Edge Hill University

Research data management guidance

Contents

This guidance should be read in conjunction with the Edge Hill University Research data management policy and Code of practice for the conduct of research.

Why manage your research data?

1. Research data management is essential for responsible research conduct, whether you are staff or student. Even if your research data does not contain sensitive personal information about participants, which needs to be stored securely, you have a professional responsibility to maximise the benefits of your findings by allowing others to use your research data, wherever possible. You

should therefore plan at the start of the project how you will safely and securely manage your research data, both during your project and after it ends. This is often essential to meet funding body requirements. Having a research data management plan will help to ensure research data is accurate, complete, reliable, secure and re-usable.

2. Research data can take many forms beyond empirical datasets (e.g. documents, field notebooks, questionnaires, transcripts, photographs, films, artefacts, data files, etc.) and be in many formats (e.g. jpeg, statistical files, text, etc.).
3. The following guiding principles should inform your research data management decisions, particularly where the data is the result of public funding:
    a. Research data is a public good produced in the public interest and should be made freely and openly available with as few restrictions as possible in a timely and responsible manner.
    b. Sharing research data is an important contributor to the impact of publicly funded research.
    c. You should be entitled to a limited period of privileged access to the data you collect to allow you to work on and publish your results. The length of this period will depend on the discipline and the nature of the research.
    d. There are legal, ethical and commercial constraints on the release of research data. To ensure that the research process is not damaged by inappropriate release of data, these constraints should be considered at all stages in the research process
    e. Institutional and project-specific data management policies and plans should be in accordance with relevant standards and community best practice and should exist for all data. Data with acknowledged long-term value should be preserved and remain accessible and usable for future research.
    f. Public funds may be used to support the preservation and management of publicly funded research data.

Who is responsible?

4. The principal investigator is responsible for making the data open and ensuring there is a metadata record on the EHU research data repository, taking all necessary steps to comply with data protection and other regulations.
5. If the principal investigator is not an Edge Hill staff member or student, the ‘lead’ EHU researcher is responsible for ensuring there is a metadata record catalogued on the EHU research data repository.
6. Other roles with key responsibilities include:
    a. Heads of Department ensure that staff and students in their areas are aware of their responsibilities

b. Departmental research leads support staff in understanding the operational aspects of their responsibilities, particularly in signposting new staff to institutional resources

c. Associate Deans for Research (or equivalent) ensure that policy developments are disseminated among all research staff and students in given faculty

d. Research Office has strategic oversight for research data management and can advise on policy

e. Learning Services has operational responsibility for the research data repository and can advise on technical/practical aspects of research data management

f. IT Services provide network access and can advise on the appropriateness of data storage solutions

g. Research ethics committees should consider data management as part of scrutiny and approval processes

h. SPPU has overall responsibility for general data protection and compliance at the University.

Research data management plans

7. Creating a research data management plan (DMP) at the start of your research project will help you identify your data management needs and areas of data protection risk. You should complete a DMP regardless of funding method, and keep the plan under review throughout the life of the project.

8. The first stage is to consider which DMP Online template to use. The default will be Edge Hill's institutional template unless your funder requires the use of a specific other template.

9. Edge Hill University has adopted the standards of the Digital Curation Centre (DCC). The DCC's online data management planning tool, DMP Online, helps you create a data management plan in accordance with the University's requirements and those stipulated by the major UK funders.

10. Your data management plan should include:

    a. Named researchers and their areas of access to, and responsibility for, research data.

    b. What research data will be collected/created?

    c. Which policies and legislation (internal and external) apply to the data?

    d. What data storage methods will you use for live research data (i.e. during the project)?

    e. What facilities, IT and equipment will you require?

f. How and where will the research data be preserved to enable re-use after the research project is completed?

Costing data management

11. All activity at the University incurs a cost – it could be for your time in managing the research data, or for the storage of live data (during the project) or the long-term preservation of it after the project.

12. The cost of research data management should be calculated for inclusion in a funding application or data management plan. Some funders will require you to factor this cost in to your funding application, while the University needs to know how much it spends on storage whether or not it recoups the cost. The UK Data Service offers costing tools to assist you with this.

Funder and stakeholder policies

13. You will need to meet your research funder's data management requirements in your data management plan. The DCC provides information on the requirements of major funders.

14. Furthermore, other stakeholders in the project may have their own research data management policies with which you will need to comply, such as external organisations supplying research participants.

15. SHERPA/JULIET lists funders' rules and requirements regarding the open access of research outputs and data.

Ethical and legal considerations

16. You are expected to maintain high ethical standards and work within the law.

17. Professional bodies, host institutions and funding organisations, usually provide ethical guidelines for research.

18. Edge Hill University Research ethics policy and Code of practice for the conduct of research contain the standards that all Edge Hill researchers are expected to maintain. You should also be familiar with the University's Data protection policy.

19. You have a legal obligation to process all data with which you come into contact in accordance with data protection legislation. For personal information, this includes – but is not limited to – the General Data Protection Legislation (GDPR) and the Data Protection Act 2018. Please consult the SPPU Information Governance wiki or contact the University's Data Protection Officer for advice.

20. Under GDPR, research participants (data subjects) should always be told the lawful basis being used to process their personal data, which can be done with reference to the University's Privacy Policy.

21. There may be a perceived tension between data sharing and data protection. However, data can be shared while upholding data protection legislation and principles of research ethics.

Confidentiality

22. You have a duty of confidentiality when handling people's personal data (for example, that of research participants).

23. As a matter of good practice:

    a. Research participants should be informed about how far they will be afforded anonymity and confidentiality.

    b. Guarantees of confidentiality and anonymity given to research participants must be honoured, unless there are clear and overriding reasons to do otherwise.

    c. You should not breach the 'duty of confidentiality' and not pass on identifiable data to third parties without participants' consent.

    d. Personal/confidential data should only be retained in its original form until it is no longer required and, following anonymisation or a reasonable retention period justified in your DMP, should be destroyed and *securely* disposed of. For example, delete the original audio files/notes once an interview has been transcribed.

24. Please note that research data given in confidence might be liable to summons by a court.

    a. Research participants should be made aware of this fact.

    b. You should guard against giving unrealistic guarantees of confidentiality and anonymity and be aware that in a legal challenge you may be compelled to disclose certain information to the authorities (this should be clarified in your participant information sheet).

25. You should ensure participants are aware that the research data acquired through your project – including that resulting from their participation – will be publicly accessible. You therefore need to ensure you take appropriate and reasonable steps to anonymise individuals' data as much as possible, reducing the risk of identification. These steps and their rights to withdraw their data at different stages of the research project should be clearly communicated to participants.

Working with sensitive data/material

26. You should consult the Edge Hill Policy on researching and handling sensitive material if your research involves material that is security-sensitive, radical or extreme.

27. You must adhere to Edge Hill IT Services' Acceptable use policy when using the University network for any purpose, including storage of your research data. If you have a legitimate need, as part of your research, to view, download, create or transmit material that would normally be defined as unacceptable use, you can

submit an *Application to access sensitive content (for research purposes)* to IT Services.

28. The University of Cambridge has produced a video guide to managing sensitive research data.

Metadata

29. Metadata is essentially data about your research data. It helps you to organise and archive your research data, and enables other users to identify it and to what it relates.

30. Sufficient metadata should be recorded and made openly available via the Edge Hill research data repository to enable other researchers to understand the potential for further research and re-use of the data. We provide more detail on this in Research data management: metadata (RO-GOV-15).

Documentation

31. You need to organise your research data on a regular basis throughout your project, so that it can be located when needed.

32. The long-term preservation of your research data, so that it can be understood and interpreted by another user, requires clear data description, annotation, contextual information and documentation. Data documentation explains how research data was created, what it means, and its content and structure. Published results should always include information on how to access the supporting data.

33. To recognise your intellectual contribution to generating, preserving and sharing key research datasets, all subsequent users of your research data should acknowledge it as a data source and abide by the terms and conditions under which they access it. You should therefore set terms and conditions for your data in a data access statement (see Research data management: metadata, RO-GOV-15).

Storing your research data

During the project

34. Live data is the research data you maintain and add to throughout the life of your project.

35. All Edge Hill staff and students have access to OneDrive: secure cloud storage accessible to you from any device with an internet connection that allows you to share your files with internal collaborators.

36. You cannot use OneDrive to share your files with external collaborators so you should consider other options for the secure sharing of files. IT Services can advise on whether the solutions you identify meet Edge Hill's requirements.

37. IT Services may be able to provide you with dedicated storage on the Y drive but this is allocated on a case-by-case basis and they will need to see your data management plan in order to make a decision.
38. Although you can, and should (see below), keep backups of your research data (e.g. on local drives and devices), you should store your live research data on the University OneDrive (or in your dedicated Y drive folder) and use that as your main data file, ensuring that copy is up-to-date.
39. You should keep clear, accurate, and secure records of the procedures followed and the approvals granted during the research process, including records of the interim results obtained as well as of the final research outcomes. This demonstrates proper research practice, but also allows you to respond definitively in case you are subsequently asked about either the conduct of the research or the results obtained. You do not normally need to deposit this record in the Edge Hill research repository.
40. If you leave the University, and wish to retain data/copies of live data for personal use, you must obtain permission from the principal investigator and head of department to do so.
    a. Where personal data is involved, the request must be refused unless it is clear that future use will be consistent with the terms of the consent (e.g. anonymisation).
    b. You will not need permission for publicly accessible data held on a research repository unless restrictions are in place for access and/or re-use.

Data backup

41. Accidents happen and can have a catastrophic effect on a research project. For example, your data could be destroyed in a fire, your computer could be irreparably damaged by malware, your cloud storage account could be compromised or closed, or you could lose a pen drive containing the data.
42. Such accidents could result in you losing all progress to date in your project and, in some cases, your funder may request reimbursement of funds paid to date because the project cannot realistically meet the approved milestones.
43. It is therefore your responsibility as a researcher to ensure that there are enough adequate, up-to-date backup copies of your research data and related documents to avoid substantial losses to the project. For example, you should:
    a. Backup and securely save electronic format research data.
    b. Digitise hard copy format research data at the earliest opportunity after collection (e.g. scan paper-based data).
    c. Photograph research data that includes physical artefacts.
44. It is good practice to keep at least two backups of your original research data on different devices and in different locations, in case at least one copy is corrupted

or accidentally destroyed. Remember: if your data exists in only one place, it does not exist.

a. The primary version must be stored on a secure drive provided by the University: your Z drive, a Y drive folder with access restricted to essential people, or on OneDrive.

b. Your backups can be stored on flash/USB/pen drives, SD cards, external hard drives, tablets, laptops, etc., all of which should be encrypted in case you misplace them (please also see 'Data security', below).

45. However you manage your data and backups, you will need to comply with data protection legislation outlined earlier.

46. You should periodically test that you can effectively restore data from your backups.

47. Because you will have multiple versions of your research data, you will need to ensure that you have a robust version control system in place to avoid confusing earlier and later versions.

48. The UK Data Service offers advice on backing up your data.

After the project

49. Archival data refers to the completed research data that is stored in a research data repository for long-term preservation after the project has ended.

50. While the research data itself may be saved in a funder or subject-specific repository, you must record the metadata for your research data in the Edge Hill research data repository (see 'Research data management: metadata', RO-GOV-15).

51. Using the UKRI common principles on data policy as a guide, primary research data (and where possible/relevant specimens, samples, questionnaires, audiotapes, etc.) that underpins publications should be retained intact in an appropriate format and storage facility for a minimum of ten years from the date of publication, although this may vary according to the funder and nature of the data:

    a. Some funders require that research data be retained for a minimum of ten years from the last time it was accessed by anybody e.g. a visitor to the Edge Hill research data repository.

    b. Research records relating to clinical or public health studies should be retained for twenty years to provide scope for longitudinal follow-up if necessary; for detailed guidance see MRC guidelines on personal information in medical research.

    c. Research data that informs national policy-making should be preserved in a research data repository permanently.

52. Your specific discipline or area of study may have additional governance and best practice guidelines – it may even be that you cannot share or make certain data public – so please consult your departmental/faculty research lead.
53. You do not need to archive research data that does not underpin a publication. You may, however, wish to retain it for further use: to do so, you must have told your data subjects, where applicable, how long you will keep the data and for what purposes it will be used in future (unless the data is anonymised).
54. If you leave the University, and wish to retain data/copies of live data for personal use, you must obtain permission from the principal investigator and head of department to do so.
    a. Where personal data is involved, the request must be refused unless it is clear that future use will be consistent with the terms of the consent (e.g. anonymisation).
    b. You will not need permission for publicly accessible data held on a research repository unless restrictions are in place for access and/or re-use.
55. Research data created electronically should ideally use common file formats to facilitate re-use which may be several years later, although specialist software may be necessary in some cases that prevents this. In all cases, details of the specific programme or software used should be recorded.

Data security

56. You have legal and ethical obligations to ensure all personal data you collect is secure from unauthorised access (e.g. before anonymisation). Your research data may even be commercially sensitive, protected by intellectual property agreements, or sensitive for security reasons.
57. You must specify and follow a process to handle and delete confidential data in line with the project's exit strategy, and the University's Information security policy. The UK Data Service offers further advice on data disposal.
58. When planning your research project, you should conduct a research data risk assessment to determine:
    a. The potential monetary value of the data (e.g. cost of collection – including your time – and cost of storage);
    b. The level of confidentiality required for the data;
    c. The steps required to provide appropriate data protection;
    d. The potential impact of unauthorised access to the data;
    e. Issues regarding access via Edge Hill's network;
    f. Issues regarding access from outside Edge Hill's network;
    g. Security of data while in transit (physically and electronically).

59. If you collect your research data off campus, you must ensure that it is kept safe between collection and the point at which you can secure it on campus (e.g. for the digitisation of hard copies; transporting backups to your office/home). For electronic research data, the device on which you store or transport the data **must** be suitably encrypted, including the equipment used to record interviews and transfer them to transcribers:

    a. Encrypted data is stored in a scrambled format and is therefore unreadable without an encryption key – set by you, this key could be a passphrase that cannot be easily guessed, or a nonsense string of text and other characters. If you misplace your encrypted device – or if it gets stolen – the lost data remains secure;

    b. Some devices allow you to use your fingerprint as the key (e.g. iPhone, some Android phones);

    c. New mobile phones and tablets are usually encrypted 'out of the box' although you will need to encrypt any SD cards you insert to expand the storage. You can check the encryption status of your device and its external storage through your device's security settings;

    d. It is possible to encrypt flash drives using third party encryption tools but you can also buy drives with encryption out of the box;

    e. If using a laptop or other device, please check its security options to see if you can encrypt it/if it is already encrypted – even if your device requires a password to *log in to your desktop*, removing an unencrypted hard drive can allow someone immediate access to the files stored on it with nothing more than a cable adapter which can be bought from any computer accessories supplier for less than £10;

    f. The UK Data Service recommends the use of technology meeting Pretty Good Privacy (PGP) standards (other encryption technology is available).

    g. The University expects to publish a policy on remote working in 2018/19, which will set out the responsibilities of staff members when handling data while working away from campus.

60. When sharing data with your collaborators – or sending it to yourself – over e-mail, you need to take all precautions necessary to keep it secure. These include:

    a. Encrypting your e-mails if the recipient is external to the University, so only the sender and recipient can view the contents. The IT Services wiki contains advice on Office Message Encryption (OME).

    b. Password protecting the dataset file and sending the password separate to the protected file.

Useful links

Policies & guidance

- Edge Hill Acceptable Use Policy
- Edge Hill Data Protection Policy
- Edge Hill Information Security Policy
- Edge Hill Research Data Management Policy
    - Edge Hill Research Data Management: Metadata

EHU resources

- Edge Hill Research Data Repository
- Edge Hill Research Governance and Ethics web pages
- Edge Hill Information Governance wiki

Other resources

- Digital Curation Centre (DCC)
- UK Data Archive
- UK Data Service

Guidance updated 28 June 2018