Collaborative GARDIAN Labs

User Manual

Authors: Sotiris Konstantinidis (SCiO) & Pythagoras Karampiperis (SCiO)

Date: 22 January 2020

Contents

1. Introduction 4
2. How to use GARDIAN Labs 4
   2.1 Preparation/ Initial Steps 4
      2.1.1 Create Globus Account 4
      2.1.2 Install Globus Connect Personal 4
      2.1.3 Sign in to GARDIAN Labs 5
   2.2 Using Available Modules 5
      2.2.1 Globus Module 7
         2.2.1.1 Brief Description 7
         2.2.1.2 Available Actions 7
      2.2.2 GARDIAN Module 11
         2.2.2.1 Brief Description 11
         2.2.2.2 Available Actions 12
      2.2.3 Analytics Module 13
         2.2.3.1 Brief Description 13
         2.2.3.2 Available Actions 14

Figures

Figure 1: Visit your profile page 5
Figure 2: Edit your account..... 6
Figure 3: Choose the option “modules” 6
Figure 4: Enable the desired modules from the list of available modules 7
Figure 5: Upload data to CGLabs Globus server 7
Figure 6: Choose the desired endpoint at the Globus user interface 8
Figure 7: Add Directory CGLabs Globus Server 8
Figure 8: Name the directory that will be created CGLabs Globus Server 9
Figure 9: Transfer data from CGLabs Globus Server 9
Figure 10: Define the desired endpoint for transferring data from CGLabs Globus Server..... 10
Figure 11: Delete a file or folder..... 10
Figure 12: Confirm the deletion of a file or a folder 11
Figure 13: Search the GARDIAN index of datasets and publications 12
Figure 14: Presentation of found datasets from GARDIAN index 12
Figure 15: Save selected datasets to the GLab Globus Server 13
Figure 16: Connect to GARDIAN’s Jupyter Lab 14
Figure 17: Jupyter Lab home page 14

1. Introduction

Collaborative GARDIAN Labs (CG Labs) is the latest offering in the GARDIAN data ecosystem. CGLabs has a built-in collaboration platform, allowing a user to create either private or public virtual spaces, invite members, receive notifications and collaborate remotely and asynchronously. Access is handled by a Single Sign-On (SSO) functionality via Globus secure data software. CG Labs offers three modules offering specific, interlinked functionalities, allowing the user to:

(1) **Find data:** Search, download, and save datasets from GARDIAN in CG Labs;

(2) **Securely share data:** This may be published (or unpublished but sensitive data) securely exchanged via a Globus implementation; and

(3) **Analyze data:** Collaboratively write scripts and run analyses via Jupyter, which has been extended to support smooth data file exchange via the CG Labs Globus Server.

The scope of this document is to provide brief guidelines on how to use CGLabs. The document starts with some preparatory steps that users should follow the first time they start using CGLabs. These actions include: (a) the creation of a Globus account (see section 2.1.1); (b) the installation of Globus Connect Personal (see section 2.1.2); and (c) the signing-in to CGLabs.

After describing these preparatory steps, the document presents the CGLabs modules developed and integrated as of Spring 2020, and details how they can be accessed and used.

2. How to use Collaborative GARDIAN Labs

2.1 Preparation/ Initial Steps

2.1.1 Create Globus Account

A Globus account is required to register and get access to CGLabs. To create a Globus Account, please visit https://www.globus.org/ and follow the instructions.

2.1.2 Install Globus Connect Personal

Globus Connect Personal turns your laptop or other personal computer into a Globus endpoint with just a few clicks. With Globus Connect Personal you can share and transfer files to/from a local machine—campus server, desktop computer or laptop—even if it's behind a firewall, and you don't have administrator privileges. A Globus Connect Personal is **required** to make full use of the CGLabs offerings.

Globus Connect Personal is available for all major operating systems. Please click on the links below for installation instructions.

- Globus Connect Personal for Mac
  https://docs.globus.org/how-to/globus-connect-personal-mac

- Globus Connect Personal for Linux
  https://docs.globus.org/how-to/globus-connect-personal-linux
- Globus Connect Personal for Windows
  https://docs.globus.org/how-to/globus-connect-personal-windows

Additional information is available in the Globus Connect frequently asked questions: https://docs.globus.org/faq/globus-connect-endpoints/

2.1.3 Sign in to CGLabs

Once you have a Globus account, the next step is to visit CGLabs (https://labs.scio.systems/index.php/user/auth/login) and sign in. The first time you sign in, CGLabs will ask for basic registration details.

2.2 Using Available Modules

Users can enable the desired available modules with the following steps:

STEP 1 Visit your Profile page

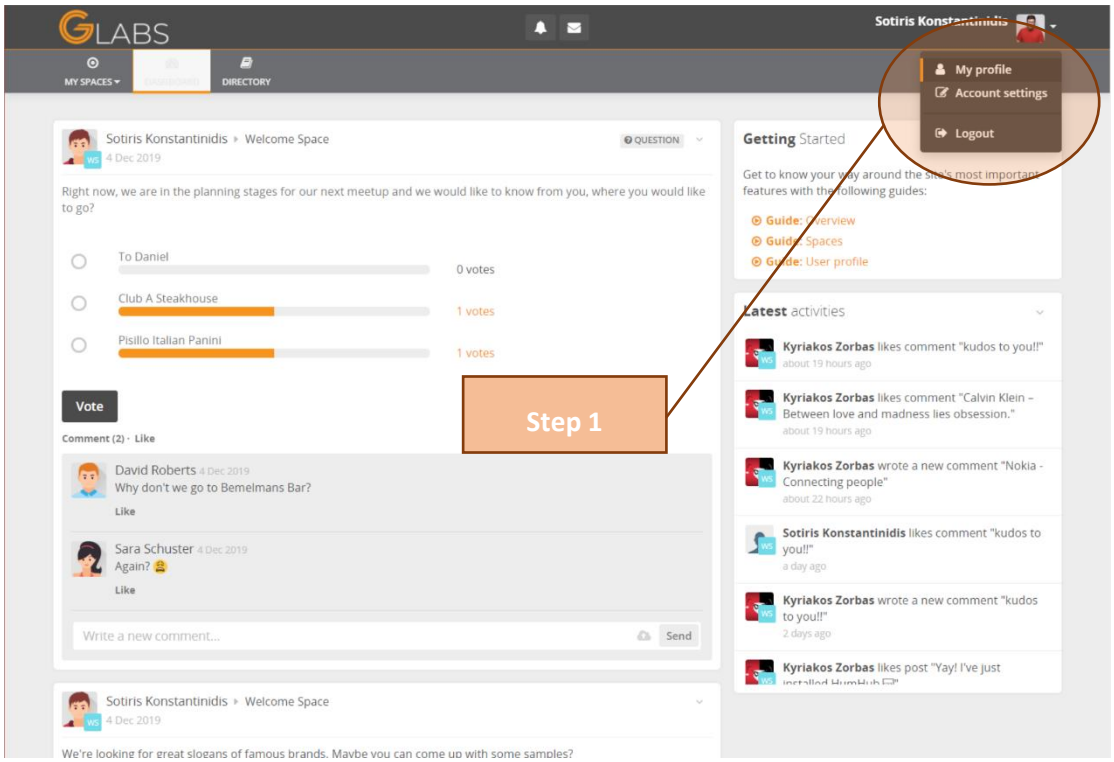

Figure 1: Visit your profile page

STEP 2 Choose Edit Account

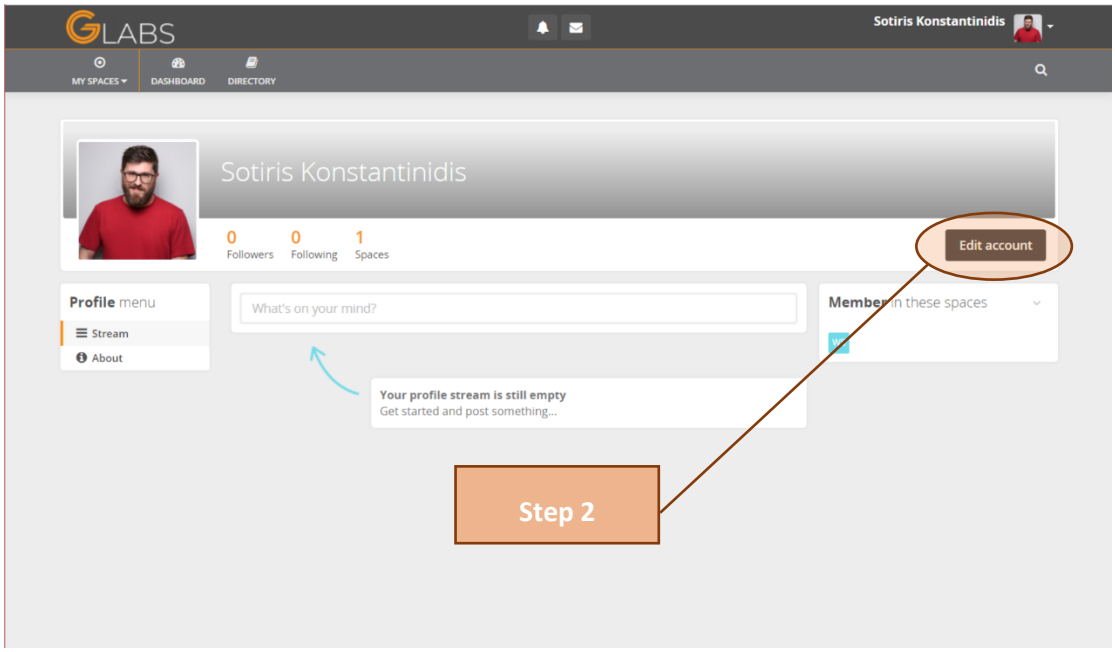

Figure 2: Edit your account

STEP 3 Choose Modules

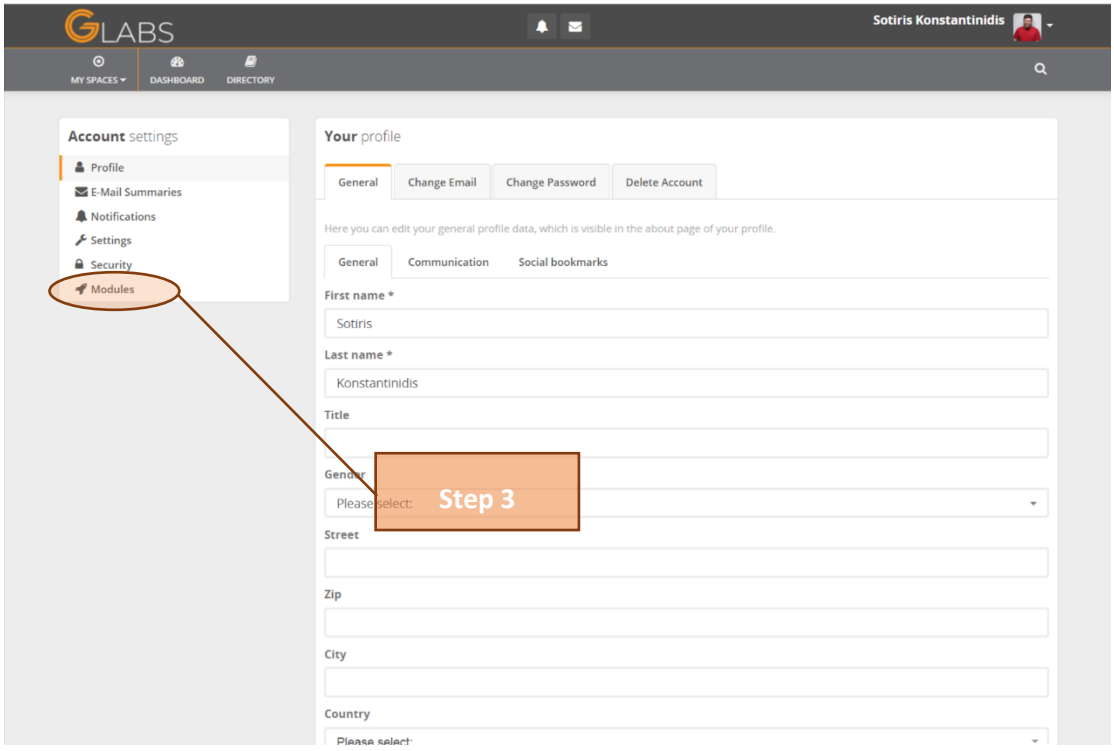

Figure 3: Choose the option "modules"

STEP 4 Click Enable for the desired module

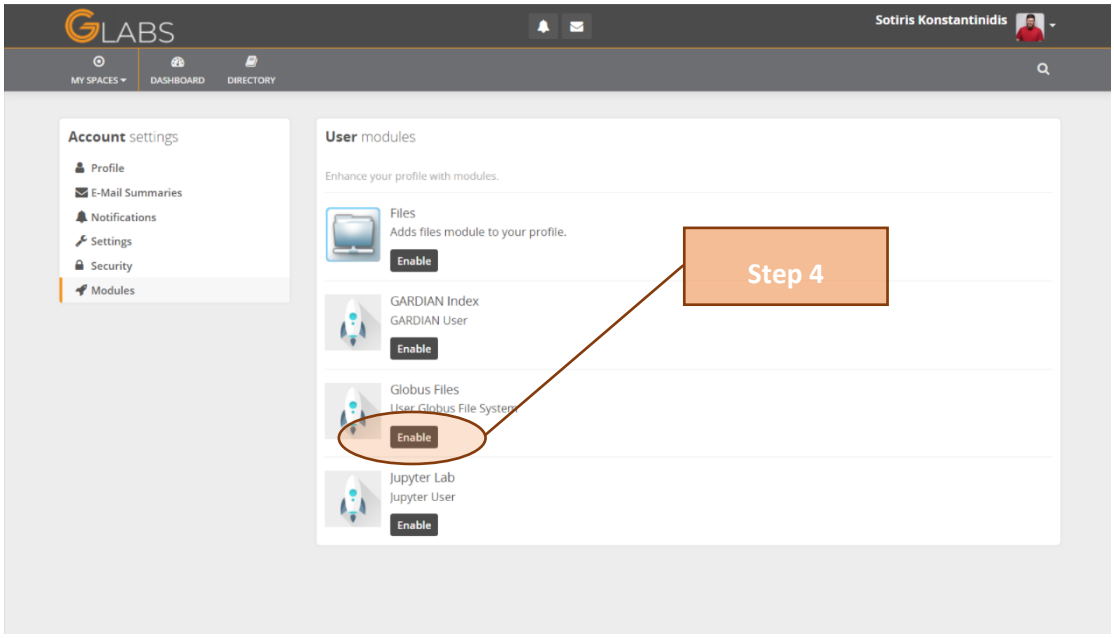

Figure 4: Enable the desired modules from the list of available modules

2.2.1 Globus Module

2.2.1.1 Brief Description

The Globus Module allows the secure & reliable upload of data from any Globus endpoint to the CGLabs Globus Server. You can upload your data, create directories and transfer data from CGLabs to any Globus Endpoint.

After enabling the module, you can access it by visiting your profile page. Under the Profile menu section click the Globus option.

2.2.1.2 Available Actions

Secure Upload: Upload your data to CGLabs Globus Server from any Globus Endpoint.

STEP 1 Click on the Secure Upload button to get to the Globus User Interface.

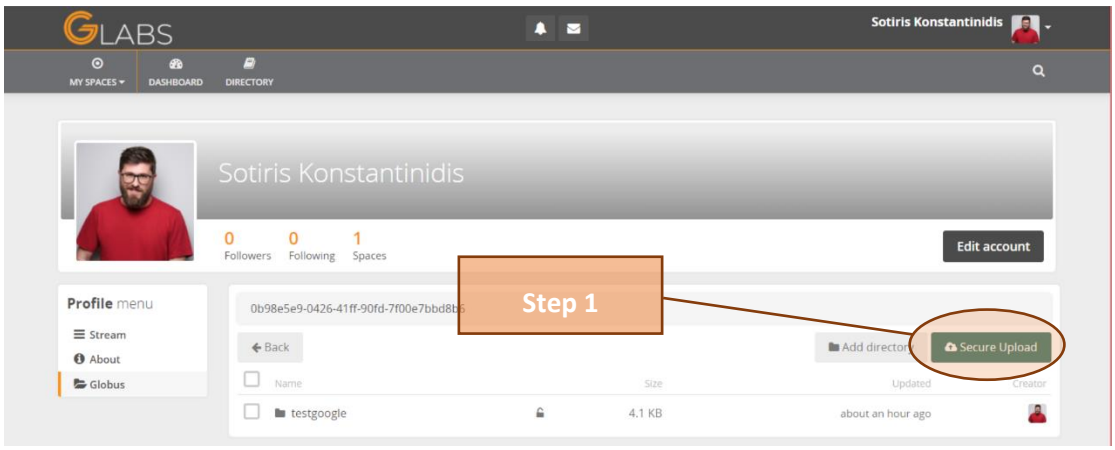

Figure 5: Upload data to CGLabs Globus server

STEP 2 From the Globus User Interface you can choose the desired endpoint and the corresponding data.

STEP 3 After clicking submit the user will be redirected to CGLabs. As soon as the transfer is completed the user will be able to see their data through CGLabs.

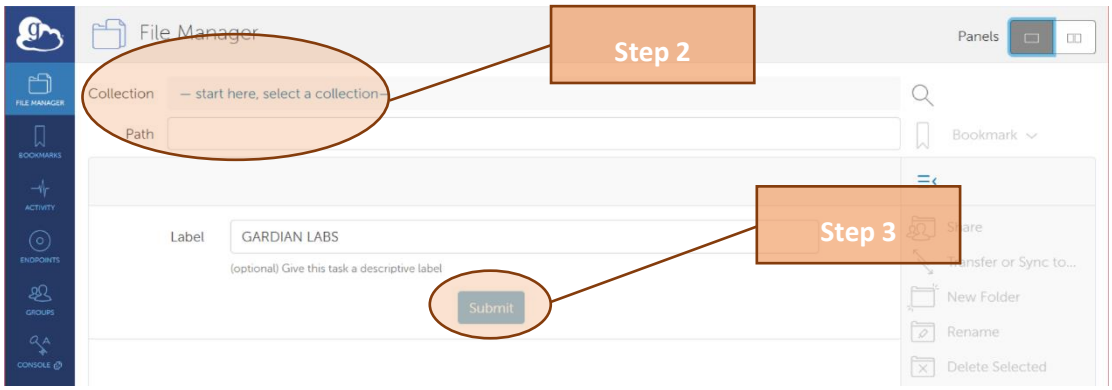

Figure 6: Choose the desired endpoint at the Globus user interface

Add Directory: Create a directory to CGLabs Globus Server.

STEP 1 Click on the Add Directory button and a modal panel will appear.

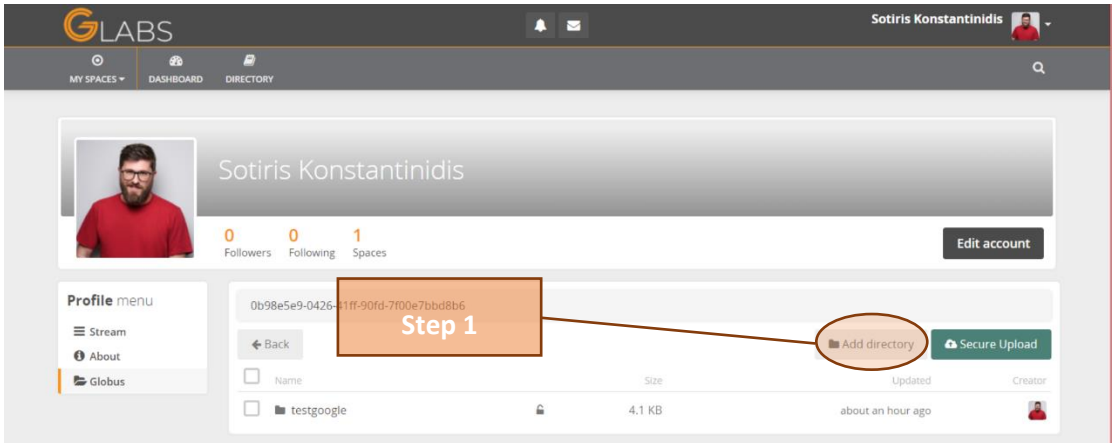

Figure 7: Add Directory CGLabs Globus Server

STEP 2 On the modal panel you can provide the desired folder name.

STEP 3 Click Save to create the directory.

Figure 8: Name the directory that will be created CGLabs Globus Server

Transfer Data: Transfer data from the CGLabs Globus Server endpoint to any Globus endpoint.

STEP 1 Make a right click to the desired file or folder and click on the Transfer option.

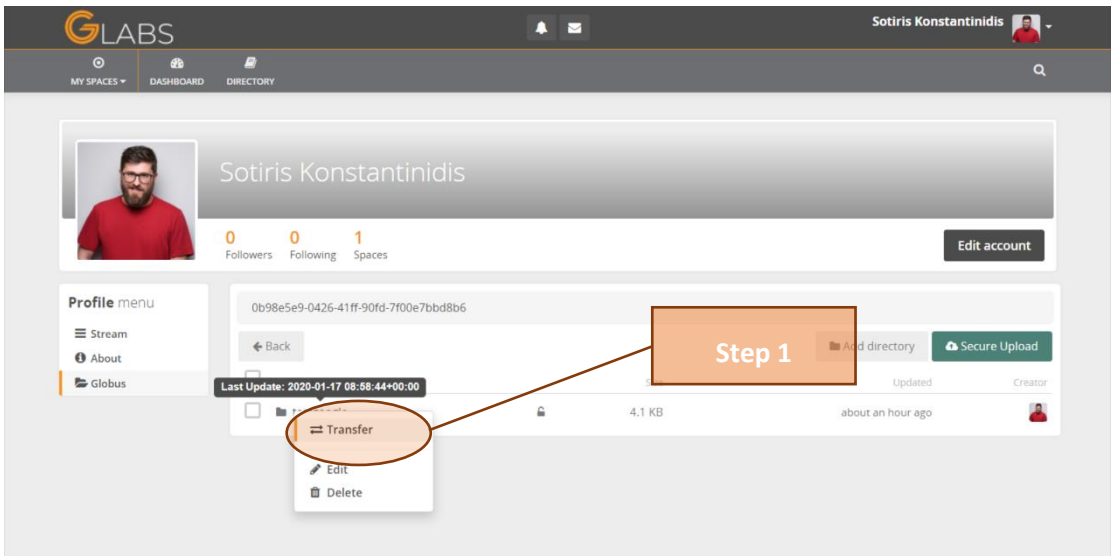

Figure 9: Transfer data from CGLabs Globus Server

STEP 2 On the modal panel you can provide the desired endpoint, then click Transfer.

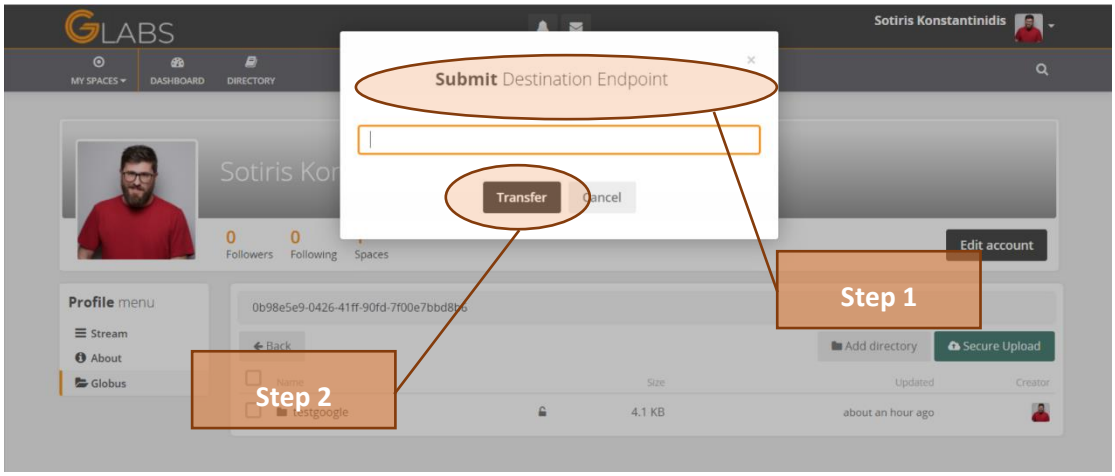

Figure 10: Define the desired endpoint for transferring data from CGLabs Globus Server

Delete Data: Delete any desired file or folder

STEP 1 Make a right click on the desired file or folder and click on the Delete option.

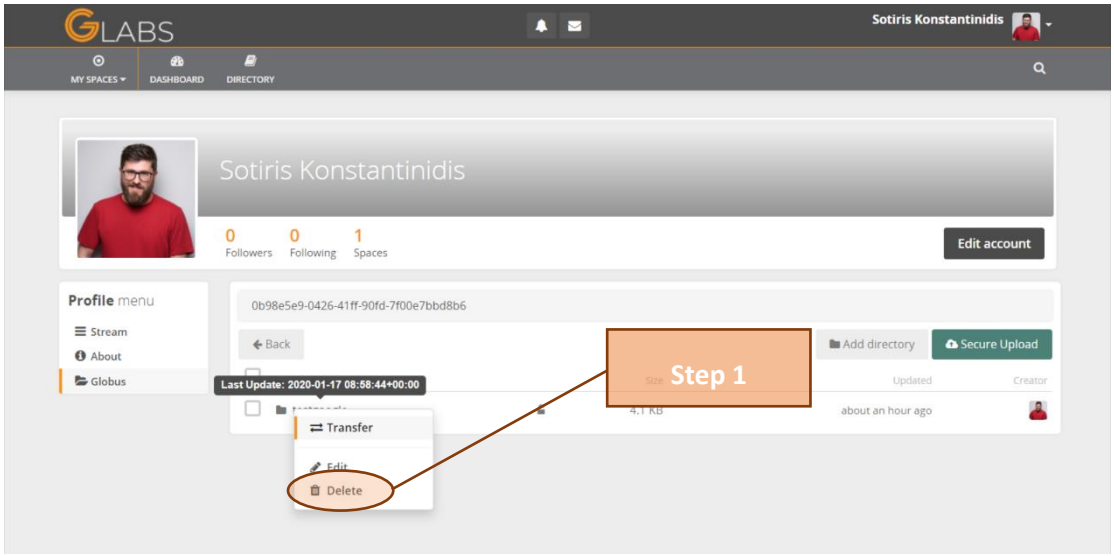

Figure 11: Delete a file or folder

STEP 2 On the modal panel confirm action.

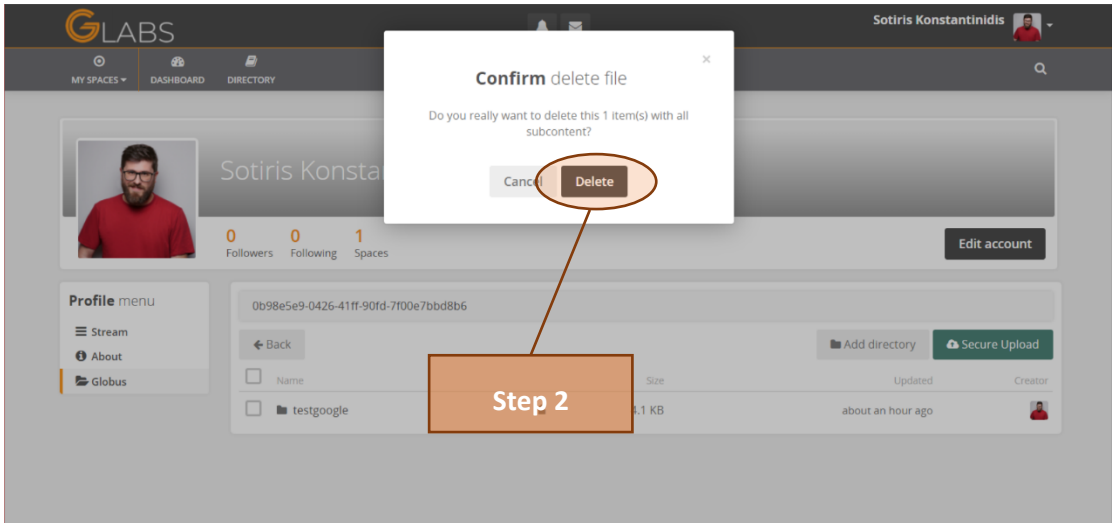

Figure 12: Confirm the deletion of a file or a folder

2.2.2 GARDIAN Module

2.2.2.1 Brief Description

The GARDIAN module allows users to perform keyword-based queries to the GARDIAN data and publications index and find resources. In the case of datasets users can choose and save the data to CGLabs Globus Server.

After enabling the module, users can access it by visiting their profile page. In the Profile menu section, click the GARDIAN option.

2.2.2.2 Available Actions

Search Resources: You can search for publication and datasets indexed by GARDIAN.

STEP 1 Type in the desired keywords in the search box.

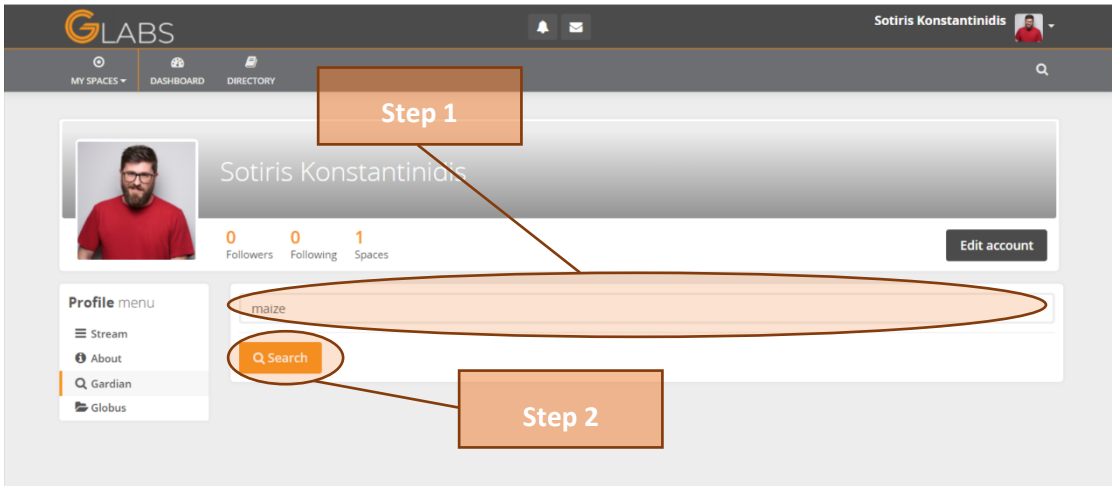

Figure 13: Search the GARDIAN index of datasets and publications

STEP 2 You can navigate to the datasets and publications you find.

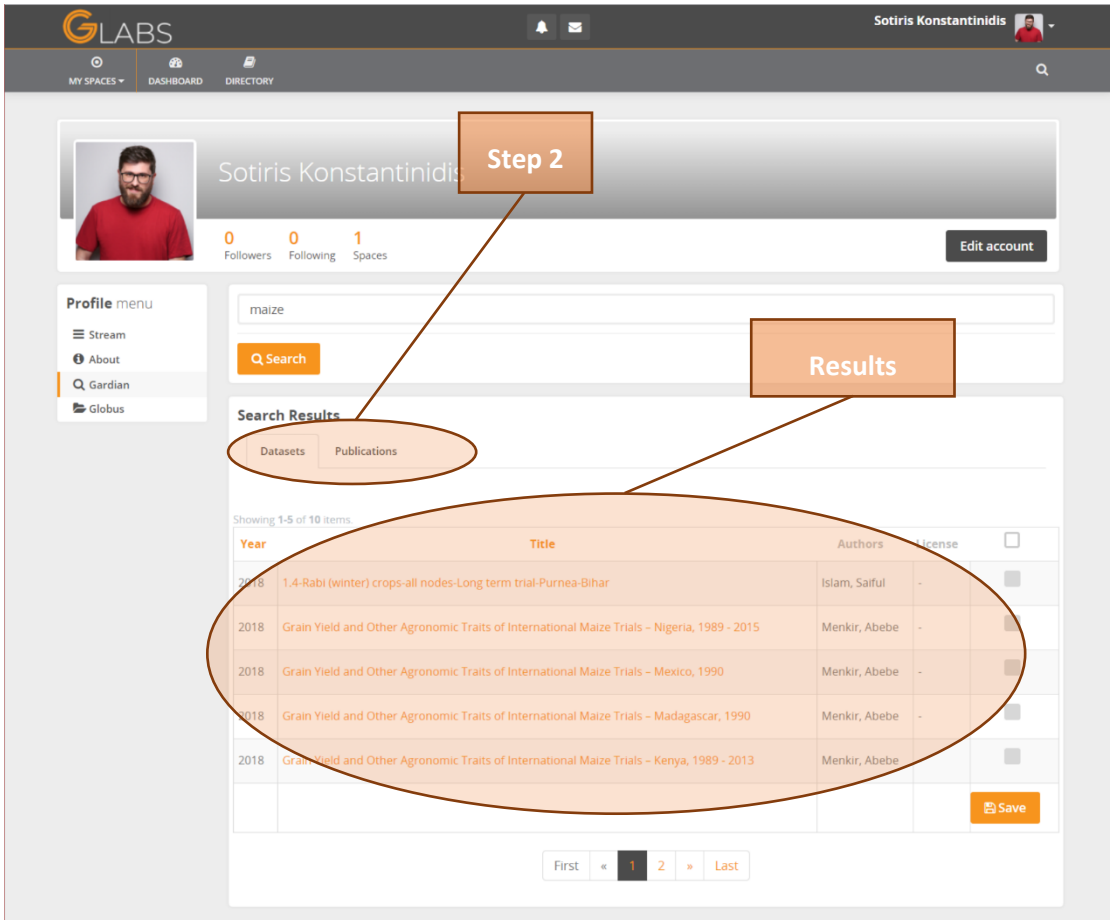

Figure 14: Presentation of found datasets from GARDIAN index

Choose & Save Datasets to the CGLabs Globus Server

STEP 1 Choose the desired datasets. Datasets that do not have appropriate licenses cannot be chosen.

STEP 2 Click Save and CGLabs will download and save the datasets to CGLabs Globus Server.

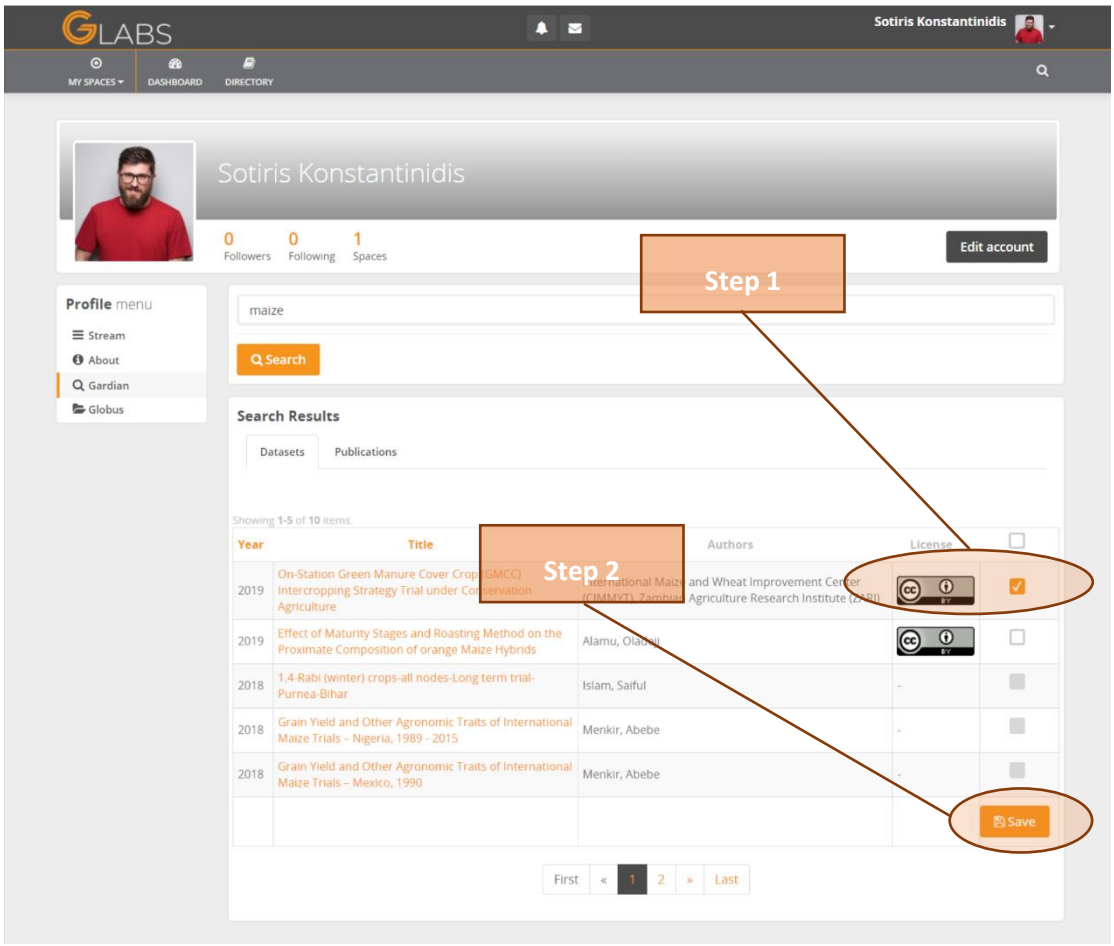

Figure 15: Save selected datasets to the GLab Globus Server

2.2.3 Analytics Module

2.2.3.1 Brief Description

The analytics module allows users to write and execute source code written in Python or R on top of CGLabs infrastructure. The module initiates a Jupyter instance that is integrated with CGLabs Globus Server, allowing users to perform data analytics.

After enabling the module, users can access it by visiting their profile page. In the Profile menu section, click the Analytics option.

2.2.3.2 Available Actions

Connect to Jupyter

STEP 1 Click the Connect button to initiate GARDIAN's Jupyter Lab.

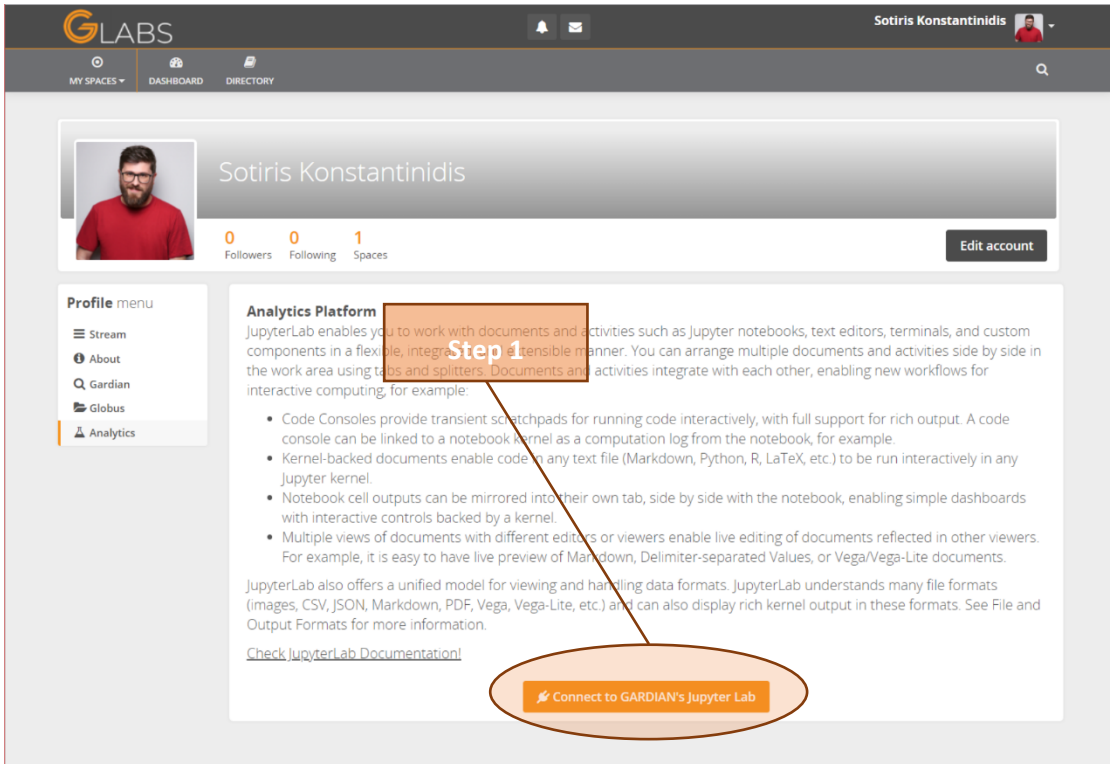

Figure 16: Connect to GARDIAN's Jupyter Lab

STEP 2 You will be redirected to Jupyter Lab, where you can execute your experiments in either python or R.

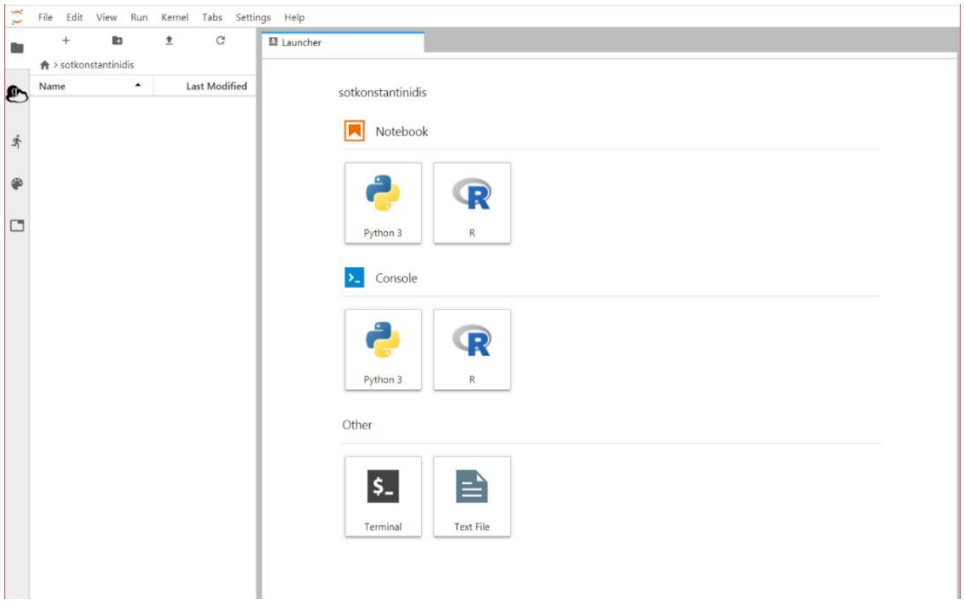

Figure 17: Jupyter Lab home page