Data Retention Policy:
Due to the large increase in file sizes and speed of data collection, we can only allow data to be kept on our storage server, Qumulo, for up to 30 days after collection has ended. This limit applies both to the raw data and any data generated by cryoSPARC if using the facility's instance. We will inform you if any raw data needs to be deleted to make sure you have made a backup. For cryoSPARC, we will delete any mrc files once the limit has been reached, but we will try to retain most metadata needed to recreate the files.
Many of our users already use their own servers for storage, so please see below of a better explanation of data's lifetime on our servers and if you will be affected by the new data retention limit:

General Data Lifetime Overview:

1. Data generated and stored on dedicated camera storage drives.
    a. These drives can store multiple days of data in the event server transfer stops working. This is meant as intermediate storage and is not meant to be directly accessible by users.
2. Transfer script recognizes new files and moves data according to each lab's rules.
    a. If a lab does not have a synced server, raw data is sent to our San Diego Supercomputing Center (SDSC) managed storage, called Qumulo. **This data is subject to deletion 30 days after data collection ends.**
    b. If a lab has a synced server, then the raw data moves to the user's server and no longer resides at the facility, which means data retention policy does not apply. If needed, we can route the data to Qumulo instead.
3. Raw data is deleted off the original storage drive, residing only on the server of lab's choosing.
    a. Metadata (various diagnostic files, including low resolution JPEGs) are stored as backup on the microscope to help us diagnose potential issues. Data can be deleted early by request.
4. If a user is using Qumulo, they also have access to the Facility's cryoSPARC instance and can pre-process data there. Any data generated by cryoSPARC also remains on Qumulo and **will be deleted after 30 days**.

cryoSPARC Policies:
If you are using the facility's cryoSPARC instance, please keep the following rules in mind:

1. Minimize the amount of data generated.
    a. Work with binned particles as much as possible. We support 4x binning during live extraction unless given prior approval.
    b. If working with EER files, do not use supersampling. Stick with 1x sampling.

    c. Clear intermediates for any larger jobs you might have run.
2. Facility's cryoSPARC instance mainly exist to help labs decide if their data looks good to collect and is not meant to be used for final reconstructions.
    a. Please keep your usage to 2D classification and Ab Initio.
    b. Further processing can be allowed, but please check in with facility staff before proceeding. Unauthorized jobs will be deleted without notice.
3. Do not abuse the number of GPUs. Each user is allowed up to 4 GPUs (two for pre-processing, one for 2D classification, as well as one more GPU for non-continuous short jobs like Ab Initio). Any sessions above 4 GPUs are subject to being stopped early by the facility staff.
4. If using Live, please make sure to stop the process after all the micrographs have been processed.

Any jobs using large amounts of data are subject to deletion before the 30-day limit. cryoSPARC data transfer unfortunately is not trivial, so we do not recommend doing much work on our servers as we cannot guarantee it will transfer without issues.

Requesting Data Transfer:
The easiest way to make sure your data is transferred correctly is to work with UCSD's Physics Computing Facility (PCF) to sync your server of choice to our data transfer script. If that is not feasible, we have multiple ways of transferring data:

1. Connecting to Qumulo directly and moving the data yourself. This requires one time set up to make sure your Active Directory account has the right permissions. Please contact us for the instructions.
2. Requesting a Qumulo storage extension. SDSC sells additional allocations of Qumulo and we make sure your data is seamlessly synced to it.
3. Use cloud services to store data. We can set up AWS transfers and Globus transfers if needed.
4. Use of physical drives. This is the last resort, as we do not allow connecting drives directly into microscope PCs, and will have to use a proxy computer to do so. This solution is usually rather slow and not recommended.

Please contact us if you have any other transfer options in mind.