Backup Guide

About this guide

Digitizing your precious cultural heritage recordings requires both time and resources. Depending on the physical condition of the media and the availability of machines to play the media, you may have few chances to attempt digitization.

It is therefore important that you back up your digitized files and the records related to them. This will protect against threats such as natural disasters, fires, storage-media malfunction, file corruption, user error, theft and vandalism. Regular backups will dramatically decrease the chance of losing data.

This section builds upon the recommendations of the *Digital Storage Guide* providing a detailed example workflow for a rotating hard drive backup.

Backup options

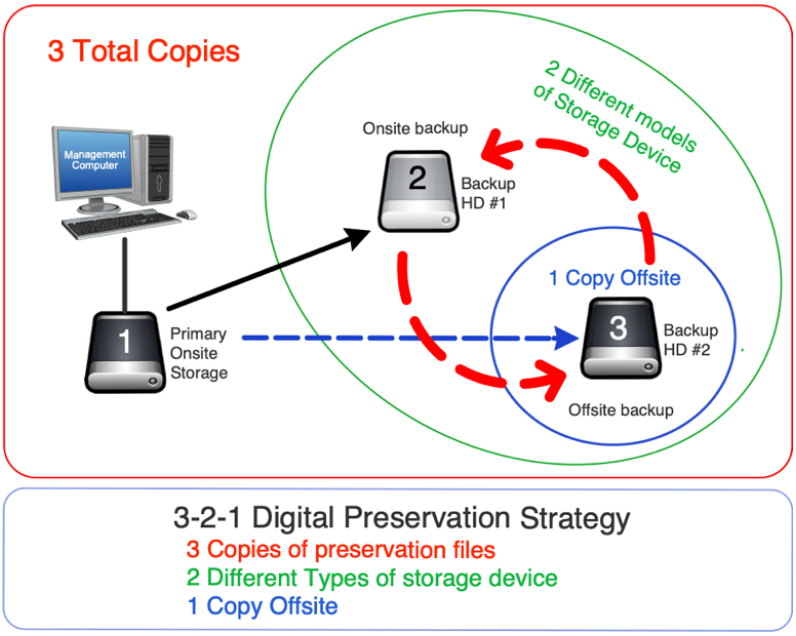
There are many backup options available for our daily digital files. Many can be used in conjunction with one another. Examples include hard drives and other storage media (CDs, DVDs, flash drives, magnetic tape), server/system imaging, virtualization, and cloud backup services.

However, backups for digitization and preservation can be more challenging. To prevent any changes to your high-quality **preservation files**, you will want to avoid any backups that involve **lossy compression**. Further, as digitization produces a high volume of data, many cloud-based backup solutions (which charge by volume) may be too expensive. See the *Digital Storage Guide* for more information about different types of storage and their uses.

Ideally, an IT professional should evaluate your setup and make recommendations regarding backups. If that is not possible, the following guide will provide a basic workflow that you can use to ensure that your data is backed up.

Example backup strategy: External hard drive primary storage with rotating external hard drive backups

The following backup strategy is based on the **3-2-1 backup rule** discussed in the *Digital Storage Guide* section. Indigitization's adapted 3-2-1 backup rule is as follows:



3: Keep a minimum of *three* copies of any important files.

2: Keep the backups on hard disk drives (HDDs), but on at least *two* different brands or models of HDD to protect against manufacturer defects in a single line.

1: Store at least *one* copy off site, preferably in a place that is geographically distant from your office to protect against data loss due to natural disasters.

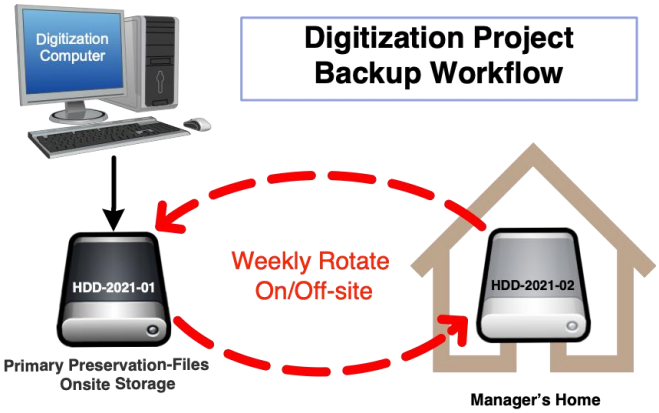
The strategy rotates backups in and out of the office, ensuring that your backups are never all in the same building. It also involves sending an annual backup to a geographically distant partner institution.

Rotating hard drives: Example Workflow

During a digitization project the digital collection of our example organization, Another Band Council Department (ABCD), is growing rapidly and requires regular backup updates. During this phase ABCD is backing up its files on a weekly basis. Its rotation is on Thursday at the end of the day. A manager has taken responsibility for protecting one hard drive offsite.

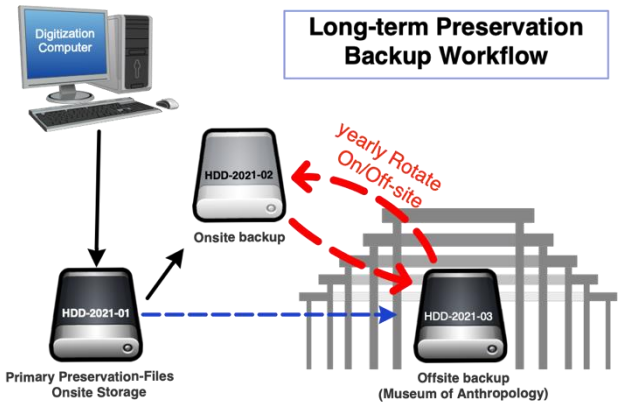
At the end of the project one backup copy of all the digitized preservation-files is sent to a partner organization (the Museum of Anthropology) for safe-keeping. On a yearly basis the on-site backup will be tested for file integrity. If any of the files have become corrupt then that hard disk will be replaced, and a new backup will be created from the primary onsite copy. Once per year the offsite backup will be returned, and the onsite backup will be sent to the trusted distant location.

Digitization project backup routine



- On Thursday, the staff makes a backup copy of the media preservation-files on a hard drive marked HDD-2021-01.
- The manager takes this hard drive (HDD-2021-01) with them when they leave at the end of the day and places the drive in a secure location at home.
- On the next Thursday, staff creates a new backup of the media files on the second hard drive (marked HDD-2021-02).
- At the end of the day, the manager takes this hard drive home, and brings the first hard drive back the next morning.
- This process continues throughout the digitization project, or until that set of hard drives are full.
    - If a set of hard drives becomes full of data, then those hard drives go into the next phase of rotation, with one drive being send to a trusted distant location.
    - A second set of drives must be purchased in order to continue backing up the digitized files.

Digitization project backup routine



- At the end of the digitization project, ABCD staff creates a second backup copy on a hard drive marked HDD-2021-03.
- This hard drive is carefully packaged and shipped, to the partner institution, the Museum of Anthropology.
- The staff make note of the hard drive's location and the date it was sent in the organization's log.
- At a predetermined time, staff arranges for the HDD-2021-03 hard drive to be returned and sends the backup-twin hard drive HDD-2021-02 to the Museum of Anthropology.
- This process continues yearly.
  Note: it is recommended that hard drives be replaced at minimum every 5 years.

Rotating hard drives: What to purchase?

You should purchase three hard disk drives. Ideally, when first starting this backup method, the size of each hard drive (in TB) will be large enough to fit all your data for the next three years. If

you have no **legacy data**, 4 TB HDDs should be enough to start for audio and image digitization projects. If your current media volume is 6 TB, and you produce about 1 TB of additional volume each year, consider getting 10 TB HDDs. At least two different brands or models of HDD should be purchased to protect against manufacturer defects. This is not a one-time purchase; hard drives should be retired after three to five years to decrease the risk of data loss due to mechanical failure or **bit rot**.

Note: Hard drives can fail at any time. Just because a hard drive is new and has always been handled properly does not mean that it is impossible for it to fail. Older hard drives are increasingly susceptible to failure, so it is important to retire hard drives after three to five years. If a hard drive that has not been backed up makes an unusual or loud sound, especially a grinding or clicking sound, power it down or disconnect it as soon as possible. Sometimes data can be recovered from drives that fail, but it will likely cost thousands of dollars, and there is no guarantee that the data will be recoverable. If one of the hard drives you are using for backup rotations makes such noises, immediately retire it.

You should also purchase software, such as TeraCopy, GS RichCopy 360, or rsync, for copying and verifying copies via **checksum**. It is important to have some sort of data validation to ensure that your backup files remain unchanged from the primary set during the backup process. See the *Folder Structures* section for more details on how to ensure fixity when copying backup files.

Optional: You may consider purchasing hard shell, foam-lined cases (such as Pelican cases) for safe transportation of hard drives during rotation. One is sufficient to start. Never use plastic or Ziploc bags to transport or store HDDs.

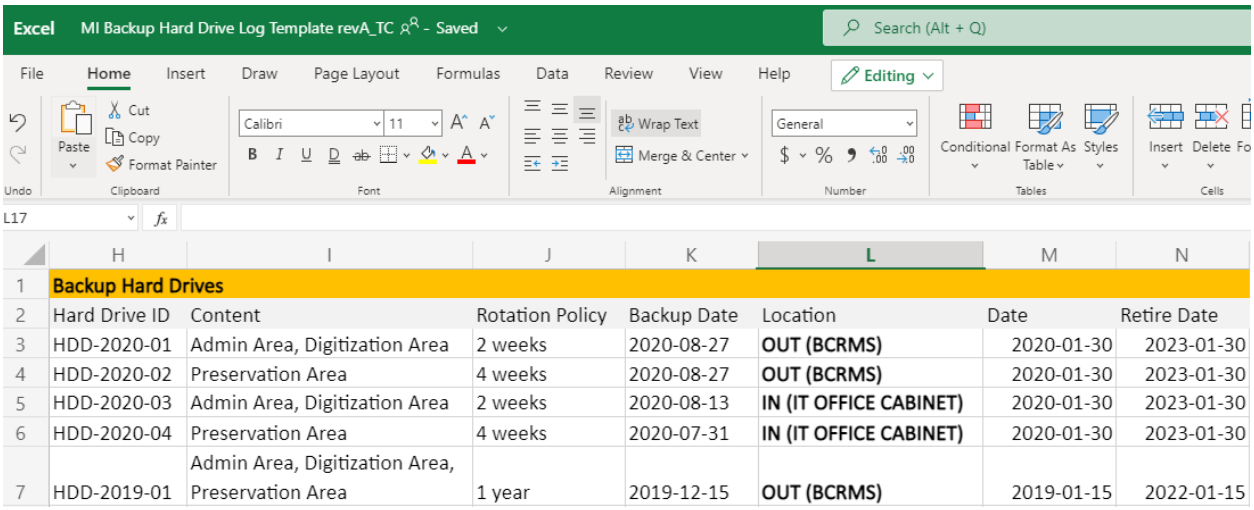
Rotating hard drives: Getting started

First, number your hard drives. This will help you keep track of where the hard drives are and what they contain. With a permanent marker, mark the hard drives with the following unique identifiers:

HDD-2021-01
HDD-2021-02
HDD-2021-03

“HDD” for “hard disk drive,” 2021 for the year purchased, and 01, 02, and 03 as sequential numbers.

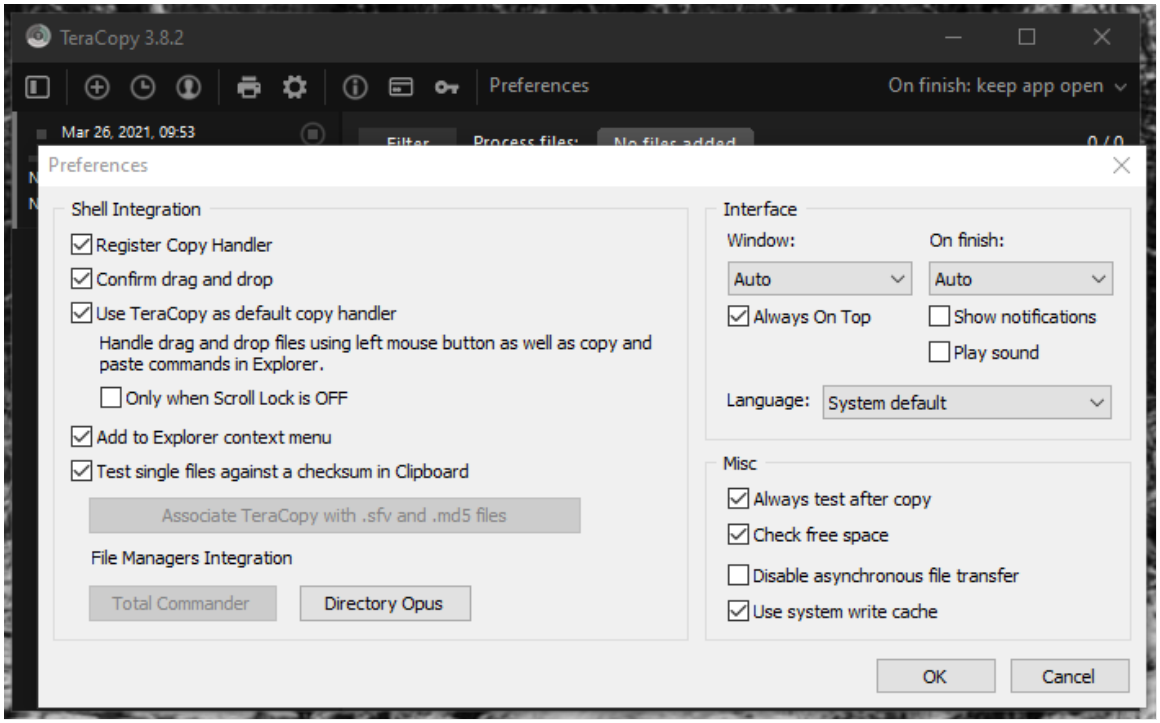
Once backup rotations begin, it is useful to have a spreadsheet to track each HDD’s content, location, backup date, in/out status, and age, as well as a schedule for the rotations. We have included template of a *Hard Drive Log* as an example tracking spreadsheet. You may edit this template as needed. Keep this spreadsheet filed in your project administration area, under 📁**-15_Backups**. You may also find it helpful to schedule reminders for backups in your calendar.



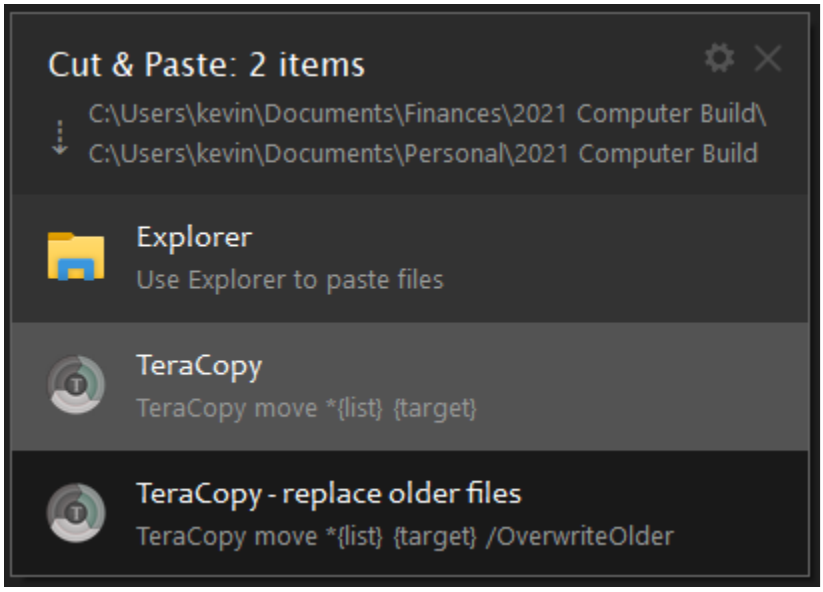
	H	I	J	K	L	M	N
1	**Backup Hard Drives**						
2	Hard Drive ID	Content	Rotation Policy	Backup Date	Location	Date	Retire Date
3	HDD-2020-01	Admin Area, Digitization Area	2 weeks	2020-08-27	**OUT (BCRMS)**	2020-01-30	2023-01-30
4	HDD-2020-02	Preservation Area	4 weeks	2020-08-27	**OUT (BCRMS)**	2020-01-30	2023-01-30
5	HDD-2020-03	Admin Area, Digitization Area	2 weeks	2020-08-13	**IN (IT OFFICE CABINET)**	2020-01-30	2023-01-30
6	HDD-2020-04	Preservation Area	4 weeks	2020-07-31	**IN (IT OFFICE CABINET)**	2020-01-30	2023-01-30
7	HDD-2019-01	Admin Area, Digitization Area, Preservation Area	1 year	2019-12-15	**OUT (BCRMS)**	2019-01-15	2022-01-15

Rotating hard drives: Making a backup copy with integrity checking (Method 1 - TeraCopy)

1. **Connecting, Renaming, and Adding Folders to your Backup Hard Drive**
   a. Plug your external hard drive into your computer using USB.
   b. Rename the external hard drive so that the name of the hard drive matches the Hard Drive ID. This can be done by right-clicking the drive in the left panel of Windows File Explorer and clicking “rename.”
   c. Create folders in the backup drive labelled “(Drive name/folder name) Backup” for each drive or folder to be backed up.
2. **Open and Configure TeraCopy**
   a. Open TeraCopy.
   b. The first time you open TeraCopy, click on the “Preferences” icon. Ensure that “Use TeraCopy as default copy handler” is selected in the Shell Integration area, and that “Always test after copy” is selected under the Misc section.

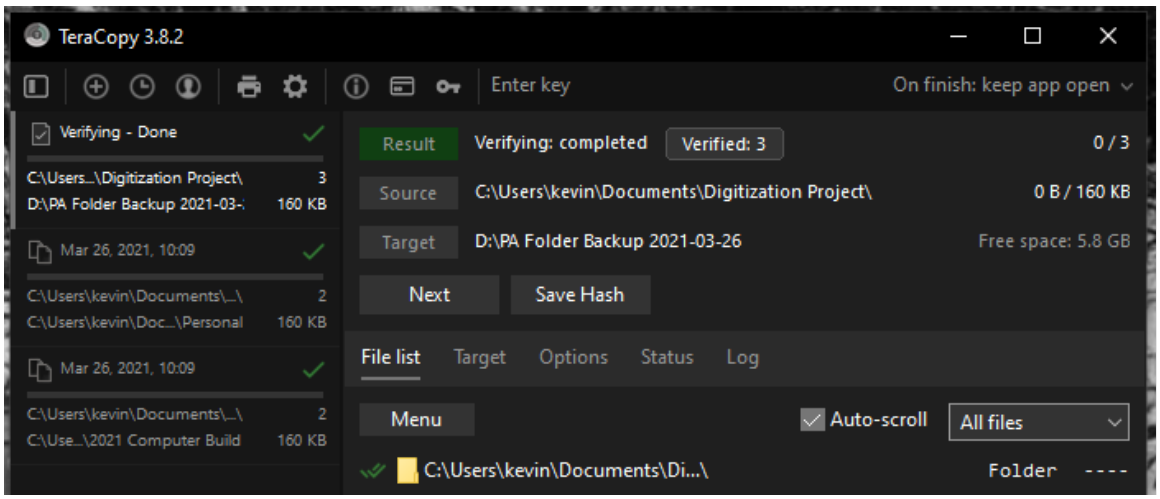


3. **Start the Copy Process**
   a. To perform your backup, copy the contents of the drive or folder you are backing up and paste them to the corresponding folder in your backup hard drive.
   b. A dialogue box may appear asking you if you want to use Explorer or TeraCopy for your action. Click “TeraCopy.”



4. **Wait**
   a. Wait for the job to finish. This may take some time if you have a large amount of data.

5. **Verify the Integrity of Copied Files**
    a. When the copy process is completed check the TeraCopy file list to verify that all files have been copied and passed the checksum verification. If any files have failed to copy, then check to see if those files are locked or corrupt. If they are locked then check to see if you or someone else has the file open, close it, and reattempt the copy of that file with TeraCopy. If the file is corrupt you will have to re-digitize the file.



    b. If there are no errors encountered or remaining after step 5(a), you have completed this backup. Be sure to also spot check the copied data in the backup hard drive to ensure that there was no user error in what was copied over.
6. **Repeat as Needed**
    a. Repeat the process for any additional drives/folders to be copied to this backup hard drive.

Rotating hard drives: Making a backup copy with integrity checking (Method 2 - GS RichCopy 360)

1. **Connecting, Renaming, and Adding Folders to your Backup Hard Drive**
    a. Plug your external hard drive into your computer using USB.
    b. Rename the external hard drive so that the name of the hard drive matches the Hard Drive ID. This can be done by right-clicking the drive in the left panel of Windows File Explorer and clicking “rename.”
    c. Create folders on your hard drive labelled “[Drive name/folder name] Backup” for each drive or folder to be backed up.
2. **Configure GS RichCopy: Create and Name a New Job**
    a. Open GS RichCopy 360.
    b. Click on the “Jobs” tab, and then on the “New” job button.
    c. A Schedule Wizard box will appear. Under “Job Name,” enter “(Drive name/folder name) Job.”
    d. In the same window, for “Source,” click on the folder icon and select the drive/folder containing the data to be backed up. For “Destination,” select the “(drive letter/folder name backup)” folder in your external hard drive. Click Next.

3. **Configure GS RichCopy: Set Copy Options**
    a. A “Set Copy Options” window follows.
    b. Click “Mirror file(s) and folder(s) from source to destination.”
    c. Under “Copy Features:”
        i. Select “File Compare Options: Timestamp + Size.”
        ii. Select “Copy with Security: Files and Folders.”
        iii. Click the boxes for “Copy open/locked file(s),” “Copy date and time stamp on destination folder(s) to match source,” and “Copy Attributes.”
    d. Click Next.
4. **Configure GS RichCopy: Set Logging, Alerting, and Scheduling Options**
    a. A “Logging, Alerting, and Scheduling” window follows.
    b. Click “Enable Logging” and “Append to Log.”
    c. Click the folder icon next to the “Log File Location” and navigate to our Project Administration folder > Backups > Logs > (Drive Name/Folder Name).
    d. In the “File Name” field, type (DriveName/FolderName)_Backup_Logs and hit open.
    e. In the “Log Level” dropdown box, select “Practical.” Optionally, you may add an email address under “Alerting” if you would like e-mail notifications to go to a particular person after the copy job is completed.
    f. Under “Scheduling,” select “Run this Job Manually.”
    g. Click Next, and then Finish.
5. **Configure GS RichCopy: Set the Verify Checksum Option**
    a. Double click the job you just created. Under “Copy Flag,” click “Verify Checksum.” Click “Save” at the bottom.
6. **Repeat the GS RichCopy Configuration for Additional Folders/Drives**
    a. If you are backing up additional folders/drives to this same backup hard drive, repeat this process, creating a new job for other drives/folders.

Perform a Backup Using Gs RichCopy

You have now set up our backup jobs. To perform a backup, simply:

1. Plug in your external hard drive.
2. Open GS RichCopy.
3. Select our first job, “[Drive name/folder name] Job,” and click the “Start” button.
4. Wait for the job to finish. This may take some time if you have a large amount of data.
5. When the job finishes, you may click the Log File next to the progress bar to see if there were any errors. GS RichCopy is configured to copy locked files and files with long file paths, so if an error is encountered check to see if the file is corrupt; if so, it will have to be re-digitized.
6. Spot check your copied data in your hard drive to ensure that there was no user error in what was copied over.
7. Repeat the process for any additional drives/folders to be copied to this backup hard drive.

Rotating hard drives: Rotating the drives

Now that you have a backup of your data, you will start the rotation by sending this hard drive off site.

The frequency of rotations should be regular (for example, every two weeks for your digitization area and every month for your preservation area). You may consider increasing the frequency of rotations if you are digitizing at a rapid pace. Any data that has been produced since your last backup is potentially at risk. If two weeks of data is an unacceptable risk, then consider setting more frequent rotations.

There are also records management service providers, such as Access Records Management. If you are in the service area of one of these companies, you could hire them to rotate your hard drives. They will arrive according to a schedule, return your previous hard drive, take your new backup hard drive, and keep it in their storage until the next rotation. The benefit to this method is that it is secure, their storage is climate controlled, and they are potentially geographically distant. The downside is that this service comes at a cost, and the distance may be insufficient depending on the natural disaster risk profile of your area.

A lower-cost option is to have a trusted employee take the hard drive home with them. While the storage conditions and security may not be as good as an off-site records management service provider, having the employee keep the hard drive in a cool, dry, secure space, like a locked drawer, is a very cost-effective substitute.

Similarly, the backup hard drive in the office should be kept in a cool, dry, and secure space.

Rotating hard drives: Annual drives

Sending an annual backup drive to a partner institution provides protection in case a natural disaster damages your server and other backups.

Consider establishing an agreement with the partner organization that sets out the details of your partnership. This agreement can include clauses about:

- your ownership of the data
- who can access the data or request it to be returned
- what they will do to protect the data from improper access
- the physical conditions of storage
- which positions on each side have responsibility for the drives and their rotation
- how costs will be handled

It is a good idea to plug in and allow hard drives to spin up every 6 to 12 months to prevent data corruption.