

MassMine: Using MassMine on UF's Research Computing Cloud Server

Login and Startup

Remote login to the UF cloud server setup specifically for MassMine research, through a Linux console on a computer with the Ubuntu OS installed. (UF provides training on accessing cloud resources through any operating system.)

```
Last login: Tue May 27 15:24:32 2014 from 75-22-119-19.lightspeed.wevloh.sbcglobal.net

Welcome to the UF Research Computing Center.

Do not run interactive jobs on the login nodes.  If you need to
run an interactive job, there are interactive/test nodes for that.

http://wiki.rc.ufl.edu/doc/Test_Nodes

UF Research Computing Center Account Policies can be found here:

http://www.rc.ufl.edu/about/policies/account
[aaronbev@gator4 ~]$
```

Startup Screen with basic text interface; users do not have to understand R code in order to collect data.

```
#####
##                                     ##
##                                     ##
##      MASSMINE                       ##
##                                     ##
##                                     ##
##      Your Access To Big Data         ##
##                                     ##
##                                     ##
#####

MassMine version 0.2.0 (2014-05-09)
https://github.com/n3mo/massmine

Copyright (C) 2014 Nicholas M. Van Horn & Aaron Beveridge
This program comes with ABSOLUTELY NO WARRANTY.
This is free software, and you are welcome to redistribute it
under certain conditions. Please see the included LICENSE file
for more information

Use previous Twitter credentials?
<Option> Yes
<Option> No
Choose =>

U:**- *R*          Bot L49      (iESS [R]: run MRev E1Doc)
```

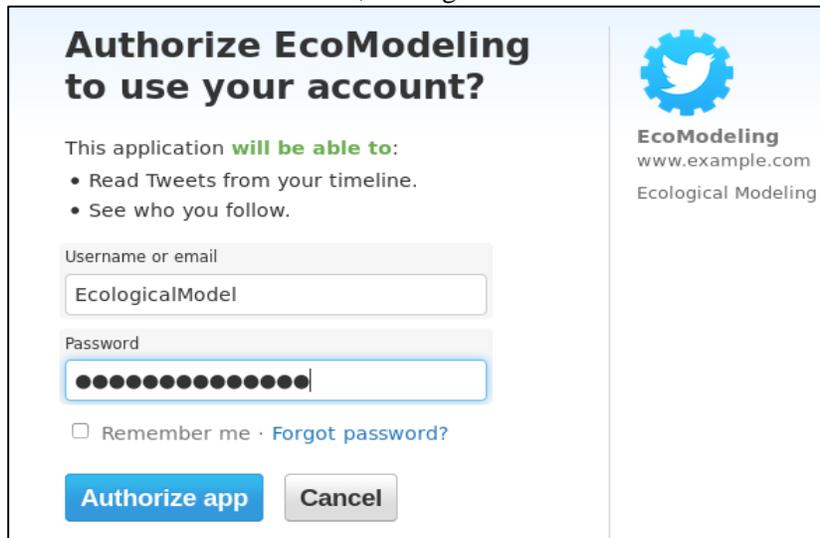
API Connection

MassMine checks the last configuration file and offers to re-authenticate API connection to restart a similar data collection.

```
Use previous Twitter credentials?
<Option> Yes
<Option> No
Choose => No

Please choose an account to authenticate:
<Option> EcoModeling
<Option> EcoModeling2
Choose => EcoModeling
To enable the connection, please direct your web browser to:
https://api.twitter.com/oauth/authorize?oauth_token=7yE8R0sbQ2Q1h0
UruStJzoyGBQ
When complete, record the PIN given to you and provide it here:
U:**- *R* Bot L126 (iESS [R]: run MRev E1Doc)
```

Based on the configuration file information, MassMine automatically displays the API account for authentication. Authentication only needs to happen once; after that, users can run the software each new time without re-authentication, as long as the same API credentials are used.



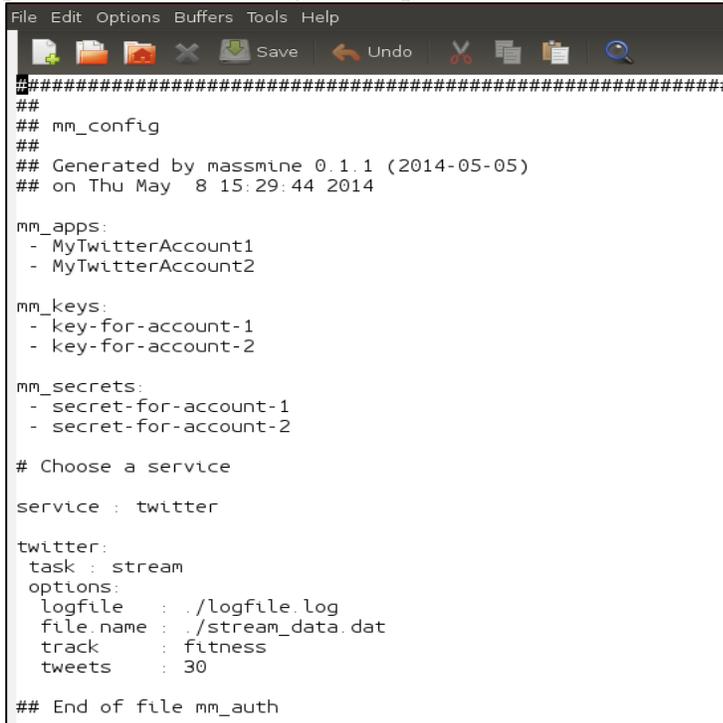
The image shows a web-based authorization form for the 'EcoModeling' application. The title is 'Authorize EcoModeling to use your account?'. Below the title, it states 'This application will be able to:' followed by two bullet points: 'Read Tweets from your timeline.' and 'See who you follow.'. To the right of the form is the application's logo (a blue Twitter bird) and the text 'EcoModeling', 'www.example.com', and 'Ecological Modeling'. The form includes a 'Username or email' field with the value 'EcologicalModel', a 'Password' field with masked characters, a 'Remember me' checkbox, and a 'Forgot password?' link. At the bottom, there are two buttons: 'Authorize app' (in blue) and 'Cancel' (in grey).

The API provides an authorization PIN, which syncs MassMine with the user's API access once the PIN is entered.

Configuration File

MassMine collects data based on the directions provided in the configuration file. The screenshot below shows the configuration file opened in a simple text editor. The configuration file is machine readable, and editable in any basic text editor on any operating system.

Users can save and re-process multiple configuration files to run different kinds of data collection activities the console application of MassMine. Templates of various basic configuration files are in process for use in trainings and experimentation for console application users.



```
File Edit Options Buffers Tools Help
#####
##
## mm_config
##
## Generated by massmine 0.1.1 (2014-05-05)
## on Thu May 8 15:29:44 2014

mm_apps:
- MyTwitterAccount1
- MyTwitterAccount2

mm_keys:
- key-for-account-1
- key-for-account-2

mm_secrets:
- secret-for-account-1
- secret-for-account-2

# Choose a service

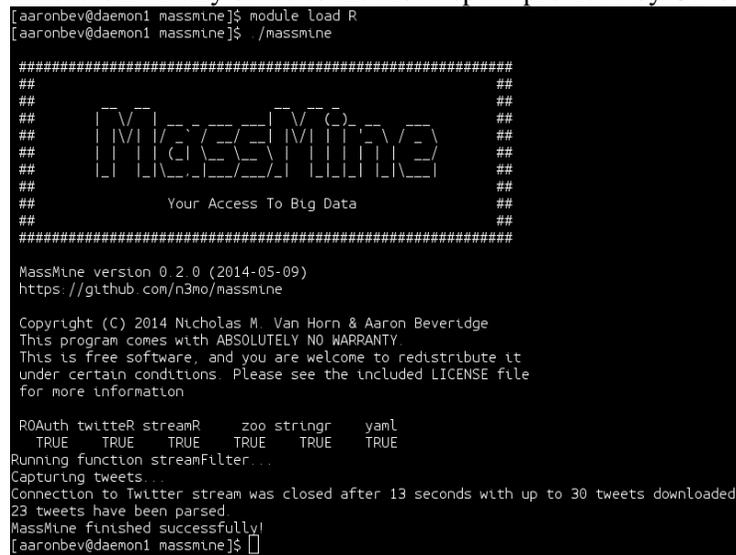
service : twitter

twitter:
task : stream
options:
logfile : ./logfile.log
file.name : ./stream_data.dat
track : fitness
tweets : 30

## End of file mm_auth
```

Success Screen

MassMine responds to let the user know when a collection finished successfully without error, with all data automatically saved to the cloud space provided by UF's Research Computing.



```
[aaronbev@daemon1 massmine]$ module load R
[aaronbev@daemon1 massmine]$ ./massmine

#####
##
##
##
##
##
##
##
##
##
##
##
##
##
##
##
#####

MassMine version 0.2.0 (2014-05-09)
https://github.com/n3mo/massmine

Copyright (C) 2014 Nicholas M. Van Horn & Aaron Beveridge
This program comes with ABSOLUTELY NO WARRANTY.
This is free software, and you are welcome to redistribute it
under certain conditions. Please see the included LICENSE file
for more information

ROAuth twitterR streamR zoo stringr yaml
TRUE TRUE TRUE TRUE TRUE TRUE
Running function streamFilter...
Capturing tweets
Connection to Twitter stream was closed after 13 seconds with up to 30 tweets downloaded.
23 tweets have been parsed
MassMine finished successfully!
[aaronbev@daemon1 massmine]$
```