
CyVerse Documentation

Release 1.0

CyVerse

Jan 31, 2020

1	Quickstart: Using CyVerse Bisque for a Shared Project	3
1.1	<i>Goal</i>	3
1.2	Prerequisites	3
1.3	<i>Get started</i>	4
1.4	<i>Managing a shared Bisque project</i>	4
1.5	Additional information, help	5
2	<i>Goal</i>	7
3	<i>Prerequisites</i>	9
3.1	Downloads, access, and services	9
3.2	Platform(s)	9
3.3	Input and example data	9
4	<i>Get started</i>	11
5	<i>Share data with project members</i>	13
5.1	Decide who has access to the data	13
5.2	Create folders and set permissions	13
5.3	Use metadata to organize your files	14
6	<i>Share tools and analyses with project members</i>	15
7	<i>Publish data from a shared project</i>	17
8	<i>Additional information, help</i>	19



CYVERSE™



Learning Center Home



CYVERSE™



Learning Center Home



Quickstart: Using CyVerse Bisque for a Shared Project

1.1 Goal

Set up a collaborative imaging project using CyVerse .

1.2 Prerequisites

1.2.1 Downloads, access, and services

In order to complete this tutorial you will need access to the following services/software

Prerequisite	Preparation/Notes	Link/Download
CyVerse account	You will need a CyVerse account to complete this exercise	
Bisque access	Request access to Bisque on the user portal	

1.2.2 Platform(s)

The following CyVerse platform(s) can be used in a collaborative project:

Platform	Interface	Link	Platform Documentation	Quick Start
Data Store	GUI/Command line			
Discovery Environment	Web/Point-and-click			
Bisque	Web/Point-and-click			

1.2.3 Input and example data

No example data are required for this quickstart.

1.3 Get started

1. Any project members who will be using CyVerse should take a look at the [and the](#) .
 2. Be sure that all project members register for CyVerse accounts at the [and](#) request access to Bisque. This action will create a directory in their home folder called *bisque_data*.
-

1.4 Managing a shared Bisque project

Note: This quickstart is in progress. Please check back soon for the complete quickstart.

Managing a shared project with [has](#) some extra considerations, because image files in Bisque are organized into Datasets. Images files that are organized and shared using the Data Store (e.g., using the DE interface) can still be accessed via the image browser in Bisque, but they will not be organized as in the Data Store ([add link to more in Bisque manual](#)). If the Bisque web viewer or API are not needed, shared imaging projects can be managed using the just as for [Managing a regular group project](#). Follow the guide below to decide how best to manage your project.

1.4.1 Choose the management workflow that best fits your project

CyVerse calls itself the “Lego building blocuks of cyberinfrastructure”, which means there are dozens of ways to set up a shared imaging project. Below are some methods that we have found work well. If you have other suggestions, [please let us know!](#)

Manage image data through the Bisque UI

When is this a good option? * Project members need to preview images * Project members need to add graphical annotations * Project uses Bisque apps for analyzing images

Manage image data through the Data Store in shared folder

When is this a good option? * Project data are shared only with project members

Manage image data through the Data Store in Community Released folder

When is this a good option? * Most of the project data are public

1.4.2 Sharing image data with project members

We strongly recommend that a single person be in charge of data management. There should also be a single person (generally the PI) who has ownership of the project folders and who sets read and write permissions for others. This ensures continuity when people move on. The PI can give ownership to a data manager for setting permissions, but should maintain their own ownership as well.

The owner of a folder has the ability to delete or rename the folder and any of its contents. If project members are given write permission to the project folder, they will be able to create their own sub-folders which they will own. In this way, project members can control access to their own data.

Tip: Before beginning your project, make a plan for how to name files and organize datasets. Remember that like traditional folders, datasets in Bisque can be nested, so you can have one overarching dataset, with sub-datasets. Agree on which metadata are needed for each type of file, and set up protocols for adding metadata when files are uploaded.

Sharing of images can be done through the interface or the via the or on the command line using . Project members also can upload and download data using the desktop application , but Cyberduck cannot be used for setting sharing permissions.

According to the , all users receive a default allocation of 100GB. Shared data is counted as part of the allocation of whoever owns the folder that contains it. To request an increase to your allocation, should that become necessary, use the . We expect that users hosting shared directories will need to request larger data allocations.

If your project needs a shared folder for data that that going to be public during the active research phase of the project (e.g., you want to share transcriptomes or draft genomes as they are created, before publication), you can request a . Community Released folders are intended for public data, not for shared projects that are kept private among collaborators.

1.5 Additional information, help

Search for an answer: or

Fix or improve this documentation

- On Github:
 - Send feedback: Tutorials@CyVerse.org
-

 [Learning Center Home](#)

CHAPTER 2

Goal

Learn the basic steps for setting up a collaborative project using CyVerse.

Prerequisites

3.1 Downloads, access, and services

In order to complete this tutorial you will need access to the following services/software

Prerequisite	Preparation/Notes	Link/Download
CyVerse account	Be sure that all project members register for CyVerse accounts at the .	

3.2 Platform(s)

The following CyVerse platform(s) can be used in a collaborative project. Be sure all project members are familiar with them.

Platform	Interface	Link	Platform Documentation	Quick Start
Data Store	GUI/Command line			
Discovery Environment	Web/Point-and-click			

Tip: If your project is image based (i.e. you are sharing a lot of images), you may want to use to manage the data. If so, see the page on [Managing a shared Bisque project](#).

3.3 Input and example data

No example data are required for this quickstart.

CHAPTER 4

Get started

Make a plan. You should have a [data management plan \(DMP\)](#) before you begin your project. There are many resources on the web to guide you on how to create a DMP, including this [CyVerse documentation](#).

The most important DMP elements for this quick start are to know how you will organize your files and folders, what metadata you will use, and to make sure you have a *written SOP* for managing data.

Tip: Before beginning your project, make a plan for how to name files and organize folders. Agree on which metadata are needed for each type of file, and set up protocols for adding metadata when files are uploaded.

Share data with project members

5.1 Decide who has access to the data

There should be a single person (generally a PI) who has ownership of the project folders and who sets read and write permissions for others. This ensures continuity when people move on. The PI can delegate responsibility for setting permissions to a data manager, but should maintain their ownership of all folders as well.

The owner of a folder has the ability to delete or rename the folder and any of its contents. If project members are given write permission to the project folder, they will be able to create their own sub-folders which they will own. In this way, project members can control access to their own data.

Warning: Anyone who has own permission on a folder can delete it or rename it!

For projects that are part of a single lab, we recommend that the PI create a CyVerse account, then create a project directory and share it with lab members. Specific sub-directories can be shared with specific lab members as desired.

For projects that are collaborations among multiple labs, one person should be create a project folder to share with all collaborators. Collaborators must decide among themselves who will host the main folder and who has read, write, and own permission for all folders.

5.2 Create folders and set permissions

Following the agreed upon folder hierarchy, create the base folders for your project and give read, write, or own permission to the appropriate project members

The sharing functionality the CyVerse Data Store can be used to share data among project members. This can be done through the via the or on the command line using . Project members also can upload and download data using the desktop application , but Cyberduck cannot be used for setting sharing permissions.

Tip: According to the , all users receive a default allocation of 100GB. Shared data is counted as part of the allocation

of whoever owns the folder that contains it. To request an increase to your allocation, should that become necessary, use the . We expect that users hosting shared directories will need to request larger data allocations.

If your project needs a shared folder for data that that going to be public during the active research phase of the project (e.g., you want to share transcriptomes or draft genomes as they are created, before publication), you can request a . Community Released folders are intended for public data, not for shared projects that are kept private among collaborators. See more below under *Publish data from a shared project*.

5.3 Use metadata to organize your files

Folder hierarchies and file naming conventions can be an important part of managing a bit data project, but you cannot rely on them to make your data FAIR (findable, accessible, interoperable, and reusable). Good metadata not only makes your data FAIR, but make it easier for you to manage your project by making files searchable and understandable.

CyVerse offers advanced metadata features for managing your data. See more details an the [CyVerse wiki metadata page](#)

Share tools and analyses with project members

Projects can use CyVerse analysis platforms to develop and share analysis tools and workflows.

The (DE) contains hundreds of application that can be used by projects. Apps can be chained together to form workflows in the DE. It is now possible for CyVerse users to integrate their own applications or any open source application into the DE, using Docker containers. Projects may create private apps and workflows, to be shared only with project members, and then make those apps public when they are ready.

In the DE, you can create a team (*add link to documentation*) and share apps with your team.

can be used to set up a virtual machine (VM) with project software, which can then be used by all project members. The VM can later be imaged (made permanent) and published along with the project.

If your project includes a lot of computationally intensive analyses, you should consider requesting an XSEDE allocation (for the U.S. national super-computer infrastructure) and setting up HPC workflows using tools such as . Compute intensive jobs that are highly pallellel can be run on the Open Science Grid (OSG) via the DE.

Publish data from a shared project

When you are ready to publish the results of your project, you should also publish the data to an appropriate repository. For sequence data, that is one of the INDSC repositories, such as NCBI's SRA. Other data types can be published to general scientific repositories or to the . See .

Group projects that are using a to share data pre-publication are encouraged to transition to fully published data (with a DOI) when the data are stable. At that point, data can move into the Data Commons repository in its own folder, or it can remain within the shared project folder, but project members will lose edit access to the dataset. For more questions on this option, contact doi@cyverse.org.

CHAPTER 8

Additional information, help

Search for an answer: or

Fix or improve this documentation

- On Github:
 - Send feedback: Tutorials@CyVerse.org
-

 [Learning Center Home](#)