How to Use Dataverse Globus Integration

Globus Connect Personal allows you to make any path on a Linux, Windows, or Mac device an endpoint in Globus that you can transfer files to/from. The Globus integration application in Dataverse allows you to use those personal endpoints to transfer files directly to a specific dataset in Dataverse. The file transfer is handled by Globus, which handles all the communication between Dataverse and the S3 bucket and removes the load of resources from the Dataverse application server.

In addition, the Globus transfer should work with any endpoint where your account can start a transfer. If you can transfer from the endpoint to your personal endpoint, you should also be able to transfer from that endpoint to the Dataverse S3 connector endpoint. For example, if you have a home directory on the Agave cluster, you should be able to transfer from there, but you will need to create a guest collection in Globus for your directory first.

To get started with transferring from local endpoints, please use the following instructions.

1. Install Globus Connect Personal on whatever device/platform you would like to transfer files to/from Dataverse. This process will allow you to configure that device as a Globus endpoint that will be available to Dataverse. For example, your local hard drive or a folder on it.

- Windows https://docs.globus.org/how-to/globus-connect-personal-windows
- Linux https://docs.globus.org/how-to/globus-connect-personal-linux
- Mac https://docs.globus.org/how-to/globus-connect-personal-mac

2. In order for the Globus application to load, you must have at least one instance of Globus Connect Personal running.

3. Go to the ASU Library Research Data Repository (QA) at https://dataverse-qa.lib.asu.edu/ and login using the institutional login.

4. Access a dataset that you have access to upload files to by clicking the link to the name of the dataset.

5. Click on the "Upload Files" button; then click "Upload with Globus".

6. The Globus app that integrates Globus endpoints to Dataverse should appear

7. Click the down arrow to see available personal endpoints. Click one to open it in the file manager.
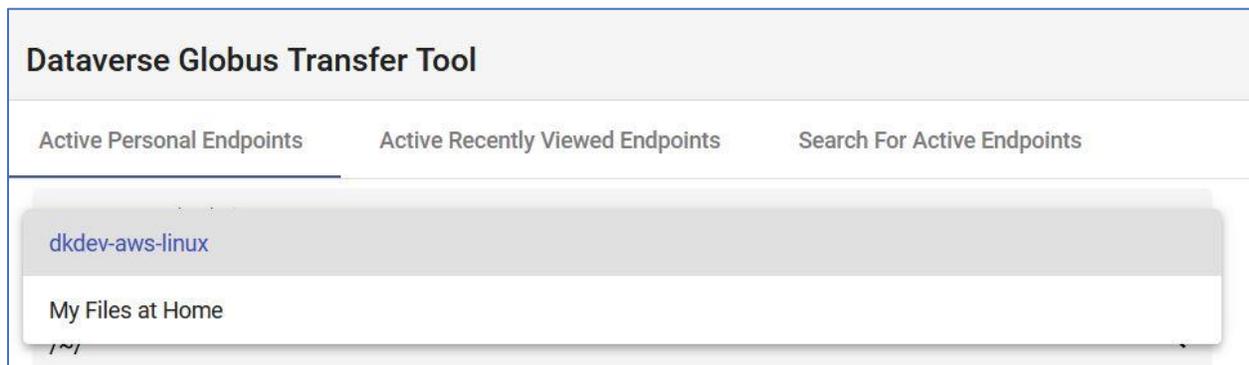
8. From the list of files in your endpoint, click on individual files to select the ones you would like to upload; the selected files will appear on the right of the screen that will be transferred.
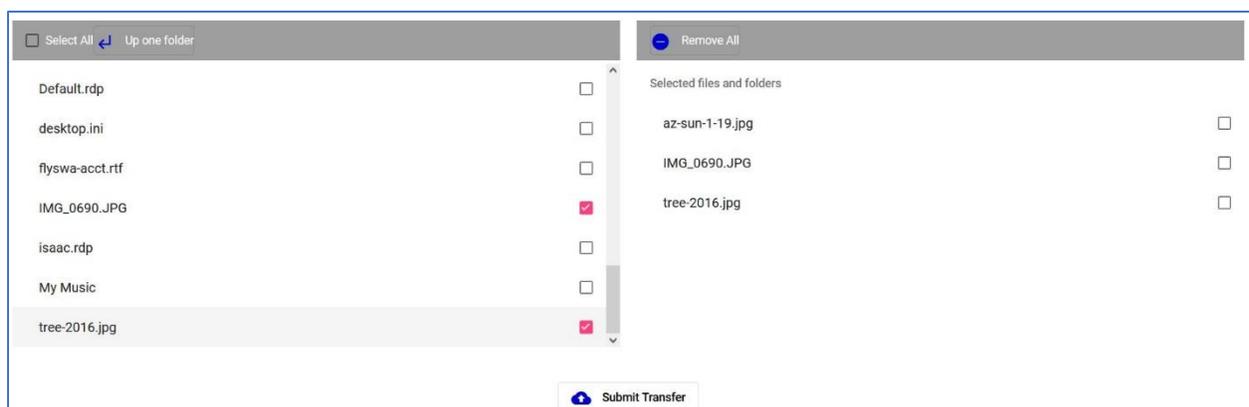
9. Click the "Submit Transfer" link at the bottom of the screen; a message will appear saying that the transfer is pending and then that it has been initiated.

10. Go back to the Dataverse console and **refresh the page** on your dataset; you should see a yellow bar indicating that the Globus transfer has been initiated; after about 5 minutes you should see a message that the dataset has an Edit in Progress.

11. After a few minutes, the yellow message will disappear, and the dataset will show the new file(s) (you may have to refresh the screen again).