv` file accordingly and attempt to resubmit.

### ***Upload was interrupted. How to re-start the upload?***

Simply re-run the last command. Samples already submitted will show a message indicating that the sample has already been assigned an accession, such as `WARNING; <SAMPLE NAME>; validation_error; covv_virus_name: already exists; existing_ids: ['<EXISTING EPI_ISL>']`