

Datasets

Key points to remember when creating and managing data

Collecting data through surveys and questionnaires is the only the first step in a complex research process. Analysis would be meaningless without strict practices for creating, preparing, storing and sharing data. For this purpose international standards have been set to ensure that comparisons between various data sources are valid. The UK Data Archive provides guidance to researchers on how to collect and look after data well.

Sharing data

Much research data about people—even sensitive data—can be shared ethically and legally if researchers employ strategies of informed consent, anonymisation and controlling access to data. Ethical guidance is provided by professional bodies, host institutions and funding organisations

Describing data

Raw data cannot be used without some accompanying context. For example, columns in spreadsheet should be meaningfully labelled, and details about how a survey sample was chosen and how data were gathered should be accurately described.

Data security

Making back-ups of files is an essential element of data management. Regular back-ups protect against accidental or malicious data loss and can be used to restore originals if loss of data does occur.

Physical security, network security and security of computer systems and files all need to be considered to ensure security of data and prevent unauthorised access, changes to data, disclosure or destruction of data.

Transferring data

Transmitting data between locations or within research teams can be challenging for data management infrastructure. Data encryption will maintain data security during transmission and should be used when transmitting sensitive or personal data. Collaborative research can pose challenges for facilitating data sharing, transfer and storage, and providing access to data across partners or institutions. Virtual research environments may offer solutions.

Disposing of data

Having a strategy for reliably erasing data files is a critical component of managing data securely and is relevant at various stages in the data cycle.

Deleting files and reformatting a hard drive will not prevent the possible recovery of data that have previously been on that hard drive. At the conclusion of research, data files which are not to be preserved need to be disposed of securely.