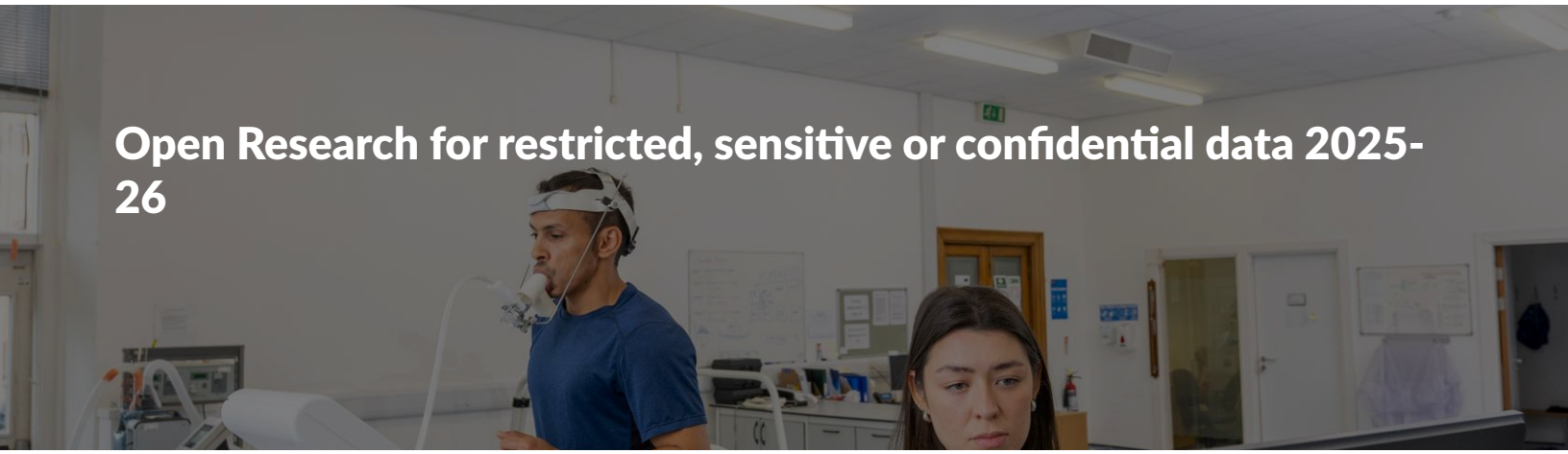


# Open Research for restricted, sensitive or confidential data 2025-26

*Welcome to our 'Open research with restricted, sensitive or confidential data' course. This course is intended for anyone who will be working with research data and information that has characteristics which mean you need to consider who should have access to it during and after your project. This might encompass any kind of data, for example, interviews, photographs, and oral histories to medical records, experiment results, and statistics.*

*This course is not intended to replace the Data Protection, Trusted Research or Research Integrity training courses or other essential training; rather, it is a supplement to help you think about best practice in data management and open research when working with restricted, sensitive or confidential data.*

**Overview**

This course should be taken in conjunction with the training on Data Protection and Research Data Management offered by the University of Glasgow. If you exit the course, your progress will be saved so you can take your time and return to it. We have put together a document containing key terms and helpful resources for you to use alongside this training, and to have for future reference.

[Link to the document here.]

If you have any concerns or questions, please contact the Research Data Management service at research-datamanagement@glasgow.ac.uk.

*Lesson 1 of 6*

# Introduction

## What is covered in this course?

What is sensitive research data?

Policy and legal compliance while working with sensitive research data

Good practices while working with sensitive research data


Preparing for the end of your project

## What is restricted, sensitive or confidential research data?

In the course of your research you may find yourself working with data that has characteristics which mean you need to consider who should have access to it during and after your project.

These kinds of data may emerge from any research area, but more often occur in research involving people or animals, research falling under the UK Government's Trusted Research agenda, research subject to data protection legislation or ethical approval, or research using data which is provided to you by a third party who sets restrictions on its use.

This covers a lot of different kinds of data! There are several definitions to keep in mind, and different sources of authority use overlapping terms about data to mean slightly different things.

Restricted, sensitive, or confidential research data is produced and handled in every area of research across the University of Glasgow. Research policy at the University is outlined in the Code of Good Practice in Research.

## As open as possible, as closed as necessary

At the University of Glasgow we encourage researchers to recognise that just because the research data needs to be restricted while research is ongoing, this should not often prevent you from depositing the data in an appropriate repository, and does not mean that all the data you deposit must be closed or subject to access controlled. In many cases, datasets arising from research using restricted, sensitive or confidential data can produce datasets which are at least partially open access.

- We support **transparency, openness, verification, and reproducibility** by facilitating early and open sharing of research data, software, code, methods, preprints, open educational resources and materials with a wide range of audiences.
- We place value on a wide range of different research output types across the whole lifecycle of research – improving value to the public as well as to other researchers.
- Our policies recognise that it is not always possible to share data where there are sensitivity concerns. A good maxim to work to is 'as open as possible, as closed as necessary': apply the appropriate restrictions to your data and other outputs, but where it is possible to do so, plan to make them available for scrutiny and reuse.

**Reproducibility and transparency**

The University of Glasgow joined the UK Reproducibility Network in 2020, and we recognise that while experimental reproducibility remains a gold standard for open research, actions that support transparency in research and scholarship vary considerably across disciplines and methodologies.

We expect researchers to pursue transparency through the most effective and appropriate means, according to the nature of their research.

"Working reproducibly means others can check your results - even early on in the research process. Thus, the full analysis and methodology is transparent. Scientific results and evidence are strengthened if they are reproduced and confirmed by several independent researchers."

(Although the quote above uses the term 'scientific', it is meant here in the European sense - to encompass all research. See https://www.nature.com/articles/s41599-018-0149-x for more on reproducibility in SHAPE research.)

**Should I really share sensitive research data?**

It is tempting to withhold all sensitive data to avoid sharing something you shouldn't. Openness, transparency, and reproducibility can be more time-consuming to plan for when your research involves restricted, sensitive or confidential data, and the impulse can be to avoid sharing data or other outputs in case there are any legal or ethical implications.

However, there are good reasons to avoid taking the 'easy' route and keeping such research hidden. For example, you can increase the reach and impact of your research, allow others to build on your research and amplify other voices and contributions - for example those of your research participants.

We want to promote the safe sharing of data wherever possible, and in this course we show you how best to plan for handling restricted, sensitive or confidential data, and how to check in what ways it can be shared.

Colleagues at the University of Glasgow discussing their research outputs.

CONTINUE

# 1.What considerations apply to your research data?

## Types of restricted, sensitive or confidential research data

You need to consider who should have access to your data during and after your project if any of the following apply:

- The research is subject to ethical approval
- The research will collect or process any personal data
- The research is subject to Trusted Research Legislation
- Outputs from the research are in consideration for commercialisation or there are other IP implications
- The research will use third-party data which comes with conditions on its reuse (eg data subject to a confidentiality clause, or any other restrictions on sharing outputs )
- The research will use or produce data which needs to remain confidential for any other reason

**Remember, some terms have technical definitions - for example, 'sensitive data' has a specific legal meaning in the context of UK data protection legislation - but they are often used in a non-technical sense in literature discussing research data more broadly.**

## Key definitions of data

We mention various different forms of data throughout this training, so read below and use the flashcards to learn the key definitions. Remember, these will be available in the accompanying handbook too.

1

## What is research data?

Research data is **any information, whether digital or physical, required to underpin research.** It is difficult to define what constitutes data for any given project, and it will vary by discipline or approach. As an expert in your field, you have a lot of leeway to determine what data you think are essential for others to understand and build upon your research.

A good starting point is to think: what information would someone need to see to understand how I reached my conclusions? For different disciplines, this may include raw data captured from instruments or collection systems, derived data, documents, spreadsheets and databases, research notebooks, visualisations, models, software, images, measurements, and numbers.

2

## What is personal data?

Personal data is defined in law, as part of the General Data Protection Regulation (GDPR), which governs data protection in the UK and EU, as **any information relating to an identified or identifiable natural person, whereby the person can be identified, directly or indirectly**. This may, for example, include photographs, email messages and data recorded by closed-circuit television (CCTV), if a person can be identified from this.

It also includes data identified by reference numbers, where a separate list can be used to match the reference numbers to named individuals. It, however, **does not** mean that all information provided during research by a person (e.g. during interviews) is personal data. If a person cannot be identified directly or indirectly from the information, then the information is not defined as personal data.

Source: UK Data Service The Data Protection Act and GDPR — UK Data Service

https://zenodo.org/doi/10.5281/zenodo.11147886

**Remember that this means that pseudo-anonymised data, in which participants are identified by pseudonyms, is considered personal data, until you delete the list that matches pseudonyms to named individuals!**

### What is metadata?

Metadata is **additional information that accompanies the digital resource to ensure that it can be accessed and understood over time**, such as keywords used as tags to enable accurate searching and retrieval. Examples include the creator of the data, the creation date and time, the size of the data file, the file type (.docx, .pdf), etc.

(Source: University of Glasgow Digital Preservation Policy v3.8)

4

### What is confidential research data?

Confidential research data is data that must be treated as confidential - for example **personal data (**confidential data may or may not be personal data; however most personal data will be confidential), or **data that is provided to you under a confidentiality agreement** - for example if it comes from an industry partner.

If your data are explicitly subject to confidentiality restrictions, either because they are considered confidential information or because they are classified as medium or high risk, they are subject to the University of Glasgow Policy on Confidential Data.

## Test your knowledge:

What is research data?

Research data is any information, whether digital or physical, required to underpin research. A good starting point is to think: what information would someone need to see to understand how I reached my conclusions?

1 of 3

What is metadata?

Metadata is additional information that accompanies the digital resource to ensure that it can be accessed and understood over time, such as keywords used as tags to enable accurate searching and retrieval.

2 of 3

What is confidential research data?

Confidential research data is personal data you're not allowed to share or data that is otherwise under a confidentiality agreement - for example if it comes from an industry partner.

3 of 3

If you are in doubt about what you can do with your data, you can contact us at research-datamanagement@glasgow.ac.uk.

CONTINUE

# 2. Policy and legal compliance for research data

You are encouraged to make your research, and your data, as **open, transparent, and reproducible as possible**, within the appropriate **legal, institutional, and ethical frameworks**.

1

## University data policies

The University's policies regarding research data appear in the Code of Good Practice in Research. The guiding principle is that data should be as **open as possible and as closed as necessary**. You should plan in advance to make data available for reuse where it is possible to do so.

We make our data available wherever possible in order to;

- Help others to avoid duplication of effort
- Encourage new collaborations between creators and users of data
- Increase the impact and visibility of our research (e.g. more citations)
- Meet the aims of national assessment exercises like the Research Excellence Framework (REF).

Making data and research as open as possible means considering the benefits to others when sharing your work, and making sure that ethical, statutory or other policy issues are taken into account when doing so. It is crucial you familiarise yourself with the relevant policies and support to make informed choices about your own research practice. While we encourage you to make your data as open as possible, you should do so in line with

existing policies. This means sharing restricted, sensitive or confidential data only with proper approval where it is required.

Reproducibility is not always achievable, but you should always aim to make your research as transparent and accountable as possible, even if you can't deposit or share all your data.

## Research Culture Priorities

Our five research culture priorities were chosen following extensive consultation with the University of Glasgow research community between 2015 and 2021 and focused purposefully on issues that are specific to the way we do research, and the way we support research careers. The five priorities are **Research Recognition, Collegiality, Research Integrity, Open Research,** and **Career Development**. These also underpin our policies relating to research.

### Find out more at our webpages:

**Conduct and Research Integrity at the University of Glasgow**

RESEARCH INTEGRITY

**Open Research at the University of Glasgow**

Open Research information and policy at the University of Glasgow

OPEN RESEARCH

**Code of Good Practice in Research at the University of Glasgow**

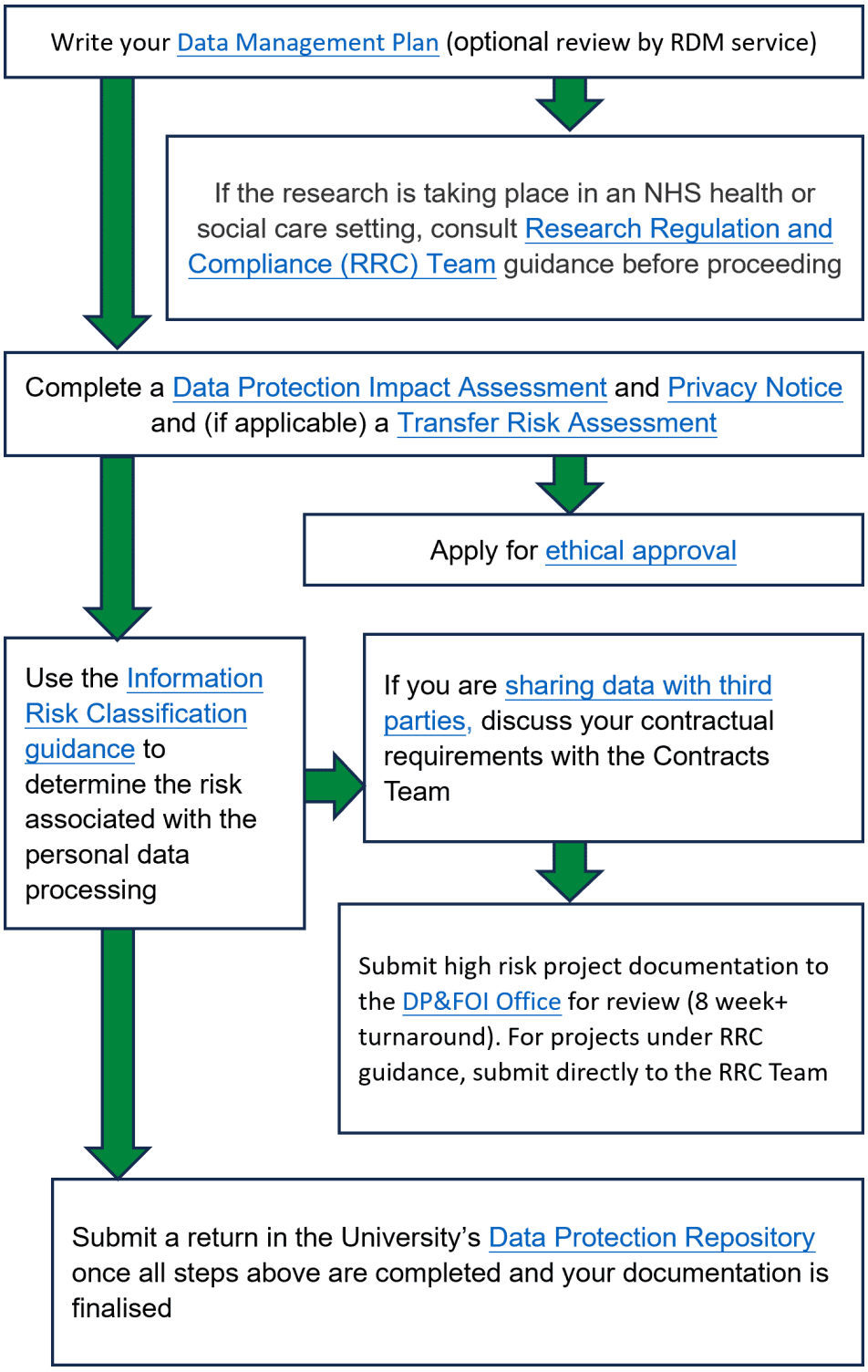
CODE OF GPR

### Deposit all data of long-term value

In order to ensure the integrity of your research, you should make sure a definitive copy of your research data is deposited at the end of the project. Before you consider which parts of your data can be shared you should consider what data should be deposited, even if it would need to remain restricted. What would someone need to see in order to confirm the integrity of your research?

2

## Project initiation workflow




Summary workflow diagram- https://doi.org/10.36399/gla.pubs.202746

---

The workflow shown here outlines the preferred process for initiating a research project involving personal data.

Projects involving personal data will need ethical approval, and you should seek consent for all the planned use of the data. The best way of planning for this is to get ready for your application for ethical approval by at least drafting

your Data Management Plan, Data Protection Impact Assessment and Privacy Notice. If you plan to share personal data outside the University, a Data Sharing Agreement should be in place before you transfer any data. **Remember that it can take some time to review and amend these documents.**

The key contacts for the project initiation process can be found in the Project Initiation Workflow document: https://doi.org/10.36399/gla.pubs.202746

3

## Ethics

Many research projects that involve restricted, sensitive or confidential data will require ethical approval, and the University has plenty of specific advice you should seek out when starting out your research. Research ethics will intersect with everything you do as a researcher, so it is important to get comfortable with the topic and the relevant policies as soon as possible.

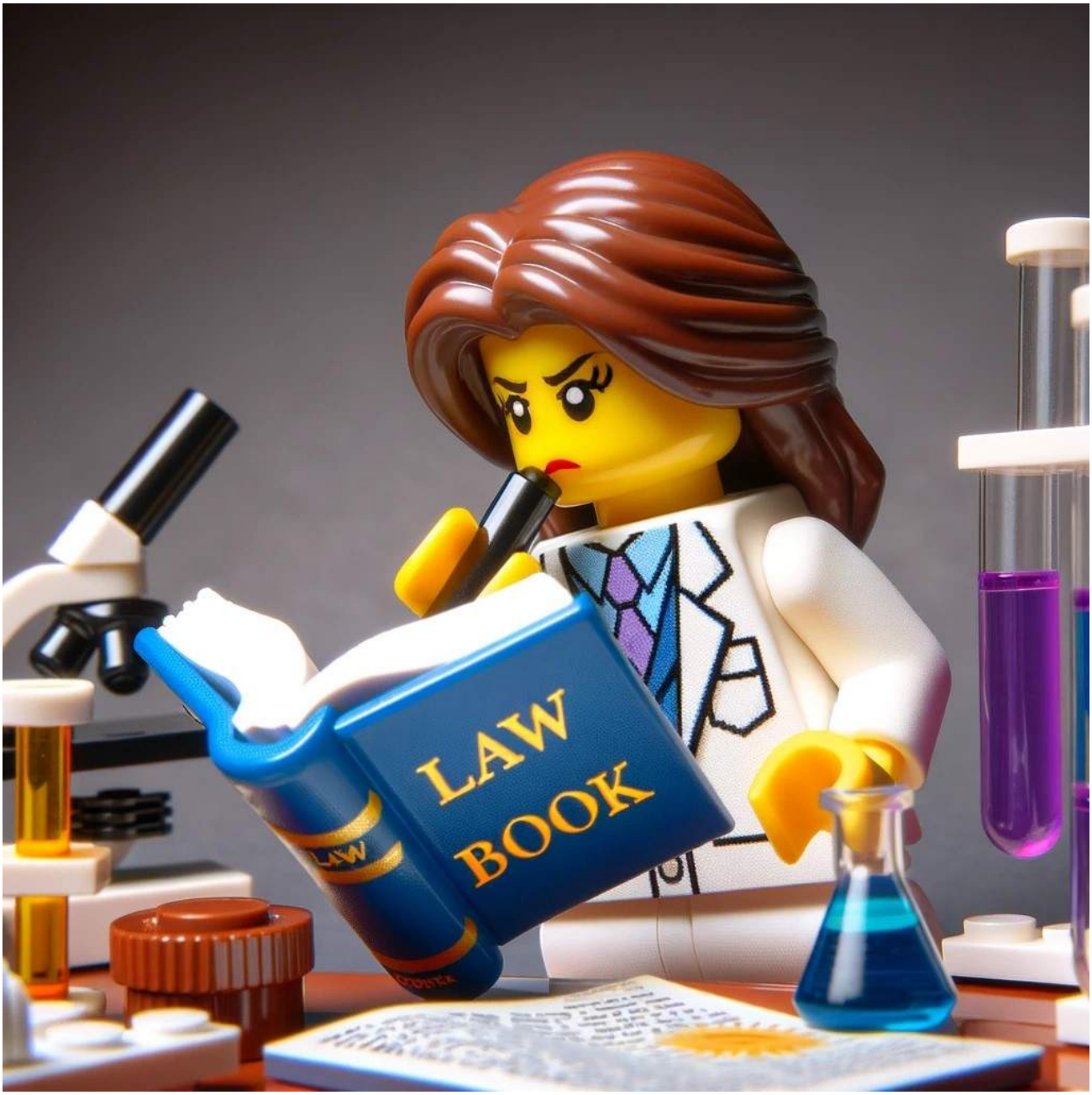
**Remember: All research involving humans, or human material, or animals has to go through an ethical review process BEFORE research can start. If you're not sure whether your project requires ethical approval, contact your College or School ethics officer in the first instance.**

You should seek clear and unambiguous ethical approval and consent for all planned uses of your data, including the long-term storage and sharing of the data.

Please look at section 5 of the Code of Good Practice in Research for information on ethical, regulatory and compliance processes for research approval for different kinds of research.

There are a range of ethics committees within the University, depending on your area and whether your research is clinical. An overview of ethics policies can be found here.

To support you in understanding ethical and other integrity issues, the University offers Research Integrity training, and the College of Social Sciences offers research ethics training for students, which is a great place to start.




**Section 7.4 of the Code of Good Practice in Research explains the University's expectations regarding your data. You can find these documents on the Policies page Our policies - Code of Good Practice in Research**

## Trusted Research

The UK Government's "Trusted Research" campaign aims to "support the integrity of the system of international research collaboration, which is vital to the continued success of the UK's research and innovation sector" (NPSA). It is particularly relevant to researchers in STEM subjects, dual-use technologies, emerging technologies and commercially sensitive research areas. More recent legislation included a list of "17 Sensitive Areas of the Economy" and this can indicate which research may be in scope.

Legislation such as Export Controls and the National Security and Investment Act mean that researchers need to consider carefully what they share with overseas partners, particularly those in countries where the UK Government has sanctions or arms embargoes (there is a list). Anyone working on controlled technology (UK government list) or where there is a clear military application (e.g. UAVs or drones) or risk of misuse (e.g. virus research, facial recognition technology) needs to take care in both how they manage their research data (particularly if travelling overseas) and how they share it.

There is a separate Trusted Research training course available that only takes 15mins which is a good first step if you have any concerns about this applying to you. Support is available from the Research Governance and Integrity Team in Research Services via the university helpdesk (search "Trusted Research"). We can help you risk assess your research data and ensure compliance

## Clinical research

If you are working with clinical research data and information, please contact the Research Regulation and Compliance Office who can provide advice and guidance on matters relating to any clinical research involving human subjects, tissue or data. The NHS has information on research planning available here.

Clinical projects involving participants selected because of their links with the NHS-should be submitted to a relevant Research Ethics Committee (either an NHS or University Ethics Committee). In addition, all studies involving human tissue (including blood samples) or data from NHS patients may also require review by a NHS Research Ethics Committee.

Clinical Research Projects may also require further approvals to take place in the NHS, for example NHS R&D Management approval, however the Research Regulation and Compliance Office will be able to advise on the relevant processes.

There is information for research involving humans available from Research, Innovation & Engagement Support, including:

- Clinical Trials
- Research Studies
- Integrated Research Application System (IRAS) Guidance
- Healthy Volunteer Studies
- Insurance and Indemnity
- Responsibilities
- Human Tissue
- Laboratories
- Student projects
- Policy on Sponsorship and Co-Sponsorship Arrangements with NHS GGC

**Data Protection**

The Introduction to GDPR training course is a requirement for all University of Glasgow staff and PGR students, and recommended for any students working with human data.
For advice on GDPR, contact the Data Protection and FOI Office in the first instance: Tel: 0141 330 6494, email: dp@gla.ac.uk, web: www.gla.ac.uk/services/dpfoioffice

**GDPR Principles and laws will apply to Personal data**, which is any information relating to a natural person who can be identified, directly or indirectly, by that information

- Name
- Identification number
- Location data
- Online identifier
- Pseudonymised data
- Factors specific to physical, physiological, genetic, mental, economic, cultural or social identity

There are also special categories of personal data, which involve personal data relating to:

- racial or ethnic origin
- political opinions

- religious or philosophical beliefs
- trade union membership
- genetic or biometric data processed for the purpose of identification
- health
- sex life or sexual orientation
- If any of the above examples of personal data or special categories or personal data are used in your research, please speak to dp@gla.ac.uk about ensuring GDPR compliance.

**Any security measures you put in place for your data should be proportionate to the risks in the data.**
·Consider your ethical approval and GDPR requirements: don't share any personal data without permission and adequate/proportionate safeguarding!
·Who will need to access the data during your project? You should plan for your data to be accessible to at least yourself (and your supervisor if you're a student), but there may be others who need to access it.
·Think about what will happen to your data in transit – is it secure?

A helpful guide to the law when working with personal data under GDPR can be found from the ICO (Information Commissioner's Office) here, which has a useful checklist to use for checking your compliance with GDPR.

Every project collecting or processing human data needs a **DPIA (Data Protection Impact Assessment).** The Data Protection team will need to review your DPIA if the processing of the personal data is high risk (for instance if you have any special category data or a very large volume of personal data). You can use the DPIA decision tool and Information Risk Classification to work out whether your DPIA needs to be reviewed.

## IP and Commercialisation

You may find that during the course of your research you produce some intellectual property (IP) that you would like to secure. Our advice for protecting your intellectual property rights is to:

1. Keep good records and gather evidence and a paper trail. Record the IPR in detail at the time it is created.
2. Don't disclose your intellectual property (e.g. in a publication or seminar) until it is evaluated. Use confidentiality agreements if talking with external parties.
3. Act quickly and decisively and consult a specialist immediately. They will help you to consider the options and costs.

Consult the IP and Commercialisation office: https://www.gla.ac.uk/myglasgow/ris/ieed/

**IP OWNERSHIP AND STUDENTS** (source: RESEARCH AND INNOVATION SERVICES Ownership of IP)

1. **Student-owned IP**: The University does not automatically own intellectual property developed by students. Students will generally own the intellectual property they develop during the course of their studies unless IP ownership is governed in some way by a third party agreement. Examples include research contracts, studentship agreements, and funding agreements.
2. **Students and invention disclosure obligations:** Invention disclosure applies where something new and useful has been conceived and developed, where the IP may need protection and/or where the invention, technology, software or multimedia product has commercial potential. It does not apply to literary works, musical creations, or works of art.

Your supervisor and the IP & Commericalisation team should be able to help if you think this may apply to you, but remember to keep your data and information secure until after discussions with the team to prevent theft of IP. You can find the necessary contact details here: Meet the IP & Commercialisation team.

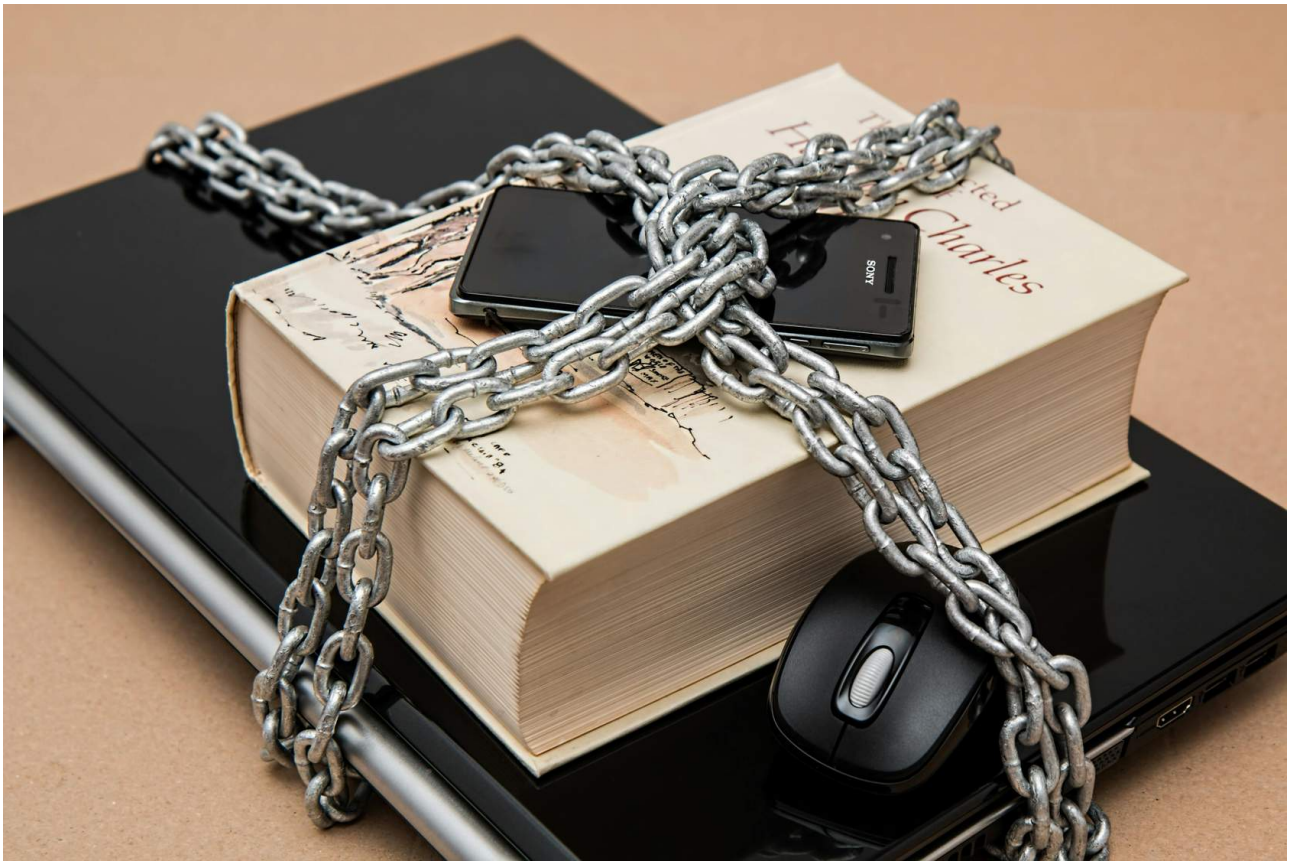
**Commercialisation**

- Research and Innovation Services - IP & Commercialisation "The IP & Commercialisation team within R&I provide a University-wide service to help protect and exploit our intellectual property through licensing and spin-out company formation. The team work closely with Colleges to select high-value opportunities and work with academic teams to develop and execute agreed exploitation plans."


**Third Party Data**

Any data owned or provided by third parties that you handle in the course of your project is likely to be subject to restrictions on its use. Even open access data are governed by the terms of the licence that they're shared under - for example, a Creative Commons CC-BY licence mandates users to credit the original creators.

If you're not sure about the ownership of your data you can contact the Research Data Management team in the first instance.




**Confidential data**

If your data are explicitly subject to confidentiality restrictions, either because they are considered confidential information or because they are classified as medium or high risk, they are subject to the University of Glasgow Policy on Confidential Data.

**Confidential Information**

This is information created in a confidential setting or disclosed under a duty of confidence, or any other information, which by its nature, should reasonably be considered as confidential including, without limitation, Technical Know-how and Trade Secrets.

**Contact compliance-support@glasgow.ac.uk if you think this might apply to you.**

**Medium or high risk data**

Confidential data are those that are classified as either medium or high risk in the University's Information Risk Classifications. Confidential data may or may not be personal data; however most personal data will be confidential. This policy applies to all confidential data used as part of University work, or stored or processed on University systems, and applies to data in electronic and other formats eg. paper.

**Confidential Data Best Practice in Brief**
Source: IT - Information Security - Confidential data

- Confidential data should be stored in central filestores, on secure servers maintained in secure physical environments
- Confidential data should not be held on local disk storage (e.g. the C:\ drive of a desktop machine)
- Confidential data should not be stored or accessed on mobile phones or tablets unless appropriate security measures are in place

Confidential data must be encrypted when:

- Stored on a laptop
- Stored on a memory stick or portable hard drive
- Stored or exchanged on other portable media such as CDs, DVDs
- or exchanged with external organisations or individuals

—

CONTINUE

# 3. Good practice when working with restricted, sensitive or confidential research data

The Introduction to Research Data Management training course covers general best practice for research data management. Below is an overview of some good practices specific to restricted, sensitive or confidential data.

## Get your ducks in a row!

The project initiation phase, and good data management planning, is essential when working with any research data, but is especially important if your data is subject to any of the considerations above.

Some of the processes you need to complete to get started when you're working with data can be time-consuming but are essential to get right. Leave yourself enough time to get your project set up correctly and you will find it helps you in the long run.

Our detailed Project Initiation Workflow contains all the information to get you started on planning a project working with personal data.

Guidance on data management planning is available on our Open Research web pages: https://www.gla.ac.uk/myglasgow/openresearch/researchdatamanagement/beforeyoubegin/

# Working with personal data

This guide from the UK Data Archive is full of valuable best practice guidance, including information on handling sensitive data. Our best practice advice is:

- Aggregate your results wherever possible

- Avoid collecting personal data if you don't need it for your research question
- Remove personal information if you don't need to retain it
- Keep a log of what types of personal information you have removed.
- Assess whether you have sufficiently de-identified your data that it can be considered anonymous.
- If your data cannot be effectively anonymised, restrict access.
- Check with the Data Protection team if you need reassurance that data are sufficiently de-identified that they can be shared.

## Organising your data

Keeping your data well organised will help you to understand which parts are sensitive. The Introduction to Research Data Management course, which is mandatory for PGRs and recommended for any researcher who wants to familiarise themselves with the University's approach to data management, describes good practice in organising, naming and handling your data. We also have a number of research data management guides covering these topics.

## Storing your data

We recommend that you always use the University's authenticated storage spaces to store your research data while you work on it, and this is especially important if your data need to be restricted during your research project. Below are examples of these storage spaces. You can always check with the Research Data Management team or the Information Security team if you are concerned about who might have access to your data.

Microsoft OneDrive/Teams.

Every member of the University has 1TB of storage space in their OneDrive account. OneDrive/Teams is suitable for most research data. OneDrive is appropriate for solo research, while Teams should be used for projects with multiple collaborators. Both OneDrive and Teams use Microsoft Sharepoint storage, contracted by the University. Log in through https://office365.gla.ac.uk.

The shared drive (usually J:\)

The University's shared drive is also suitable for most research data. Some departments have their own local servers that are also suitable. Contact your local IT support for information on setting up a shared folder that is accessible only to those who need to have access to your data.

| | |
|---|---|
| NextCloud | This is an encrypted transfer and storage tool that is installed on University servers. IT Services can provide an on-campus cloud service for people who cannot use Teams, or have local data protection requirements. It provides a simple cloud file-sharing system, accessible through your web browser or through a desktop sync app. Students will need their supervisors to arrange access via the IT Helpdesk. |

More information on the University's storage is available in this IT Services guide. If none of these options look suitable to you, or if you expect to collect more than 1TB of data, contact your local IT support.

The University's Confidential Data best practice guidance provides advice on the storage of confidential research data.

**Free personal cloud storage must not be used for confidential data.** Paid personal cloud storage is discouraged as the location of the servers and the ownership conditions vary between providers and are not always obvious. Common examples of these kinds of storage include Notion and Google Drive.

This means that you **should not store any personal data somewhere if you're not sure what territory the data will be kept in.** If you're not sure whether a storage area or tool is compliant with data protection requirements, you can check with the Data Protection and FOI Office at dp@gla.ac.uk.

https://ico.org.uk/for-organisations/uk-gdpr-guidance-and-resources/international-transfers/

Researchers travelling outwith the European Union should take care not to download any personal data to the local drives of their computer, as this would constitute taking personal data outside of the EU. Find our advice here.

The University Information Security team has provided extensive guidance on storing confidential data. Confidential research data is personal data or data that is otherwise under a confidentiality agreement - for example if it comes from an industry partner.

In brief, you should:

- Use central filestores on **secure servers** maintained in **secure physical environments**
- Confidential data should **not** be held on **local disk storage** (e.g. C:\)
- Confidential data should **not** be **stored or accessed on mobile devices** unless adequate security measures are in place.

Confidential data must be encrypted when:

- Stored on a laptop, phone or tablet
- Stored on portable media (e.g. memory stick, external hard drive)
- Exchanged with external organisations and individuals

Guidance on data storage during your project: https://www.gla.ac.uk/myglasgow/openresearch/researchdatamanagement/datastorageforresearchinprogress/

Take a look at the guidance on handling confidential data: https://www.gla.ac.uk/myglasgow/it/informationsecurity.

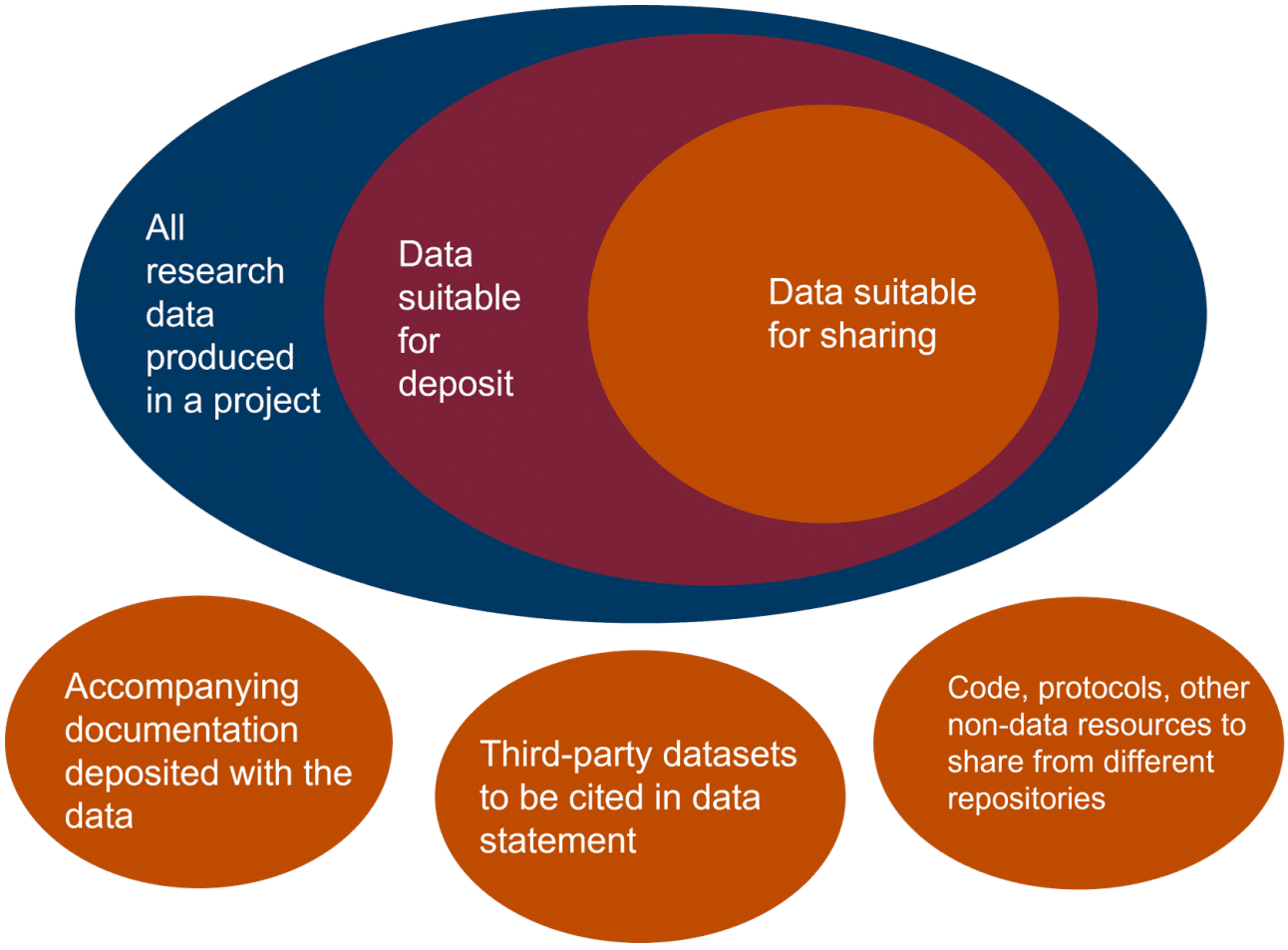
CONTINUE

# 4. Preparing for the end of your project

## Selection and appraisal

It is not usually necessary to keep all the versions of data that you produce during a project. You should consider your obligations for data retention - these might come from the University's or a funder's policies, from a collaboration agreement between partners, or from your ethical approval. You may also have statutory (legal) obligations - for example, certain types of medical data need to be retained for a statutory period.

It is not all-or-nothing - you may decide that while parts of your data cannot be retained, others can. For more information on data appraisal and selecting what to keep, see our guide: https://edshare.gla.ac.uk/1385/



A venn diagram showing the different types of data you might produce during a project, and accompanying outputs.

---

**If you have any legal or ethical reasons to destroy all or part of your data, you need to make sure you do so properly. Having your data well organised and documented will help you to identify any of your data that you cannot keep.**

## Retention of restricted, sensitive or confidential research data

As per the Code of Good Practice in Research, data of long-term value should be retained for at least ten years from the conclusion of the project.

Having restricted, sensitive or confidential research data does not necessarily mean that it needs to be destroyed at the conclusion of your research, and in fact the majority of data produced at the University of Glasgow is suitable for retention.

For information on retention of research data with regard to data protection, please see: https://www.gla.ac.uk/myglasgow/dpfoioffice/guidanceforstaffandstudents/research/retentionofresearchdata/

**Do not promise your participants that you'll destroy your research data at the conclusion of your project unless it's absolutely necessary - you should plan to retain data of long-term value wherever possible.**

## Repositories for restricted, sensitive or confidential data

The University requires that **all your data of long-term value be deposited in an appropriate repository,** where it will be stored for at least ten years after the end of your project. You should make your data available for sharing wherever it is appropriate to do so.

A **repository** is somewhere that stores and manages files. Repositories usually accept a *definitive version* of a dataset, and they are usually *open*, in the sense that they have a public list of their contents, even if the content can't all be freely accessed.

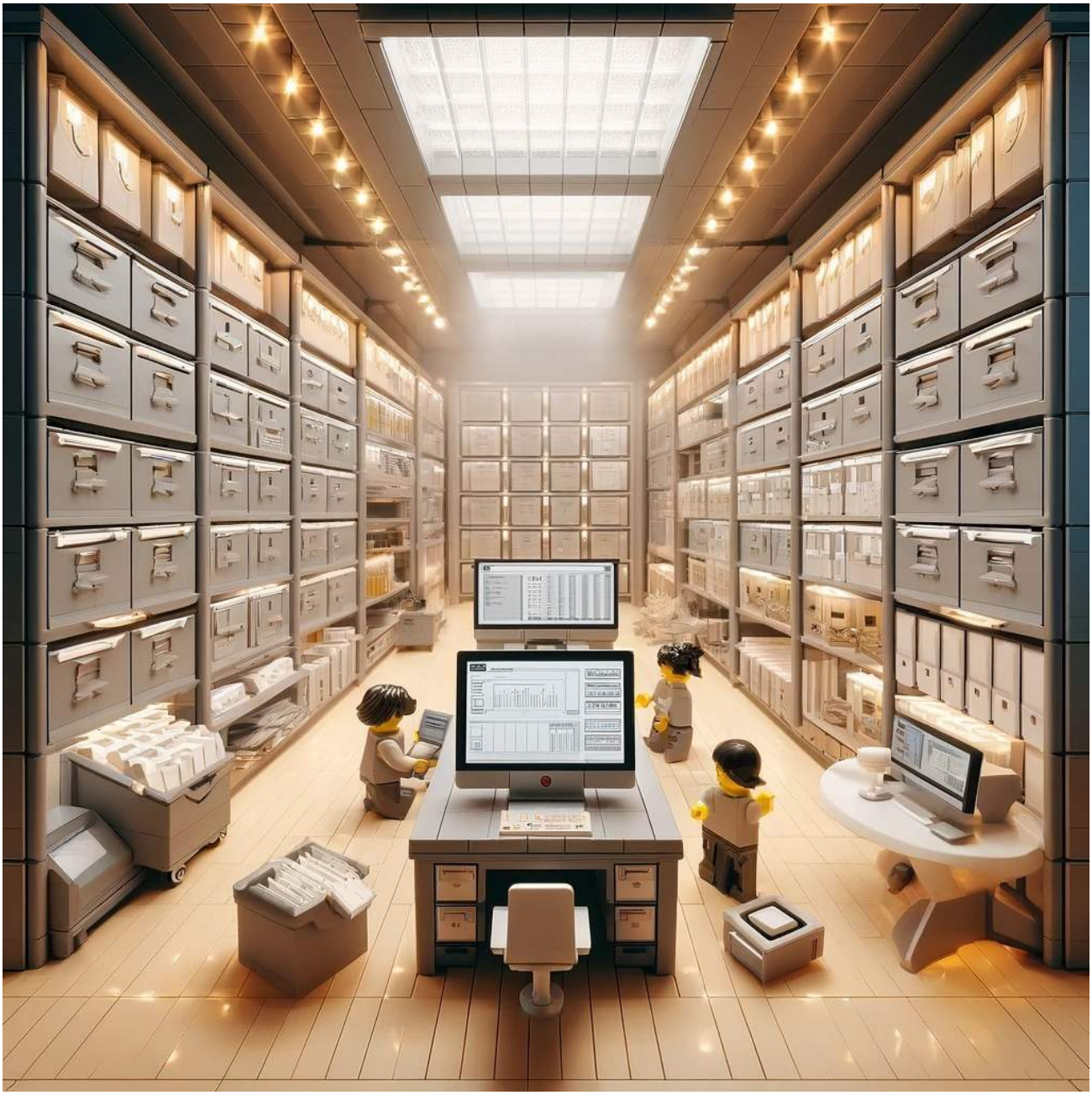
> Do you need to keep personal data, or is your de-identified data enough to show how you reached your conclusions? *Removing personal data is a good way to reduce the risk of accidentally disclosing personal information.*

Most researchers retain their raw data, the final processed version, and any intermediate processed versions that someone would need to see to understand how the dataset was produced.

**If your data was provided by a third party, or if it is under third-party copyright, you may not have permission to keep it after the project ends**. In this case you should make sure that your documentation is good enough that someone who wanted to access the data would know where to find it.

If your dataset is going to be restricted (for example if it contains confidential data), or if it is too large to easily download, you might want to make a small (shareable) sample available for someone to explore before they request access to the whole dataset.

> *If you have any legal or ethical reasons to destroy all or part of your data, you need to make sure you do so properly. Having your data well organised and documented will help you to identify any of your data that you cannot keep.*



https://doi.org/10.5281/zenodo.11147887

Please see our Introduction to **Research Management Training** for more information on finding the correct licence and repository for your data after your project.

When you put your data in a repository, you will need to give it a **licence**. A licence tells potential users of the data what they can do with it. When you're planning to deposit data you can start by dividing it into two categories: data that can be *open access*, and data that should be *controlled*. You should take into account the preferences of your funders and any project partners when choosing licences for your data.

Open access data is that which is freely available for other people to reuse, with no barrier to entry. A user will not need to complete a licence, register an account or pay any fee to get open access data. You should make your data open access unless you have a compelling reason for restricting it. The most common open access licences are **Creative Commons**, which are identified by their grey logos. The default licence for open-access data in the University's repository is CC-BY, which means that users must attribute the original author of the material but are otherwise free to re-use it as they wish.

Controlled data is that which cannot be accessed without some additional procedures like completing a user licence or submitting an access request. In some cases, users requesting access to restricted data will need to submit a new application for ethical approval.

Closed data is that which is deposited for long-term storage but where it is not possible to provide access to outside users. These data are usually retained for integrity purposes only.

If you think you have some data that needs to be controlled or closed, you should look at the policies of your chosen repository to see what options are available.

While some repositories only accept open access data, others can handle a variety of different access conditions. For example, our institutional repository is called Enlighten: Research Data. The repository has a range of access options for data, including open access, controlled, closed and dark data: https://www.gla.ac.uk/myglasgow/openresearch/researchdatamanagement/enlightenresearchdatainformation/#accessoptions

All staff and students can deposit research data in Enlighten, but might choose a different repository if it's more appropriate. Another good example of tiers of access can be found on the UK Data Archive.

Enlighten: Research Data meets a high level of information security which means that we can store very sensitive data, provided you have permission to deposit it. Some deposit charges may apply; contact us if you think that your dataset will be larger than 100GB.

You can also make deposits with mixed access conditions, depending on the affordances of your dataset. For example, if some of your participants consent to open access data sharing and others refuse, but are happy for their data to be retained, you could deposit all of your data to Enlighten: Research Data, with some of it closed and some open access.

You're almost done with this course. Let's summarise what you've learned by reviewing some key takeaways before you

take the quiz.

CONTINUE

# Summary

You have now read through the entire course, and hopefully feel more confident in working with restricted, sensitive or confidential research data. You can come back to this training and the Introduction to Research Data Management course at any time if you have questions. We would recommend that you save or bookmark a copy of the **handbook that accompanies this course** so that you can refer back to it throughout your research.

Remember that you can contact the Research Data Management team anytime with questions relating to your research data, open access for publications and other research outputs, working with data, and open research more broadly. You can reach us at research-datamanagement@glasgow.ac.uk.

Please return to Moodle to complete a short quiz so that your attendance can be recorded. If you've any follow-up questions, please email us at the address above, and watch for our regular drop-in advice sessions, which will be advertised through your College.

Good luck with your research!