Planning Historical Digitisation Projects

- Introduction
- Project Management
- Data Development
- Technical Support
- Backup
- Preservation

Backup

Backing up data is a basic precautionary step that everybody working with computers should take. Backup copies are an insurance policy against the possibility of your data being lost, damaged or destroyed.

Requirements for a Good Backup Policy

A good backup policy will protect your data from a large range of mishaps. The range of events that you should consider when planning how to backup your data includes:

- Accidental changes to data
- Accidental deletion of data
- Loss of data due to media or software faults
- Virus infections and interference by hackers
- Catastrophic events (fire, flood etc.)

A good backup policy should provide protection against all of these threats.

Frequency of Backup

Backups should be made regularly to ensure that they remain up-to-date. The more frequently data is being changed the more frequently backups should be made. If your data is changing significantly every day you should consider a daily backup, but if you are prepared and can afford to redo a longer period of work then less frequent backup may be appropriate.

As well as backing up frequently, you should keep several backup copies made at different dates. Doing this guards against the danger that your backup copy will incorporate a recent, but as yet undiscovered problem, from your working copy.

Multiple Backup Copies

A backup copy may suffer the same mishaps as the working copy of your data, so it is a good idea to spread the risk by maintaining several backup copies. A minimum of two backup copies should be maintained in addition to your working copy of the data.
**Offsite Backups**

More serious events, such as a fire in the office, will destroy both the working copy of the data and any backup copies stored at the same location. Some backup copies should be stored ‘offsite’ (offsite is a relative term, dependent on the level of protection you want).

As well as storing some copies offsite, it is useful to keep a backup copy onsite. This copy can be quickly retrieved and work recommenced if there is a minor mishap, such as the accidental deletion of an important file.

**Media**

Backup copies should be made on new media. Do not continue to use media once they start to develop faults. Specifically, floppy disks are not a good media for backup copies. If they are used, they should be replaced often.

Store backup copies on multiple media (e.g. zip disk and CD-ROM) to avoid all your backup copies becoming corrupted by the same drive or disk fault.

**Multiple Formats**

Store backup copies in both the software formats that you are using and in exported formats (many spreadsheets and database packages can exported to delimited text for example). This will help protect you from subtle faults that can sometimes develop in complicated data formats (such as database file formats) that may not become apparent until after they have been included in both the working copy and the backup copies.

**Institutional Backup Policy**

Projects should never assume that their institution's policies will be appropriate to their needs. Always check.

- Institutions may maintain backups for a limited period
- Institutions may only provide backups to protect against complete loss of data, and not individual users losing data
- Institutions may not backup all data held on their network

Many organisations advise their users to make their own backups of critical data. This is good advice and should be followed.

**Check Your Backup!**

A backup that does not actually work is of no use at all. Always test your backup procedures to ensure that your backup can be retrieved and is useable.

**Backup is not Preservation**

A backup copy is an exact copy of the version of the data you are working on. If your working copy becomes unusable, you should be able to start using your backup copy immediately, on the same computers, using the
same software.

In contrast, a preservation version of the data is designed to mitigate the effects of rapid technology change that might otherwise make the data unuseable within a few years.

Next: Preservation