galaxy-upload

Release 1.0.1

The Galaxy Project

Jul 21, 2023

CONTENTS:

GALAXY-UPLOAD - GALAXY COMMAND-LINE UPLOAD UTILITY

A utility for uploading files to a Galaxy server from the command line. galaxy-upload supports Galaxy servers 22.01+, which include support for resumable uploads with the tus protocol.

1.1 Installation

1.1.1 Using pip

Python 3.7 or later is required.

To install:

```
$ pip install galaxy-upload
```

This installs two commands: `galaxy-upload`, used to upload file(s) to a Galaxy server, and `galaxy-history-search`, a helper utility for finding Galaxy histories to pass to the `galaxy-upload` command.

To make your life easier, you are encourged to install into a Python virtual environment. The easiest way to do this is with Python's built-in venv module:

```
$ python3 -m venv ~/galaxy-upload
$ . ~/galaxy-upload/bin/activate
$ pip install galaxy-upload
```

1.1.2 Using Conda

Alternatively, galaxy-upload can be installed using the Conda package manager. The galaxy-upload Conda package can be found on the bioconda channel and installed like so:

```
$ conda create -n galaxy-upload -c conda-forge -c bioconda galaxy-upload
$ conda activate galaxy-upload
```

1.1.3 Using Containers

It is also possible to run galaxy-upload in either a Docker or Singularity/Apptainer container. The galaxy-upload BioContainer is automatically built and maintained by the BioContainers project.

To use the Docker container:

```
$ docker run -it --rm -v "$(pwd):$(pwd)" -w "$(pwd)" -u "$(id -u):$(id -g)" \
    quay.io/biocontainers/galaxy-upload:1.0.0--pyhdfd78af_0 /bin/bash
```

Or as a single command without entering an interactive shell:

```
$ docker run --rm -v "$(pwd):$(pwd)" -w "$(pwd)" -u "$(id -u):$(id -g)" \
    quay.io/biocontainers/galaxy-upload:1.0.0--pyhdfd78af_0 galaxy-upload
```

Adjust the values of `-v` and `-w` according to where the data you want to upload are located. In the example above, it is assumed they are in the current working directory.

To use the Singularity container:

```
$ singularity run https://depot.galaxyproject.org/singularity/galaxy-upload:1.0.0--
→pyhdfd78af_0
```

Or as a single command without entering an interactive shell:

```
$ singularity run https://depot.galaxyproject.org/singularity/galaxy-upload:1.0.0--
→pyhdfd78af_0 galaxy-upload
```

Additional (newer) versions of the container may be available, BioContainers does not use the `latest` tag, but you can find all tags (which are valid for the Singularity images hosted on depot.galaxyproject.org as well as the Docker images) at the galaxy-upload quay.io page

1.2 Usage

Upload a pair of fastq.gz files:

```
$ galaxy-upload --url https://usegalaxy.org \
    --api-key 70ffeec0ffeea11e1eaccede1337loaf --history-name 'Run 2' \
    RUN2_F_001.fastq.gz RUN2_R_001.fastq.gz
RUN2_F_001.fastq.gz  100/100 mB ? eta 0:00:00
RUN2_R_001.fastq.gz  100/100 mB ? eta 0:00:00
```

Required arguments are the Galaxy server URL and API key, and a history ID or name. Your API key can be found in the Galaxy UI after logging in, by navigating to **User Preferences Manage API Key**.

You can set the URL and API key options as environment variables to avoid retyping and prevent the key from being visible to other users in `ps(1)` output:

```
$ export GALAXY_URL=https://usegalaxy.org
$ export GALAXY_API_KEY=70ffeec0ffeea11e1eaccede1337loaf
$ galaxy-upload --history-name stuff reads.bam
```

When selecting a history by name, regular expression matching is used. If the name matches multiple histories, `galaxy-upload` will exit will output details about the matched histories and then exit with an error. You can then select the correct history ID using the `--history-id` option:

```
$ galaxy-upload --history-name stuff reads.bam
                        Active Histories

 ID                  Name            Last Modified

  70ffeec0ffeead07 │ Newer stuff │ Mon Jul 11 15:54:05 2022
  a11e1eaccedeble8 │ Older stuff │ Wed May 25 18:03:46 2022

ERROR: Multiple histories matching stuff found! Use --history-id to select one.
$ galaxy-upload --history-id 70ffeec0ffeead07 reads.bam
reads.bam  3.6/3.6 gB ? eta 0:00:00
```

If you want to find the correct history without attempting an upload, use the `galaxy-history-search` command. The `--ignore-case` option can be used to perform a case-insensitive search:

```
$ galaxy-history-search --ignore-case trinity
                        Active Histories

 ID                  Name                Last Modified

  084649feb42d4295 │ : test trinity inputs │ Wed Dec  9 10:02:35 2020
  f697f94ca47080cf │ automate_trinity      │ Mon Dec 21 17:40:24 2015
  c79278c7a37e619e │ TrinityRun            │ Fri Mar 10 14:21:56 2017
  ee31286b26ff3352 │ trinity               │ Wed Sep 30 09:04:03 2020
```

Regular expressions are supported, for example, to find only the histories with names *ending* with `trinity`:

```
$ galaxy-history-search --ignore-case 'trinity$'
                        Active Histories

 ID                  Name                Last Modified

  f697f94ca47080cf │ automate_trinity │ Mon Dec 21 17:40:24 2015
  ee31286b26ff3352 │ trinity          │ Wed Sep 30 09:04:03 2020
```

Multiple options mirror those of the Galaxy UI's upload dialog, including `--file-type`, `--dbkey`, and `--space-to-tab`. The `--file-name` option can be used when uploading single files to control the file name in the history (by default it will be the same as the name on the local filesystem).

To support resuming interrupted uploads, use the `--storage` option to point to a state file (it will be created if it does not exist):

```
$ galaxy-upload --file-type bam --file-name Reads --storage /data/upload.txt /data/reads.
→bam
```

If the upload is interrupted, simply repeat the same command to resume uploading from the point of interruption.

Note that if you are trying to re-upload (not resume) a file that you have already uploaded once before, you will need to remove it from the storage file or use a different storage file.

HISTORY

2.1 1.0.1

- Document installation/usage with conda and containers by @natefoo in https://github.com/galaxyproject/galaxy-upload/pull/7

- Use click to require file arg to exist by @natefoo in https://github.com/galaxyproject/galaxy-upload/pull/9

- Pin tuspy to !=1.0.1 due to issue with upload_chunk(). by @natefoo in https://github.com/galaxyproject/galaxy-upload/pull/10

2.2 1.0.0

- Initial release

2.3 0.0.1

- Initial prerelease for comment and testing

2.4 Indices and tables

- genindex
- search