

Data Depot File Uploading Instructions for EMBRIO

Scott Bolton

EMBRIO Institute

v. 2025-11-13

1. Uploading bulk raw data to Data Depot

The EMBRIO Data Depot is called “/depot/embrio”, which is a subfolder of the Data Depot and has its own storage quotas and group management credentials. The current size of the depot is 4 TB and can be increased as needed by request through Purdue RCAC.

This depot is viewed through Globus: a web-based data transfer tool. The repository is here:

<http://transfer.rcac.purdue.edu/>

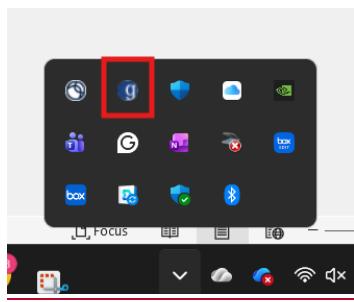
NOTE: this requires Purdue career account login credentials.

Transferring files into the Data Depot requires a background system application running on your computer called Globus Connect Personal, which connects your local filesystem to Globus through the web interface. It is downloaded here:

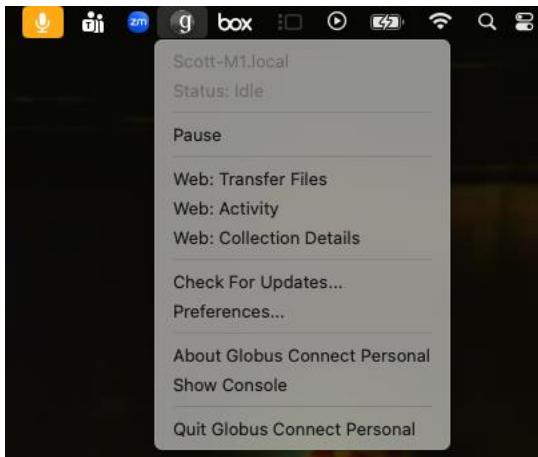
<https://www.globus.org/globus-connect-personal>

Follow the instructions on the Globus website to download, install, and connect to Globus for your particular operating system. You should see the “g” icon loaded in your “system tray” or list of background apps.

Windows:



MacOS:



TODO: HOW TO SET UP GUEST COLLECTION AT INSTALLATION, BEATRIZ EXAMPLE

2. Configure the file transfer windows

Launch the Globus Connect Personal application and verify it is running in the background.

Next, log into the Globus website <http://transfer.rcac.purdue.edu/>

The Panels icons at the upper right side of the window can be configured to show a left panel, right panel, or both panels simultaneously. The transfer process is performed by choosing a source location, a destination location, and then a transfer of selected files.

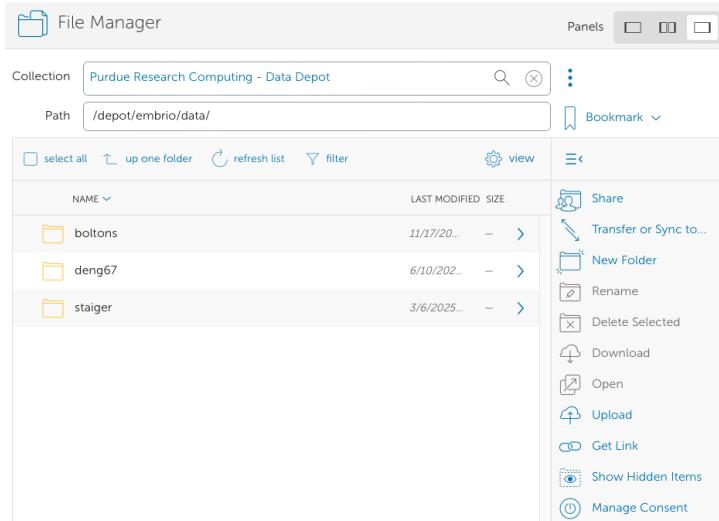
To select the destination location, select the right panel icon, then click on the “Search” text in the Collection field. This brings up a Collection Search window. Type “data depot” to search for the Data Depot collection. Several choices will be presented:

1. Select Purdue Research Computing – Data Depot.

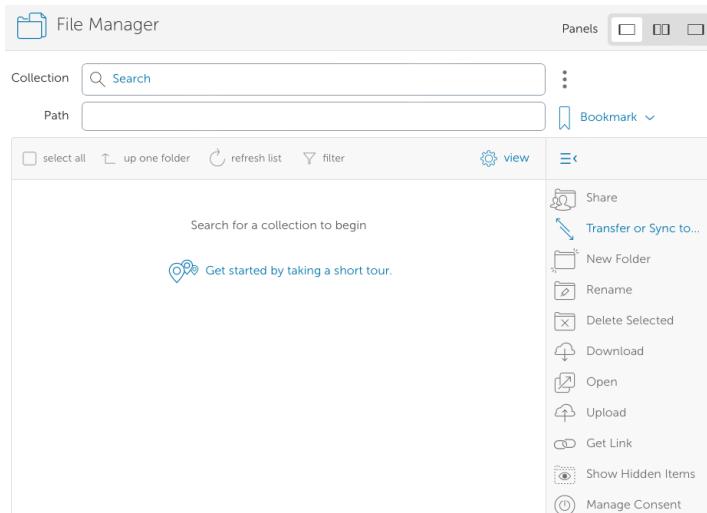
TODO SNAPSHOT: FIRST TIME USER NOT AUTHORIZED. Group Manager of /data/embrio must add username to group.

2. The path will now show “/depot/”. Navigate to the subfolder “embrio”:

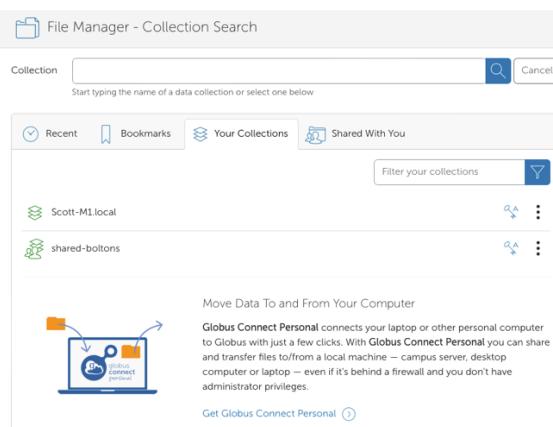
3. These folders can be organized as the EMBRIO group requires, but data storage will be in the data subdirectory, organized by lab. For example:



4. Create subfolders as desired to hold the data to be uploaded.
5. Now we set the source location by clicking on the left panel icon in the upper right of the window.

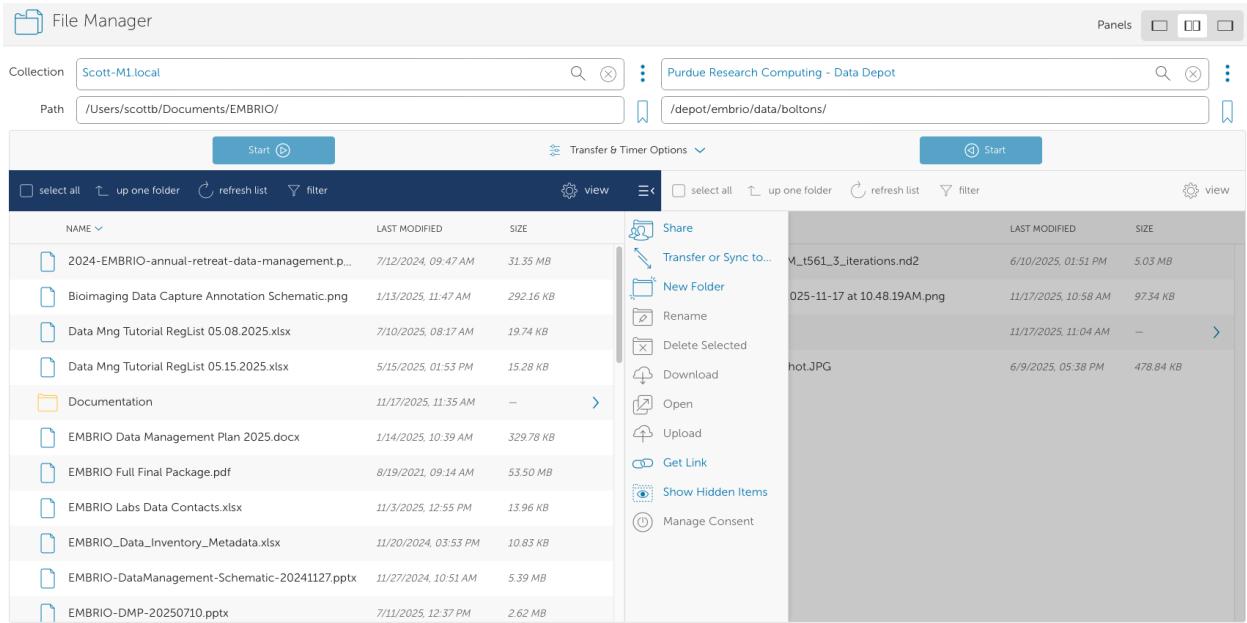


Click on the “Search” text in the Collection field. This brings up a Collection Search window. Select the “Your Collections” tab:

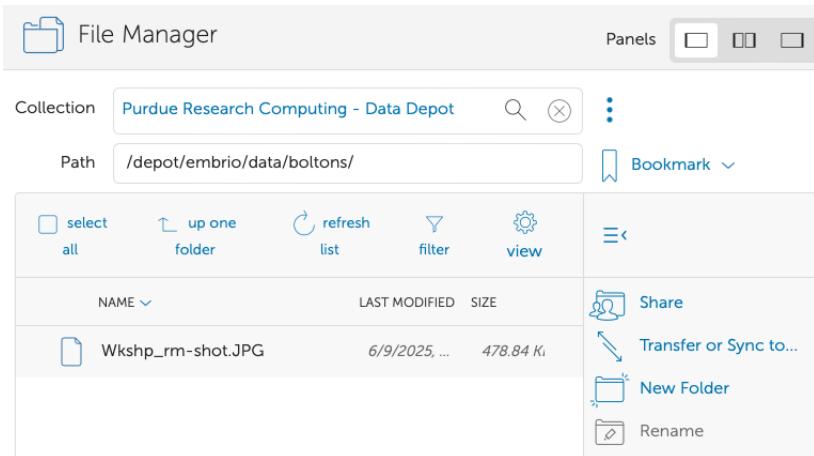


In this example, the local computer is “Scott-M1” and the computer’s filesystem is mounted at “Scott-M1.local”. Choose this location, then navigate to the location of the data to be uploaded.

- Now click the dual panel icon in the upper right window to show both source and destination panels.



- Select the folder or data file(s) in the left panel to be transferred to /depot/embrio/data, then click the green “Start Button” above the panel to begin the transfer. Depending on the data size this can take a significant amount of time. Globus will send an email to your account once the transfer has been completed.
- Verify the transfer by clicking on the refresh button on the right side.
- Set permissions for the data to be shared. Select the right single panel view again, select the data by checking boxes, then click the “Share” button in the side bar popover menu:



- Next, select the Guest Collection. For each URL locator for a set of data, a new Guest Collection mount point is needed, which can be done by clicking the “Add Guest Collection” button:

 Create new Guest Collection

You are creating a guest collection on **Purdue Research Computing - Data Depot** to share data

User Credential **Scott Charles Bolton (boltons)**

Globus Identity **Scott Charles Bolton (boltons@purdue.edu)**

Directory   

Display Name

Description

Keywords

Information Link 

Contact Email 

Organization

Department

Other Contact Info

Default Directory   

This is relative to the Directory you set above. If this field is left empty it will default to / (i.e. the Directory itself)

Force encryption on transfers to and from this collection

11. Upon clicking “Create Collection”, the share point is created and this window appears:

 New Test Data Share

 Overview  Permissions  Roles

Shared With

Mapped Collection: Purdue Research Computing - Data Depot
Base Path: /depot/embrio/data/boltons/Test/



USER OR GROUP	CREATED	EXPIRATION	READ	WRITE
Permissions granted by role				
Scott Charles Bolton (boltons@purdue.edu)	-	never expires	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
Purdue Data Depot Endpoint (d6a36558-6489-4ded-8ceb-71194bc6e58a@clients.auth.globus.org)	-	never expires	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>

12. Select the Add Permissions – Share With button, then select the icon to add sharing:

Path: /

Share With: user - share with specific individuals
 group - make data accessible to members of a group
 all users - make data accessible to all logged in users of Globus
 public (anonymous) - make data accessible to everyone

Permissions: read write

13. Select the “public (anonymous)” radio button, then click “Add Permission”:

Path: /

Share With: user - share with specific individuals
 group - make data accessible to members of a group
 all users - make data accessible to all logged in users of Globus
 public (anonymous) - make data accessible to everyone

Permissions: read write

14. The data is now publicly accessible via the link found by clicking “Show link for sharing”.

New Test Data Share

Overview Permissions Roles

Shared With

Mapped Collection: Purdue Research Computing - Data Depot
 Base Path: /depot/embrio/data/boltons/Test/

USER OR GROUP	CREATED	EXPIRATION	READ	WRITE
> Permissions granted by role				
Path: /		11/17/2025, 03:22 PM	never expires	<input checked="" type="checkbox"/> <input type="checkbox"/>
Public				<input type="button" value=""/>

It should have this format:

https://transfer.rcac.purdue.edu/file-manager?origin_id=...&origin_path=%2F

15. Copy this link to the clipboard and use it for the PURR record Raw Data value.