

Northwestern

Northwestern
INFORMATION TECHNOLOGY

Globus in Action: Introduction to Globus for Research Data Management

Brian Roland

Data Management Specialist

Northwestern IT

Research Computing and Data Services

Agenda

- What is Globus?
- Why use Globus?
- How to get started with Globus
- Live Demo
 - Globus Collections
 - Setting up a Globus transfer
 - Transfer status/history
 - Data Sharing in Globus
 - Setting up Globus Connect Personal
- Advanced Globus Features

What is Globus?

Globus is a powerful data management, transfer, and collaboration tool designed to simplify research workflows and empower users to securely manage and share large amounts of data from a variety of data storage platforms.

Why use Globus?



Used by research institutions and government facilities across the globe.

2600+ Connected Institutions



Makes data storage usage more accessible to researchers. Allows researchers to efficiently manage usage on a variety of storage systems.

Institutional Storage

Local Storage

Commercial cloud storage platforms



Developers can use Globus to automate data workflows leveraging the Globus API, SDK, and command line tools (Advanced)

Globus Features

Globus enables users to transfer and manage data between endpoints using advanced protocols that provide users with:

- High performance data transfers optimized via GFTP
- Automatic fault recovery and error checking
 - Fire and forget functionality
- Encryption in transit (HTTPS)
- Post-transfer file integrity verification

Globus Features

Additional key features of Globus provides include:

- Web based user interface
- Downloadable client that enables transfers to local machine storage (Globus Connect Personal)
- Automation and scheduling functionality
- Data Sharing capabilities

Getting started with Globus

Globus uses your netID and Password for authentication via Northwestern's Single Sign-On Service (SSO)

Use of Globus at Northwestern University can only be used with data that is NOT legally or contractually restricted

Use of Globus with RDSS requires user to complete the [Request Access to RDSS via Globus form](#).

<https://www.globus.org/>



The screenshot shows the Globus website homepage. The header features the Globus logo (a stylized 'g' in a cloud) and the text 'globus a uChicago non-profit service'. To the right of the logo are two buttons: 'GET STARTED' with a rocket icon and 'LOG IN' with a login icon, the latter of which is highlighted with a yellow circle. Below the header is a navigation bar with links: 'Solutions', 'Resources', 'Pricing', 'Newsroom', 'Developers', and 'About'. The main content area has a dark blue background with a network diagram. On the left, a central figure labeled 'On Demand Computing Resources' is connected to five other nodes: 'leadership class computing', 'commercial computing', 'research computing', 'institutional computing', and 'personal computing'. To the right of the diagram, the text reads: 'Go beyond data', 'Globus Compute', 'Reliable, distributed Function-as-a-Service', and 'COMPUTE ANYWHERE: EDGE TO SUPERCOMPUTER' with a right-pointing arrow icon. At the bottom of the page, the text 'Research IT. Reimagined.' is displayed.

globus
a uChicago non-profit service

GET STARTED

LOG IN

Solutions ▾ Resources ▾ Pricing ▾ Newsroom ▾ Developers ▾ About ▾

On Demand Computing Resources

leadership class computing

commercial computing

research computing

institutional computing

personal computing

Go beyond data

Globus Compute

Reliable, distributed Function-as-a-Service

COMPUTE ANYWHERE: EDGE TO SUPERCOMPUTER >

Research IT. Reimagined.

Logging in via NU Credentials



Log in to use Globus Web App

Use your existing organizational login

e.g., university, national lab, facility, project



Northwestern University

North Carolina State University

North Dakota State University Main Campus

Northeastern University

Northern Arizona University

Northern Illinois University

Northern Michigan University

Northwestern University

CILogon and Globus. You also agree for CILogon to issue a certificate that allows Globus to act on your behalf.

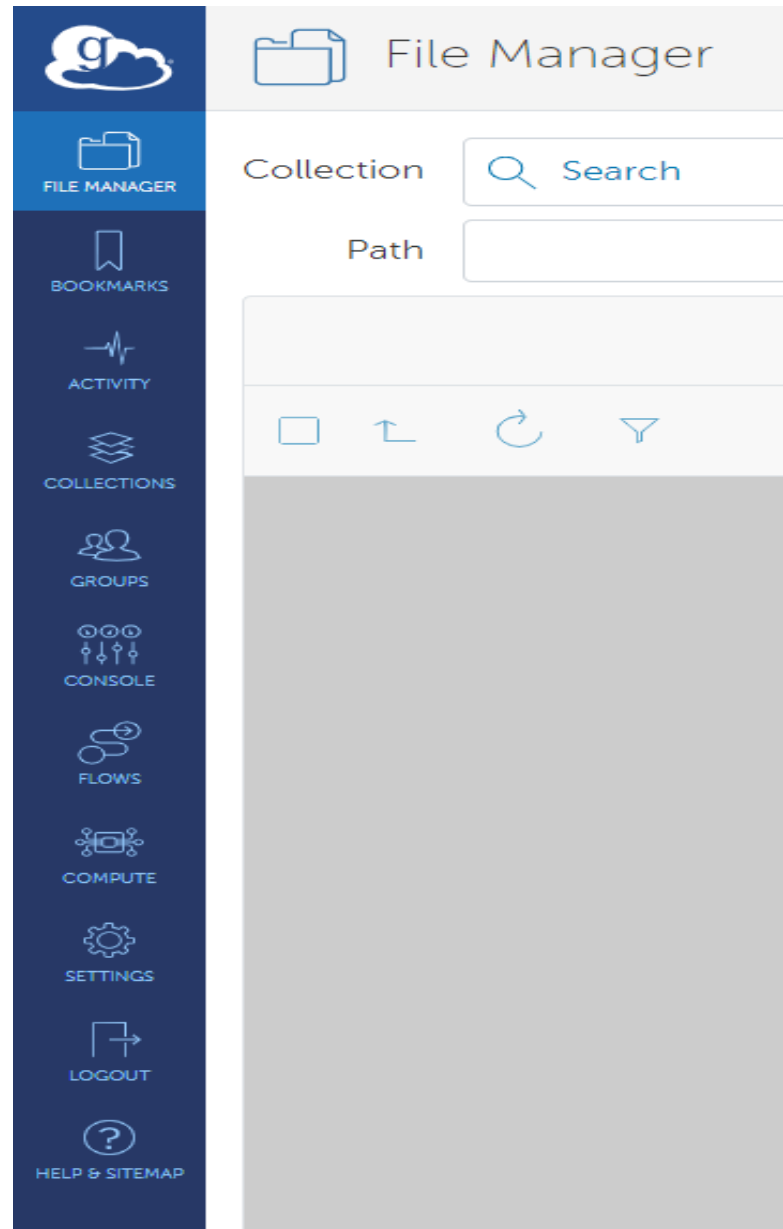
OR



Didn't find your organization? Then use [Globus ID to sign in](#). ([What's this?](#))

Navigation Menu

- **File Manager** is where you will setup and submit globus transfers
- **Bookmarks** is where you can find and manage shortcuts you to various Collections frequently accessed
- **Activity** is where users can find the status of current transfers and a history of recently completed transfers.
- **Collections** is where you can search information on publicly listed Globus Collections.
- **Groups** is where you can create, manage, or request access to Globus groups that help manage permissions on Globus Collections
- **Settings** is where you manage your Globus account including consents, Globus Group memberships, and the developer console.

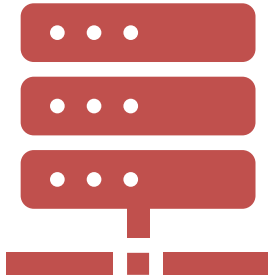


How Does Globus Work?

Globus allows you to manage and transfer data between nodes they call *Collections*. Globus supports three types of Collections:

- Managed Collections
- Guest Collections
- Globus Connect Personal Collections

Managed Collections



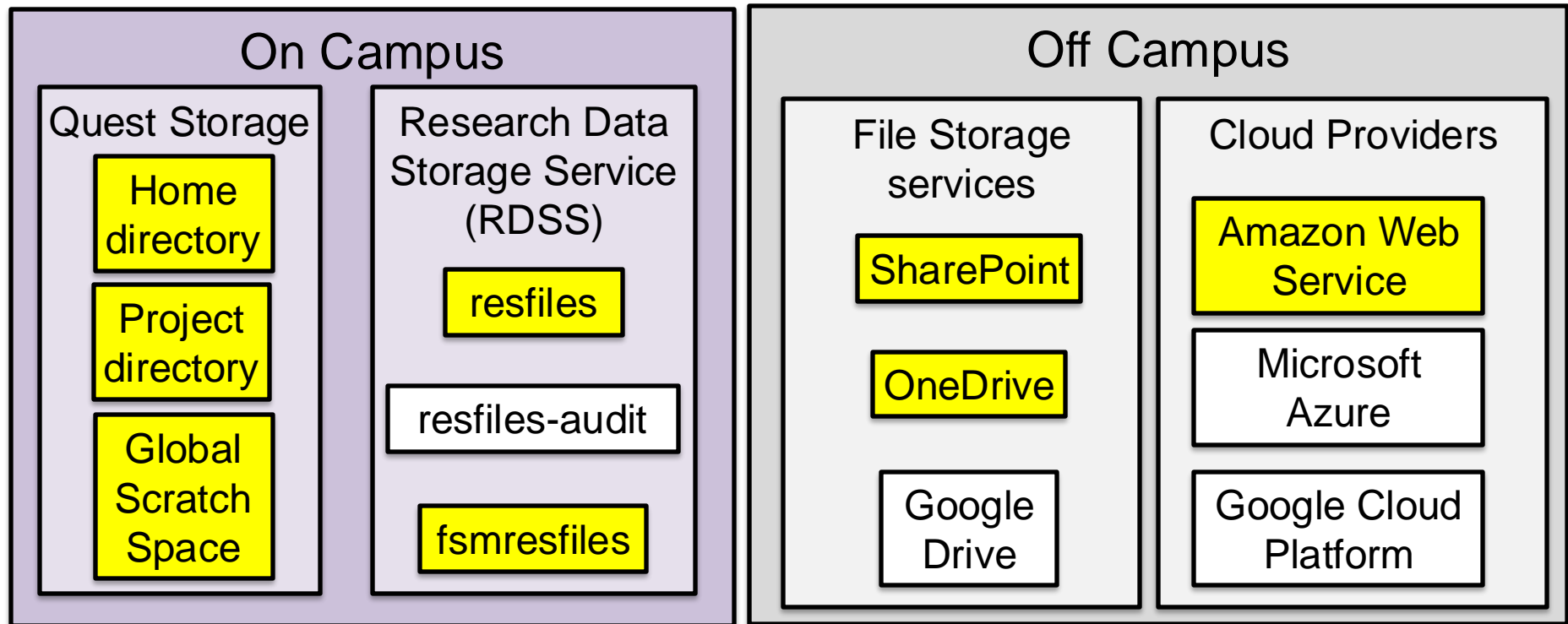
A Globus Managed Collection provides access to centrally administered storage platforms.



NU IT supports Managed Collections that provide access both on-prem and cloud-based storage systems

Globus at Northwestern University

NU IT supports Globus Managed Collections that access both on-premises and cloud-based storage



Northwestern Managed Collections

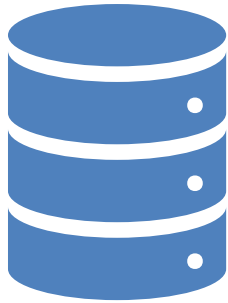
<u>Name</u>	<u>UUID</u>
Northwestern Quest	d5990400-6d04-11e5-ba46-22000b92c6ec
Northwestern Quest RDSS	8f796c9e-f5c8-11e5-9842-22000b9da45e
Northwestern AWS us-east-1 N. Virginia	bf0474ea-df2f-45f6-82e3-41d062d84ce7
Northwestern AWS us-east-2 Ohio	a14148f0-23dc-488f-affd-bd981f919b10
Northwestern AWS us-west-2 Oregon	880cd765-5d14-43f6-97a6-2dbb355f3560
Northwestern OneDrive Pilot*	930c2fcb-416e-4540-a757-496f86acbe70

**Provides access to both OneDrive and SharePoint*

Globus and RDSS

- Globus access to RDSS must be requested so that NU IT can configure your Globus access to specific RDSS share paths.
- To submit a request for access, please submit the Request [Access to RDSS via Globus Form](#) found in the NU IT Service Catalog. In the form you will need to include the following information:
 - NetID
 - Full share path:
 - `//resfiles.northwestern.edu/SHARENAME/`
 - `//fsmresfiles.fsm.northwestern.edu/fsmresfiles/DEPARTMENT/LAB/`

Guest Collections



A Globus Guest Collection is a shared storage space within Globus that provides collaborators to access designated data.



Allows researchers to collaborate with external partners without compromising data security.

Globus Connect Personal Collections



A Globus Connect Personal Collection is enabled by installing the Globus Connect Personal client on a local machine.



Globus Connect Personal Collections allow for efficient transfers between local storage and other Globus endpoints.

Globus Connect Personal

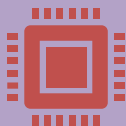


[Download Globus Connect Personal](#)

Supports Mac, Windows, and Linux machines



During installation, the user will need to log into Globus and provide consent for Globus to access the local file system

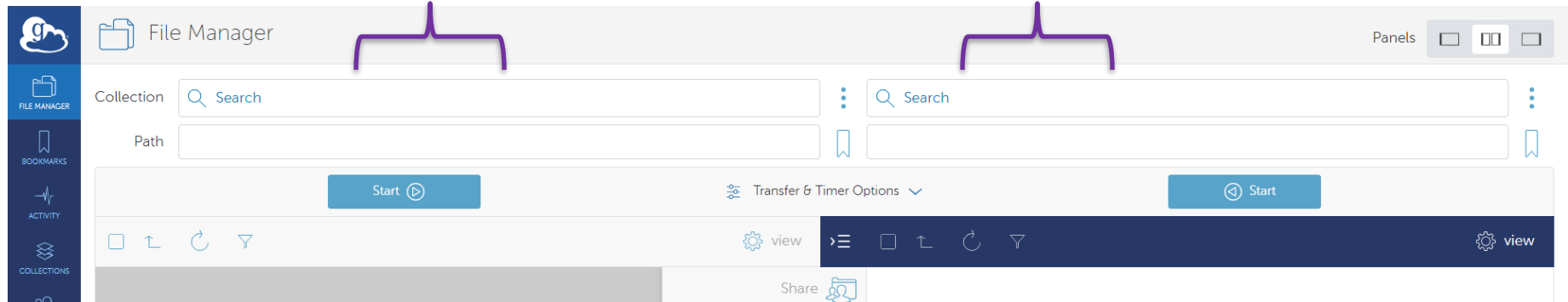


It is recommended that a globally unique name is used for the Globus Connect Personal Collection during the installation process

File Manager: Setting up your transfer

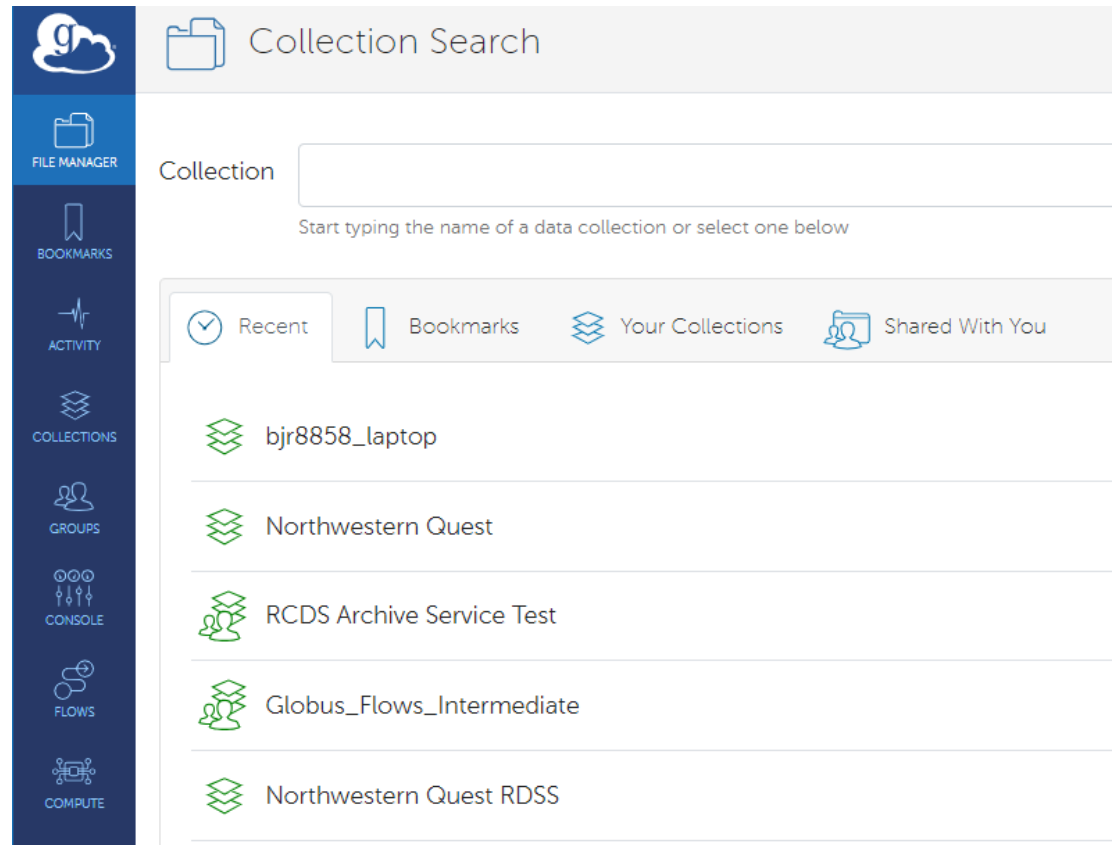
Source Collection

Destination Collection



File Manager: Selecting Globus Collections

- Clicking on the Collection field for either the Source or Destination of the data transfer, opens a new page that includes a search field where you can enter the name or UUID for the Collection you wish to access
- The Collection Search page will have tabs that will list:
 - Recently used Collections (**Recent**)
 - **Bookmarks** you created for quick access to Collections and paths on that Collection
 - **Your Collections** that you own. These can be Guest Collections or your Globus Connect Personal Connections
 - Guest Collections created by other users where you have been provided access permissions (**Shared With You**)



File Manager: Selecting Data Source

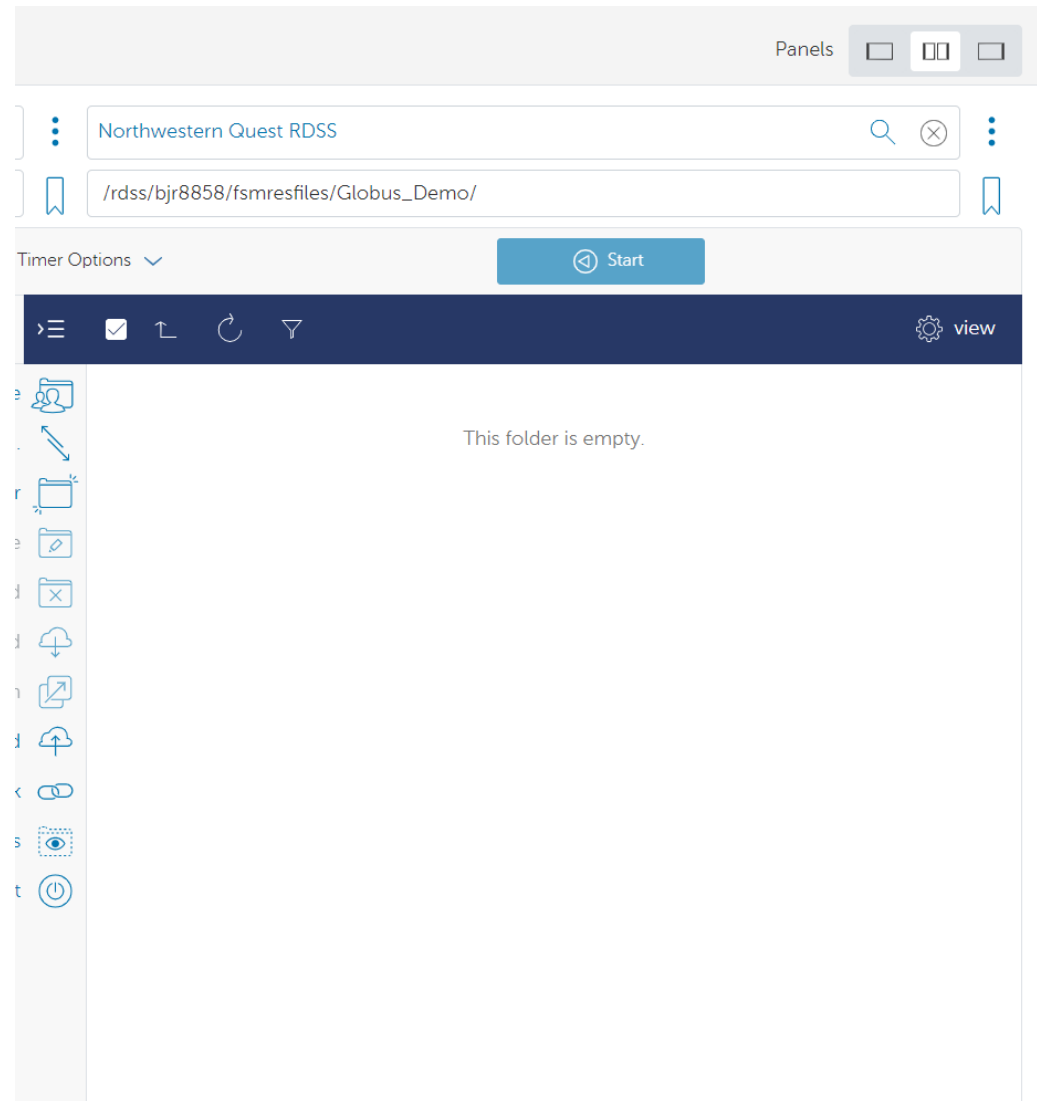
- Once you select a Collection, you can navigate to the specific path they want to access. This can be done manually in the **Path** field or by navigating the file explorer using the Globus interface.
- Globus will display the files and subdirectories in the path provided. The user will be able to select which files and directories they want to transfer from the data source.

The screenshot displays the Globus File Manager interface. On the left is a dark blue sidebar with navigation icons for FILE MANAGER, BOOKMARKS, ACTIVITY, COLLECTIONS, GROUPS, CONSOLE, FLOWS, COMPUTE, SETTINGS, LOGOUT, and HELP & SITEMAP. The main panel has a header 'File Manager' and two input fields: 'Collection' set to 'Northwestern Quest' and 'Path' set to '/home/bjr8858/Globus_Demo/'. Below these is a 'Start' button and a 'Transfer & Timer Options' link. The central area shows a table of files with columns for NAME, LAST MODIFIED, and SIZE. To the right of the table is a vertical toolbar with icons for Share, Transfer, New Folder, Rename, Delete, Download, Open, Upload, Get Link, Show Info, and Manage.




NAME	LAST MODIFIED	SIZE
sampleFile_2.txt	5/4/2023, 04:26 AM	0 B
sampleFile_3.txt	5/4/2023, 04:26 AM	0 B
sampleFile_4.txt	5/4/2023, 04:26 AM	0 B
sampleFile_5.txt	5/4/2023, 04:26 AM	0 B
sampleFile.txt	11/4/2022, 01:34 PM	0 B

File Manager: Selecting Data Destination

- Follow the same steps as selecting the data source for the data destination. You need to select the Collection name or enter the UUID in the blank Collection field in the 2nd panel.
- Once you select a Collection, you define the path or navigate to the folder that will be the final destination for the data transfer.
- Globus will display any files and subdirectories that already exist in the path provided.




File Manager: Transfer & Timer Options


 1 Transfer & Timer Options  


Label This Transfer


Transfer Settings


NOTE: These settings will persist during this session unless changed.


☐ sync - only transfer new or changed files 


☐ delete files on destination that do not exist on source 


☐ preserve source file modification times 

☐ do NOT verify file integrity after transfer 


☐ encrypt transfer 


☐ Skip files on source with errors 


☐ Fail on quota errors 

☐ Apply filter rules to the transfer 


Notification Settings


☐ Disable success notification 


☐ Disable failure notification 

☐ Disable inactive notification 

Schedule Start

mm/dd/yyyy --:-- -- 


 Repeat

days 

every

days

end

never 

Recommended Transfer Options

- **Sync:** Files will only be transferred if it meets a selected criteria when comparing the source to the destination of the transfer. Criteria include:
 - Modification time is newer
 - File size is different
 - File does not exist on destination
 - File checksum is different
- **Preserve Source file modification times**
- **Encrypt Transfer:** Data in your files will be encrypted while in transit. Both endpoints must support encryption to ensure a successful transfer.
- **Skip files on source with errors:** Files on source with "file not found" and "permission denied" errors will be skipped rather than causing the entire transfer procedure to fail.

File Manager: Submitting Transfer

The screenshot displays the File Manager interface with a transfer configuration. The left sidebar contains navigation icons for File Manager, Bookmarks, Activity, Collections, Groups, Console, Flows, Compute, and Settings. The main area shows a transfer setup between two locations:

- Collection:** Northwestern Quest
- Path:** /home/bjr8858/Globus_Demo/
- Transfer & Timer Options:** A dropdown menu is open, showing options: Share, Transfer or Sync to..., New Folder, Rename, Delete Selected, Download, Open, and Upload.
- Start Button:** A blue button with a play icon and the text "Start" is circled in yellow.
- File List:** A table with two columns: NAME and LAST MODIFIED. It lists five files, all of which are checked for selection.

NAME	LAST MODIFIED
<input checked="" type="checkbox"/> sampleFile_2.txt	5/4/2023, 04:26 AM
<input checked="" type="checkbox"/> sampleFile_3.txt	5/4/2023, 04:26 AM
<input checked="" type="checkbox"/> sampleFile_4.txt	5/4/2023, 04:26 AM
<input checked="" type="checkbox"/> sampleFile_5.txt	5/4/2023, 04:26 AM
<input checked="" type="checkbox"/> sampleFile.txt	11/4/2022, 01:34 PM

On the right side of the interface, there is a section for the destination location:

- Collection:** Northwestern Quest RDSS
- Path:** /rdss/bjr8858/fsmresfiles/Globus_Demo/
- Start Button:** A blue button with a play icon and the text "Start".
- Message:** "This folder is empty."

Monitoring Globus transfer status

The screenshot shows the Globus Activity monitoring interface. On the left is a dark blue sidebar with icons for FILE MANAGER, BOOKMARKS, ACTIVITY (selected), COLLECTIONS, GROUPS, CONSOLE, FLOWS, COMPUTE, SETTINGS, and LOGOUT. The main area has a light gray header with a green pulse icon and the word 'Activity'. Below this are tabs for 'Tasks' and 'Timers'. A search bar is present with the placeholder text 'search tasks by label, collection, or owner'. Below the search bar are filter sections: 'QUICK FILTERS' with a 'TYPE' dropdown set to 'TRANSFER' and checkboxes for 'DELETE', 'STATUS' with checkboxes for 'ACTIVE', 'INACTIVE', 'FAILED', and 'SUCCEEDED'. To the right of these is an 'Advanced Filters' button with a gear icon. Below the filters is a 'List Update in 20' button with a refresh icon and a pause icon. Below that is a 'Sort' dropdown. The main content area displays a list of three completed transfer tasks, each with a green checkmark icon, a title, a path, and a status 'transfer completed'. Each item has a right-pointing chevron icon.

Activity

Tasks Timers

search tasks by label, collection, or owner

QUICK FILTERS TYPE ☒ TRANSFER ☒ DELETE STATUS ☒ ACTIVE ☒ INACTIVE ☒ FAILED ☒ SUCCEEDED













Advanced Filters


List Update in 20


Sort


- ✓ **Northwestern Quest to bjr8858_laptop**
/projects/a9009/GlobusReports/ to /C/Globus/
transfer completed
- ✓ **Northwestern Quest to RCDS Archive Service Test**
/projects/b1159/ to /rcds-archive-bjr8858-123/
transfer completed
- ✓ **Northwestern Quest to Globus_Flows_Intermediate**
/home/bjr8858/ to /
transfer completed

Monitoring Globus transfer status: Overview



**Northwestern Quest RDSS to Northwestern Quest RDSS**
transfer completed

 Overview

 Event Log

Task Label	Northwestern Quest RDSS to Northwestern Quest RDSS
Source	► Northwestern Quest RDSS
Source Local User	bjr8858
Destination	► Northwestern Quest RDSS
Destination Local User	bjr8858
Task ID	67f04aaa-9e8d-11ee-8801-a52c65340a88
Owner	► Brian Roland (bjr8858@northwestern.edu)
Condition	SUCCEEDED
Requested	12/19/2023, 10:41 AM
Completed	12/19/2023, 10:41 AM
Duration	5 seconds
Base Paths	Source /rdss/bjr8858/resfiles/broland/ Destination /rdss/bjr8858/fsmresfiles/broland/
Transfer Settings	<ul style="list-style-type: none">• verify file integrity after transfer• transfer is not encrypted• overwriting all files on destination

1 Files

0 Directories

1 Files Transferred

170.65 MB Bytes Transferred



30.96 MB/s Effective Speed


0 Skipped files on sync



0 Skipped files on error

[View debug data](#)


Monitoring Globus transfer status: Event Log








FILE MANAGER
BOOKMARKS
ACTIVITY
COLLECTIONS
GROUPS


 Northwestern Quest RDSS to Northwestern Quest RDSS
transfer completed

 Overview  Event Log

All Events Fault Events Skipped Errors Successful Transfers

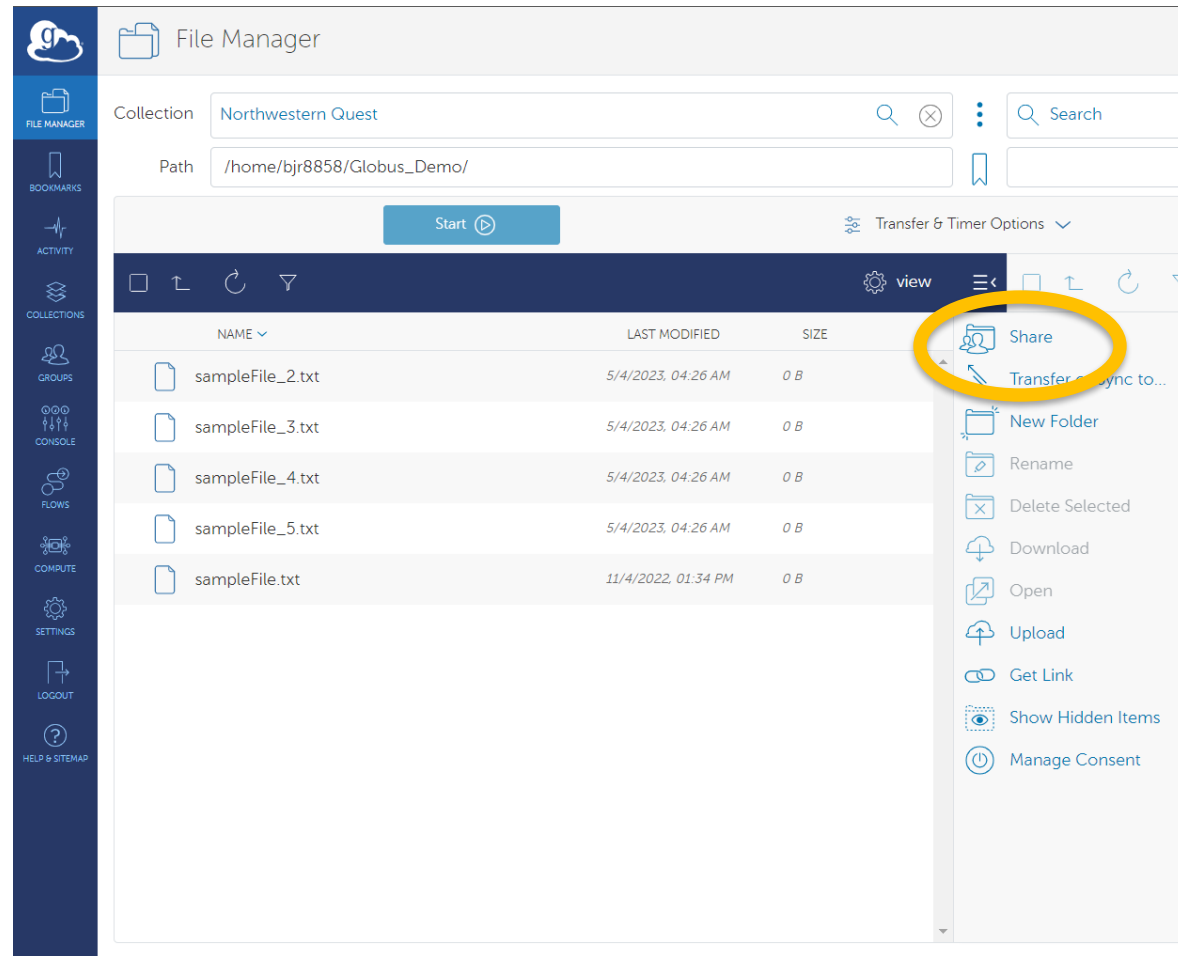
 ⌕ First < Prev Next > Last >⌕

12/19/2023, 10:41 AM	succeeded	 View details 
12/19/2023, 10:41 AM	progress	 View details 
12/19/2023, 10:41 AM	started	 View details 

Data Sharing

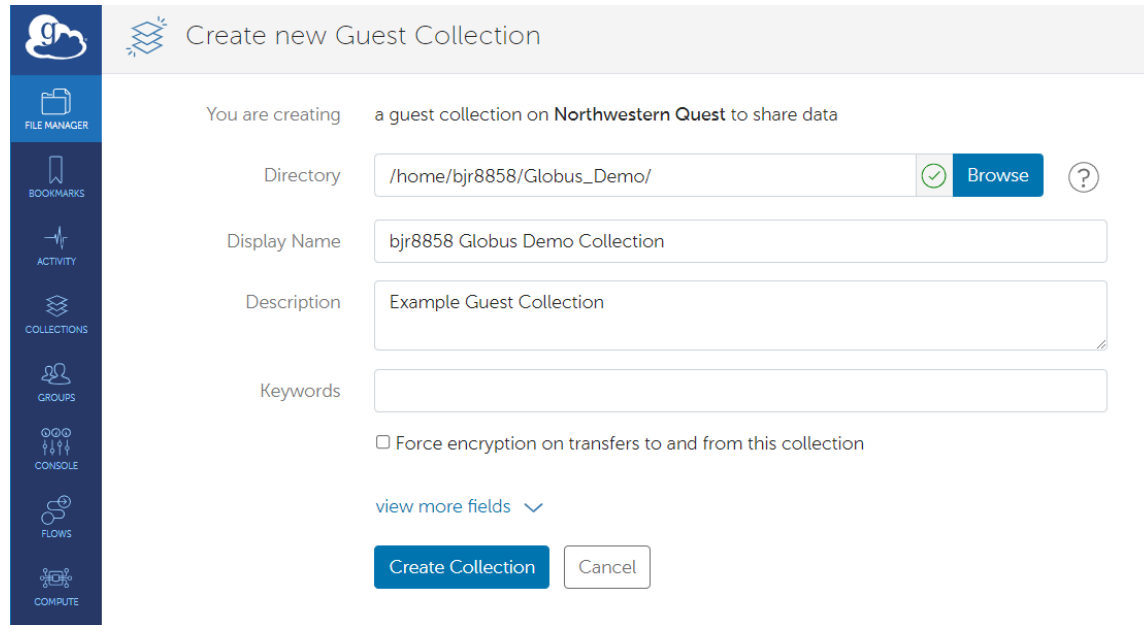
Sharing data with Northwestern collaborators as well as external collaborators is made easy via Globus Guest Collections

- To setup a Guest Collection, navigate to the **File Manager** and select a Managed or Globus Connect Personal Collection in the Collection field.
- Once a Collection is selected, click on the **Share** link/icon to the right of the Collection
- The Globus interface will next show a list of publicly listed Guest Collections. You will need to click on **Add Guest Collection** in the upper right-hand corner of the page



Data Sharing: Setup Guest Collection

- You will be asked to define the root directory for the Guest Collection. This will be the highest-level directory that the Guest Collection can access.
- You are required to enter a display name for the Guest Collection. This name does not need to be Globally unique, but a unique name is recommended.
- Once you have defined the Directory and Display Name for the Guest Collection, you will be able to then create the Guest Collection by clicking the **Create Collection** button
- You will then be shown Permissions information on the Guest Collection. To share this Collection with collaborators you need to click on **Add Permissions – Share With**



The screenshot shows the 'Create new Guest Collection' interface. On the left is a vertical sidebar with icons for FILE MANAGER, BOOKMARKS, ACTIVITY, COLLECTIONS (highlighted), GROUPS, CONSOLE, FLOWS, and COMPUTE. The main area has a header 'Create new Guest Collection' and a sub-header 'You are creating a guest collection on Northwestern Quest to share data'. Below this are form fields for Directory (with a 'Browse' button), Display Name (filled with 'bjr8858 Globus Demo Collection'), Description (filled with 'Example Guest Collection'), and Keywords. At the bottom, there is a checkbox for 'Force encryption on transfers to and from this collection', a 'view more fields' link, and 'Create Collection' and 'Cancel' buttons.

Data Sharing: Guest Collection Permissions

- A new page will be displayed that will have a field to define the highest-level directory on the Guest collection that the collaborator(s) can access.
- You then select the individual user or Globus group that you want to access the Guest Collection.
- Read and Write permissions can be assigned to the user or group you are providing access via the Guest Collection.
- By clicking **Add Permission**, you finalize providing the user or group access to the Guest Collection and the underlying data.

The screenshot shows the 'Add Permissions - Share With' interface in the Globus ecosystem. On the left is a dark blue sidebar with navigation icons for File Manager, Bookmarks, Activity, Collections, Groups, Console, Flows, Compute, Settings, Logout, and Help & Sitemap. The main content area is titled 'BJR8858 GLOBUS DEMO COLLECTION' and 'Add Permissions - Share With'. It features a 'Path' field set to '/', a 'Browse' button, and a file list with columns for file names, dates, sizes, and types. The file list includes 'sampleFile.txt' through 'sampleFile_5.txt'. Below the file list, the 'Share With' section has radio buttons for 'user' (selected), 'group', 'all users', and 'public (anonymous)'. The 'Username or Email' field contains a search prompt and a magnifying glass icon, followed by an 'Add' button. The 'Permissions' section has checkboxes for 'read' (checked) and 'write'. At the bottom are 'Add Permission' and 'Cancel' buttons.

BJR8858 GLOBUS DEMO COLLECTION
Add Permissions - Share With

Path: / Browse

up one folder refresh list

File Name	Date	Size	Type
sampleFile.txt	11/4/...	0 B	file
sampleFile_2.txt	5/4/2...	0 B	file
sampleFile_3.txt	5/4/2...	0 B	file
sampleFile_4.txt	5/4/2...	0 B	file
sampleFile_5.txt	5/4/2...	0 B	file

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

Username or Email: enter a username or search for user Search Add

Permissions

- ☒ read
- ☐ write

Add Permission Cancel

Data Sharing: Setup Guest Collection

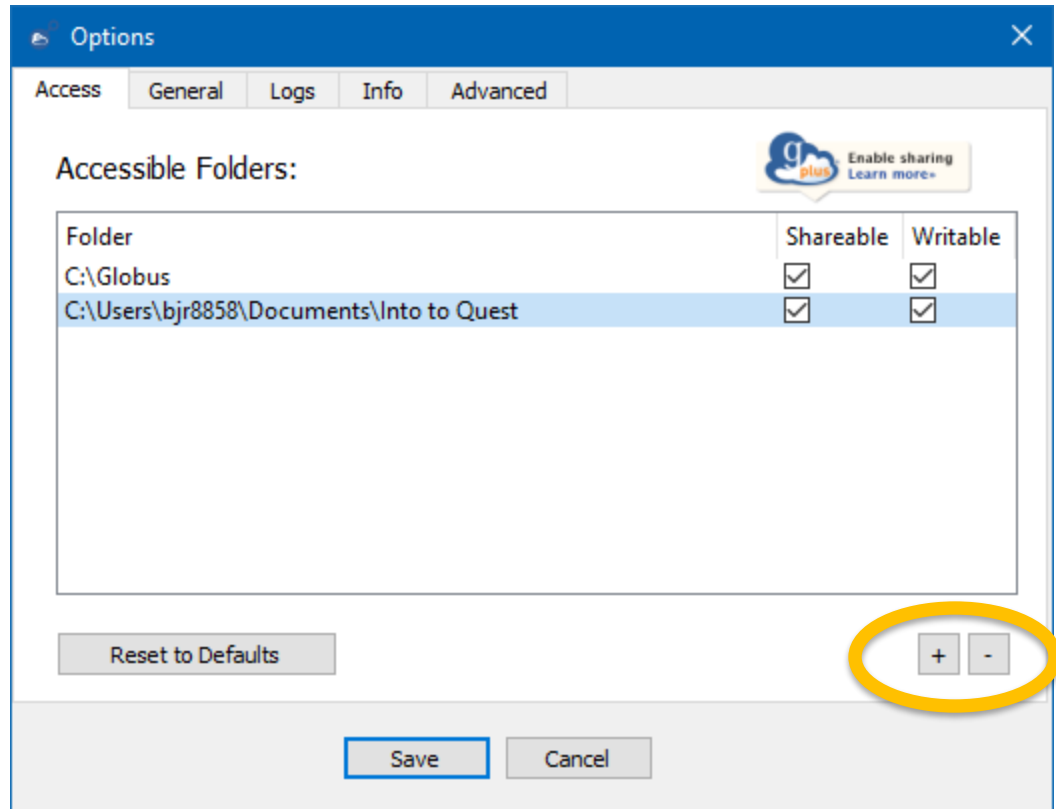
- When done adding permissions, Globus returns you to the Permissions page of the Guest Collection where it lists all who have access to the Guest Collection

The screenshot shows the Globus interface for a 'Guest Collection' named 'bjr8858 Globus Demo Collection'. The 'Permissions' tab is selected, showing a table of users and groups with their access levels. The table has columns for 'USER OR GROUP', 'CREATED', 'READ', 'WRITE', and 'ROLE'. There are three entries: Brian Roland (bjr8858@northwestern.edu), Northwestern Quest - qglobus11 (10cd4d6a-317b-43be-b9ab-04c804d94031@clients.auth.globus.org), and Brian Roland (brian.j.roland@gmail.com). The first two entries have 'READ' and 'WRITE' permissions checked, while the third entry has 'READ' checked and 'WRITE' checked. The 'ROLE' column shows a key icon for the first two and a trash can for the third. A 'Show link for sharing' button is visible next to the 'Path: /' header.

USER OR GROUP	CREATED	READ	WRITE	ROLE
Path: /				Show link for sharing
Brian Roland (bjr8858@northwestern.edu)	-	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	Key icon
Northwestern Quest - qglobus11 (10cd4d6a-317b-43be-b9ab-04c804d94031@clients.auth.globus.org)	-	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	Key icon
Brian Roland (brian.j.roland@gmail.com)	1/15/2024, 12:14 PM	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	Trash icon

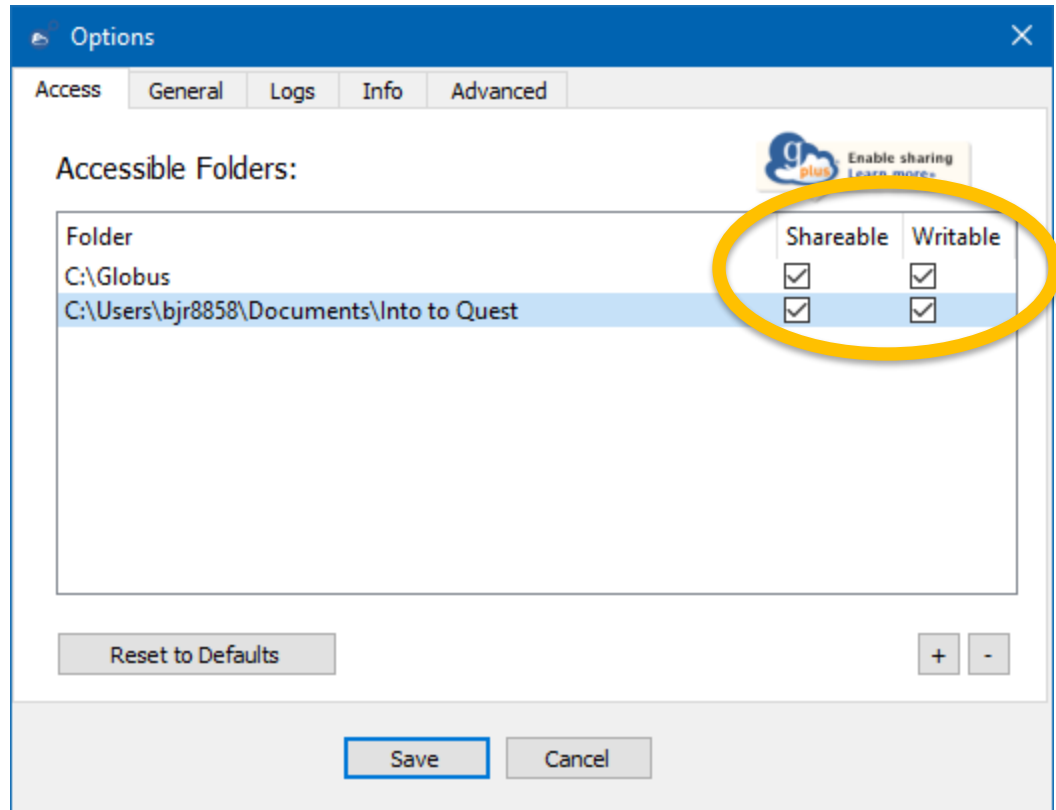
Globus Connect Personal: Settings

- Once you have completed the Globus Connect Personal installation a Globus icon of a lower-case g in a blue circle will appear in the system tray or menu bar of the operating system
- To configure your Globus Connect Personal Collection, use the right mouse button on the icon to display the context menu. You will then want to select Options or Preferences
- A new window will appear that will list the local folders accessible by Globus Connect Personal. You can keep the default location or add/remove folders via the +/- buttons in the lower right-hand corner of the page.



Globus Connect Personal: Settings

- You have the option to make the accessible folders writable or shareable by clicking the checkbox in the appropriate column
- Once locations have been added/removed, as well as the shareable/writable options, you can commit the changes by clicking Save
- After committing the changes, the locations on your local machine will be available via your Globus Connect Personal Collection
- If Shareable is enabled on a location, you will be able to share that location with users via a Guest Collection



Advanced Globus Features



Globus Groups

Collaborative spaces within the Globus platform that facilitate organized and secure teamwork.



Globus CLI

A command-line tool provided by Globus for interacting with the Globus service through the terminal.



Globus SDK

A collection of libraries, tools, and resources provided by Globus to simplify the development of applications and services that interact with the Globus platform.

Globus Groups

- Enables you to create groups, invite members, and manage access permissions for efficient data sharing.
- Provides control over who can access shared resources, ensuring data security and privacy.

Globus CLI

- Globus provides the Globus CLI, a command line tool for interacting with the Globus service through the terminal.
- The Globus CLI is available on Quest as part of software and applications commonly used on Quest.
 - Available version: Globus-cli/3.5.0

Globus SDK

- The Globus SDK is a collection of libraries, tools, and resources provided by Globus to simplify the development of applications and services that interact with the Globus platform
- Offers pre-built functions and utilities for common tasks

Globus Resources

Globus Resources

- [Globus Documentation](#)
- [Globus Connect Personal](#)
- [How To Share Data Using Globus](#)
- [How to configure firewall policy for Globus Connect Personal](#)
- [Globus Command Line Interface \(CLI\) Reference](#)

NU KB Articles

- [Using the Globus Data Transfer Tool](#)
- [Using the RDS Globus Endpoint](#)
- [Using Globus with Amazon S3 Object Storage](#)
- [Using the OneDrive/SharePoint Globus Collection](#)

Questions?

globus-help@northwestern.edu

We are excited to be holding in-person consultation hours on Mondays from 3-4pm at the Mudd Library GIS Lab (2nd Floor across from the bridge to Tech).