

Using the Bionimbus PDC For ICGC Downloads:



cdis

CENTER FOR
DATA INTENSIVE SCIENCE



This user guide contains information on:

- Gaining the proper credentials for access to the Bionimbus PDC v1.1
- Applying for “downloader only” access
- Finding the correct directories for the datasets

Getting Credentials for Access

It's very important that a PDC user is the only one who logs in with an account. If you have members of your lab that would like to participate in this research, we ask that they each get separate accounts on the PDC. Having an account used by a single prescribed user is an important step in keeping the PDC secure and protected.

Before your account is created on the Bionimbus Protected Data Cloud (PDC), you'll need the following credentials:

- 1) An eRA Commons account
- 2) dbGaP access
- 3) TCGA access for your research project
- 4) dbGaP / TCGA access assigned to all researchers in your Lab requiring access to the PDC

If you do not have these credentials, we provide detailed instructions in <https://bionimbus-pdc.opensciencedatacloud.org/static/BionimbusAccessFAQ.pdf>. Only once these credentials are obtained can you be granted access to Bionimbus PDC v1.1.

Getting Access to the Bionimbus PDC v1.1

Once you have obtained the credentials listed above please apply for “ICGC Download ONLY Access” at <https://bionimbus-pdc.opensciencedatacloud.org/>

- Select “ICGC Download ONLY Access” in the application process
- When prompted upload a copy of your public ssh key (Note: there is no webconsole for the ICGC/Bionimbus PDC, command line access only).
- A member of our team will be in communication with you once the application is submitted. You will be asked to review and sign some legal documents before access is granted.

Getting Access to the Bionimbus PDC v1.1

Once your account is created you can SSH into the system via <username>@ bionimbus-pdcv1.1.opensciencedatacloud.org. SSH-based tools such as scp and rsync can be used to transfer the data from Bionimbus PDC v1.1.

The following directories have the data for download:

Results from <https://gtrepo-osdc-icgc.annailabs.com> are in /glusterfs/data/ICGC1/gnos/data/<uuid>

Results from <https://gtrepo-osdc-tcga.annailabs.com> are in /glusterfs/data/ICGC2/gnos2/data/<uuid>

Data Responsibility

Once data is transferred out of the Bionimbus PDC, it is the user's responsibility to follow all data use agreements prescribed for TCGA/ICGC, including the Acceptable Use Policy, Code of Conduct and Non-Disclosure agreement for the Bionimbus PDC (signed during application process), and the [NIH GDS Policy](#) which reflects NIH's commitment to responsible data stewardship.