|  |
| --- |
| Research Data Management Guidance |
|  |

Contents

[Summary 3](#_Toc97021028)

[Why manage your research data? 3](#_Toc97021029)

[Who is responsible? 4](#_Toc97021030)

[Research data management plans 4](#_Toc97021031)

[Costing data management 5](#_Toc97021032)

[Funder and stakeholder policies 5](#_Toc97021033)

[Ethical and legal considerations 6](#_Toc97021034)

[Confidentiality 6](#_Toc97021035)

[Working with sensitive data/ material 7](#_Toc97021036)

[Metadata 7](#_Toc97021037)

[Documentation 8](#_Toc97021038)

[Storing your research data during the project 8](#_Toc97021039)

[Data backup 9](#_Toc97021040)

[Storing your research data after the project 10](#_Toc97021041)

[Data security 11](#_Toc97021042)

[Useful links 12](#_Toc97021043)

[Appendix A: Metadata 14](#_Toc97021044)

[Endmatter 16](#_Toc97021045)

# Summary

This guidance should be read in conjunction with the Edge Hill University [Research Data Management Policy](https://www.edgehill.ac.uk/documents/research-data-management-policy/) and [Code of Practice for the Conduct of Research](https://www.edgehill.ac.uk/documents/code-of-practice-for-the-conduct-of-research/).

# Why manage your research data?

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.).

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.

The UKRI [common principles on research data](https://www.ukri.org/funding/information-for-award-holders/data-policy/common-principles-on-data-policy/) should inform your research data management decisions, particularly where the data is the result of public funding:

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.

Sharing research data is an important contributor to the impact of publicly funded research.

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.

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

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.

Public funds may be used to support the preservation and management of publicly funded research data. It should be noted that all costs associated with research data management are eligible to be funded by UKRI grants, subject to conditions.

# Who is responsible?

The principal investigator is responsible for making the data open and ensuring there is a metadata record (as defined in section 10 below) on [Figshare](https://figshare.edgehill.ac.uk/), Edge Hill’s repository for a range of research materials including research data. You must take all necessary steps to comply with data protection and other regulations.

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 Figshare.

Other roles with key responsibilities include:

* + Heads of Department ensure that staff and students in their areas are aware of their responsibilities
	+ Departmental research leads support staff in understanding the operational aspects of their responsibilities, particularly in signposting new staff to institutional resources
	+ Associate Deans for Research (or equivalent) ensure that policy developments are disseminated among all research staff and students in given faculty
	+ Research Office has strategic oversight for research data management and can advise on policy
	+ Learning Services has operational responsibility for Figshare and can advise on technical/practical aspects of research data management
	+ IT Services provide network access and can advise on the appropriateness of data storage solutions
	+ Research ethics committees should consider data management as part of scrutiny and approval processes
	+ Strategic Planning and Policy Unit (SPPU) has overall responsibility for general data protection and compliance at the University.

# Research data management plans

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. All research projects should have a DMP, and the plan should be reviewed throughout the life of the project.

Edge Hill University has adopted the standards of [the Digital Curation Centre (DCC)](http://www.dcc.ac.uk/resources/data-management-plans). The DCC’s online data management planning tool, [DMP Online](https://dmponline.dcc.ac.uk/), helps you create a data management plan in accordance with the University’s requirements and those stipulated by the major UK funders.

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.

Your data management plan should include:

* + Named researchers and their areas of access to, and responsibility for, research data.
	+ What research data will be collected/created?
	+ Which policies and legislation (internal and external) apply to the data?
	+ What data storage methods will you use for live research data (i.e. during the project)?
	+ What facilities, IT and equipment will you require?
	+ How and where will the research data be preserved to enable re-use after the research project is completed?

# Costing data management

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. However, the university provides Figshare as a digital research data storage platform (incorporating cataloguing and repository functions) and most storage/preservation needs are anticipated to be covered by this.

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](https://www.ukdataservice.ac.uk/manage-data/plan/costing) offers costing tools to assist you with this.

# Funder and stakeholder policies

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](http://www.dcc.ac.uk/resources/data-management-plans/funders-requirements).

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.

[SHERPA/JULIET](http://www.sherpa.ac.uk/juliet) lists funders' rules and requirements regarding the open access of research outputs and data.

# Ethical and legal considerations

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

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

Edge Hill University [Research ethics policy](https://www.edgehill.ac.uk/documents/research-ethics-policy/) and [Code of practice for the conduct of research](https://www.edgehill.ac.uk/documents/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](https://www.edgehill.ac.uk/documents/data-protection-policy/).

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](https://go.edgehill.ac.uk/display/compliance/Home) or contact the University’s Data Protection Officer for advice.

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](https://www.edgehill.ac.uk/privacy/).

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

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

As a matter of good practice:

* + Research participants should be informed about how far they will be afforded anonymity and confidentiality.
	+ Guarantees of confidentiality and anonymity given to research participants must be honoured, unless there are clear and overriding reasons to do otherwise.
	+ You should not breach the 'duty of confidentiality' and not pass on identifiable data to third parties without participants' consent.
	+ 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.

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

* + Research participants should be made aware of this fact.
	+ 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](https://www.edgehill.ac.uk/research/governance/)).

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](https://ico.org.uk/media/for-organisations/documents/1061/anonymisation-code.pdf) 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

You should consult the Edge Hill [Policy on researching and handling sensitive material](https://www.edgehill.ac.uk/documents/policy-researching-handling-sensitive-material/) if your research involves material that is security-sensitive, radical or extreme.

You must adhere to Edge Hill IT Services’ [Acceptable use policy](https://www.edgehill.ac.uk/documents/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.

The University of Cambridge has produced a [video guide to managing sensitive research data](https://www.sms.cam.ac.uk/media/1122568).

# Metadata

Metadata is essentially data *about* your research data. It helps you to organise and archive your research data and helps other people discover, interpret, and use it.

Sufficient metadata should be recorded and made openly available via Figshare to enable other researchers to understand the potential for further research and re-use of the data. We provide more detail on this in [Appendix A: Metadata](#_Appendix_A:_Metadata).

# Documentation

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

The long-term preservation of your research data, so that it can be understood and interpreted by other users, 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.

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 Appendix A: Metadata).

# Storing your research data during the project

The IT Services wiki contains information on [storage, file sharing and collaboration tools](https://go.edgehill.ac.uk/display/itservices/Storage%2C%2Bfile%2Bsharing%2Band%2Bcollaboration%2Btools).

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

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.

You cannot use OneDrive to share your files with external collaborators so you should consider other options for the secure sharing of files. The first option you should consider is Figshare, which is Edge Hill’s digital data storage platform.

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 to make a decision.

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 (Pure).

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.

* + 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).
	+ 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

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 storage account could be compromised or closed, or you could lose a pen drive containing the data.

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.

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:

* + Backup and securely save electronic format research data.
	+ Digitise hard copy format research data at the earliest opportunity after collection (e.g. scan paper-based data).
	+ Photograph research data that includes physical artefacts.

While you can save secure backups of research data on appropriate mobile devices and pen drives, it is not necessary as the University maintains two copies of your material on its servers behind its secure firewall. Your data should be saved either on your Z:drive or on OneDrive. If you choose to save portable backups, you must ensure they are suitably secure (e.g. encrypted).

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

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

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.

The [UK Data Service](https://www.ukdataservice.ac.uk/manage-data/store/backup) offers advice on backing up your data.

# Storing your research data after the project

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.

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 Figshare (see ‘Research data management: metadata’, RO-GOV-15).

Using the [UKRI common principles on research data](https://www.ukri.org/funding/information-for-award-holders/data-policy/common-principles-on-data-policy/) as a guide, primary research data (and where possible/relevant specimens, samples, questionnaires, audio files, etc.) that underpins publications should be retained intact in accordance with legal, contractual, regulatory or ethical requirements, in an appropriate format and storage facility. This may vary according to the funder and nature of the data:

* + Some funders (for example, EPSRC) 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 Figshare pages.
	+ 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](https://mrc.ukri.org/research/policies-and-guidance-for-researchers/data-sharing/).
	+ Research data that informs national policy-making should be preserved in a research data repository permanently.

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.

You do not need to archive research data that does not underpin a publication. If you think you may wish to use this further, you should ensure your participants are aware of the possibility through participant information sheets with details of how long you will keep the data and for what purposes it will be used in future. All your data storage and archiving must conform to data protection legislation (including but not limited to GDPR).

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.

* + 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).
	+ 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.

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

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](https://www.sms.cam.ac.uk/media/1122623), or sensitive for security reasons.

You must specify and follow a process to handle and delete confidential data in line with the project's data management plan, and the University’s [Information security policy](https://www.edgehill.ac.uk/documents/information-security-policy/). The UK Data Service offers further advice on [data disposal](https://www.ukdataservice.ac.uk/manage-data/store/disposal).

When planning your research project, it is recommended that you conduct a research data risk assessment to determine:

* + The potential monetary value of the data (e.g. cost of collection – including your time – and cost of storage);
	+ The level of confidentiality required for the data;
	+ The steps required to provide appropriate data protection;
	+ The potential impact of unauthorised access to the data;
	+ Issues regarding access via Edge Hill’s network;
	+ Issues regarding access from outside Edge Hill’s network;
	+ Security of data while in transit (physically and electronically).

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](https://go.edgehill.ac.uk/display/itservices/Encrypting%2Byour%2Bfiles), including the equipment used to record interviews and transfer them to transcribers:

* + 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;
	+ Some devices allow you to use your fingerprint as the key (e.g. iPhone, some Android phones);
	+ 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;
	+ It is possible to encrypt pen drives using third party encryption tools but you can also buy drives with encryption out of the box;
	+ 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;
	+ The [UK Data Service](https://www.ukdataservice.ac.uk/manage-data/store/encryption) recommends the use of technology meeting Pretty Good Privacy (PGP) standards (other encryption technology is available).
	+ While working away from campus, you should follow [the University’s advice on remote working](https://go.edgehill.ac.uk/display/compliance/Working%2BRemotely).

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:

* + 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).
	+ Password protecting the dataset file and sending the password separate to the protected file.
	+ Multi-factor authentication should be used by all collaborators across the university.

# Useful links

## Policies & guidance

* [Edge Hill Acceptable Use Policy](https://www.edgehill.ac.uk/documents/acceptable-use-policy/)
* [Edge Hill Data Protection Policy](https://www.edgehill.ac.uk/documents/data-protection-policy/)
* [Edge Hill Information Security Policy](https://www.edgehill.ac.uk/documents/information-security-policy/)
* [Edge Hill Research Data Management Policy](https://www.edgehill.ac.uk/documents/research-data-management-policy/)

## EHU resources

* [Figshare (EHU research data repository)](https://www.edgehill.ac.uk/services/figshare/)
* [Edge Hill Research Governance and Ethics web pages](https://www.edgehill.ac.uk/research/governance)
* [Edge Hill Information Governance wiki](https://go.edgehill.ac.uk/display/compliance/Home)
* [Edge Hill Storage, file sharing and collaboration tools wiki page](https://go.edgehill.ac.uk/display/itservices/Storage%2C%2Bfile%2Bsharing%2Band%2Bcollaboration%2Btools)

## Other resources

* [Digital Curation Centre (DCC)](http://www.dcc.ac.uk/)
* [UK Data Archive](http://data-archive.ac.uk)
* [UK Data Service](https://www.ukdataservice.ac.uk/)

## Contacts

* For further assistance on research data management in general, please contact the Research Office.
* If you have questions about the University’s research data repository, please contact Learning Services

#

# Appendix A: Metadata

## What is metadata?

Metadata is essentially data *about* your research data. It provides information on who, where, when, why, and how data was collected, access conditions, and terms of use of a data collection.

Metadata supports the management, discovery, understanding, and re-use of data. Funders and publishers have their own policy requirements governing metadata. Providing standard metadata means end-users can correctly cite and attribute your research data, as they would with research outputs. For example:

*Richardson, Elizabeth A. (2009). Carstairs deprivation scores for Scotland by CATT2, 1981, 1991, 2001 [Dataset]. University of Edinburgh. School of GeoSciences.* [*http://hdl.handle.net/10283/19*](http://hdl.handle.net/10283/19)*.*

## Types of metadata

Broadly, metadata shows how the data is generated and managed. The most common types include:

* Provenance metadata: including when and where the data was collected, and by whom.
* Rights and access metadata: information on rights and access usage rules.
* Structural metadata: information for a person or computer to read the data (e.g. data formats).
* Preservation metadata: information that allows the long-term use of the data, including what software has been used to access the data.

Research data repositories require specific metadata, which would be impractical to list here in detail. However, some examples of key metadata you will normally need to provide are:

* Title of the dataset;
* Author/creator names;
* Keywords that would help the dataset appear in search results;
* A summary description/abstract for the dataset, detailing how and why it was collected;
* The relationship of the data to other resources e.g. to a related research article on the research repository;
* Research funders (may be internal or external);
* The start and end dates of the period covered by the data;
* Accessibility status e.g. can the data be used by those with visual impairments, etc.?
* Licence for re-use e.g. Creative Commons CC-BY

## Metadata requirements

**Regardless of where the research data itself is stored, you must still create a record containing its metadata in Figshare.**

Funders may have their own research metadata requirements which exceed Edge Hill’s, and they may require you to store your research dataset in their own repository. Please familiarise yourself with your funder’s policy on research data management. Alternatively, you may store the research data in a subject-repository instead of Edge Hill’s.

## Data access/availability statement

The data access statement provides the end user of your research data with the information necessary to be able to make sense of the dataset(s). There is no template for this but such statements may include:

* Access restrictions – should the research data be available only to certain groups? If so, what is the reason and how can they request access?
* Data creation or observation conditions – which software or survey instruments were used to collect the research data?
* Code label and field descriptors – to what do the research data variables refer? Can someone with knowledge in the same field make use of the research data using the information provided?

# Endmatter

|  |  |
| --- | --- |
| Title | Research Data Management Guidelines |
| Document Owner | Director, Research Office |
| Date for Review  | February 2025 |