Introduction to Research Data Management (RDM)

gro@gold.ac.uk

RDM at Goldsmiths

Research and Knowledge Exchange offers support on funder expectations for research data management res-services@gold.ac.uk

If you need guidance on using GRO to deposit research data, or advice on using an alternative data repository gro@gold.ac.uk

Contact IT&IS for provision of secure storage facilities during data collection and throughout the data lifecycle https://www.gold.ac.uk/it/

Contact the Data Protection Officer for advice on data protection and GDPR dp@gold.ac.uk

What is Research Data Management?

“**Research data** are the evidence that underpins the answer to the research question, and can be used to validate findings regardless of its form (e.g. print, digital, or physical)… The primary purpose of research data is to provide the information necessary to support or validate a research project's observations, findings or outputs” UKRI, *Concordat on Open Research Data (2016)* https://www.ukri.org/wp-content/uploads/2020/10/UKRI-020920-ConcordatonOpenResearchData.pdf

“**Research data management** is simply the effective handling of information that is created in the course of research.” JISC, *How and why you should manage your research data: a guide for researchers* (2021) https://www.jisc.ac.uk/guides/how-and-why-you-should-manage-your-research-data

What is Research Data Management?

Importance of research data will vary depending on the nature of your research

- For some it will require a lot of thinking about what, how and where data is collected, stored and shared
- For others it will involve learning best practice for filing, file naming and data description
- All researchers should think about Research Data Management at the start of a project

If you're new to research data management don't worry - much of what's involved should be familiar

Examples of Research Data

Examples of research data may include:

“Electronic text documents, spreadsheets, laboratory notebooks, field notebooks and diaries, questionnaires, transcripts and codebooks, audiotapes and videotapes, photographs and films, examination results, specimens, samples, artefacts, slides, database schemas, database contents, models, algorithms and scripts, workflows, standard operating procedures and protocols, experimental results, metadata and other data files like e.g. literature review records and email archives” *CESSDA Data Management Expert Guide* https://dmeg.cessda.eu/Data-Management-Expert-Guide/1.-Plan/Research-data

Examples of research data produced by practice researchers may include:

“Notebooks, ecological survey data, anonymised audience surveys, prototype software, work in progress sound recordings, score drafts, storyboards, interview transcripts, or any other datasets that were instrumental in the process of the research inquiry” James Bulley and Ozden Sahin (2021). *Practice Research - Report 2: How can practice research be shared?* https://research.gold.ac.uk/id/eprint/30222/1/2021_Bulley-Sahin_Practice-Research-reports.pdf

Benefits of Research Data Management

- Makes it easier to produce research outputs using data recorded and documented consistently throughout a project
- Protects yourself and others: you can reduce the risk of costly/embarrassing/damaging accidents, such as losing data, or disclosing confidential data
- Makes it easier to select the data that should be preserved or which can be disposed to save space
- Confidence in the robustness and completeness of your data now will allow future projects based on your research to be built on solid foundations
- Transparency of publicly funded research: funder requirements https://goldmine.gold.ac.uk/AdviceInformation/Pages/Funder-Requirements-.aspx

Funder and Publisher Requirements

- Funder requirements – applying for funding (Data Management Plan), preserving your data, sharing your data https://goldmine.gold.ac.uk/AdviceInformation/Pages/Funder-Requirements-.aspx
  - AHRC, British Academy, EPSRC, ESRC, Horizon Europe, Wellcome Trust

- UKRI Open Access policy - a Data Access Statement needs to be included in research articles covered by the policy

- Publisher requirements Imperial list of publishers with known research data policies

Funder Requirements

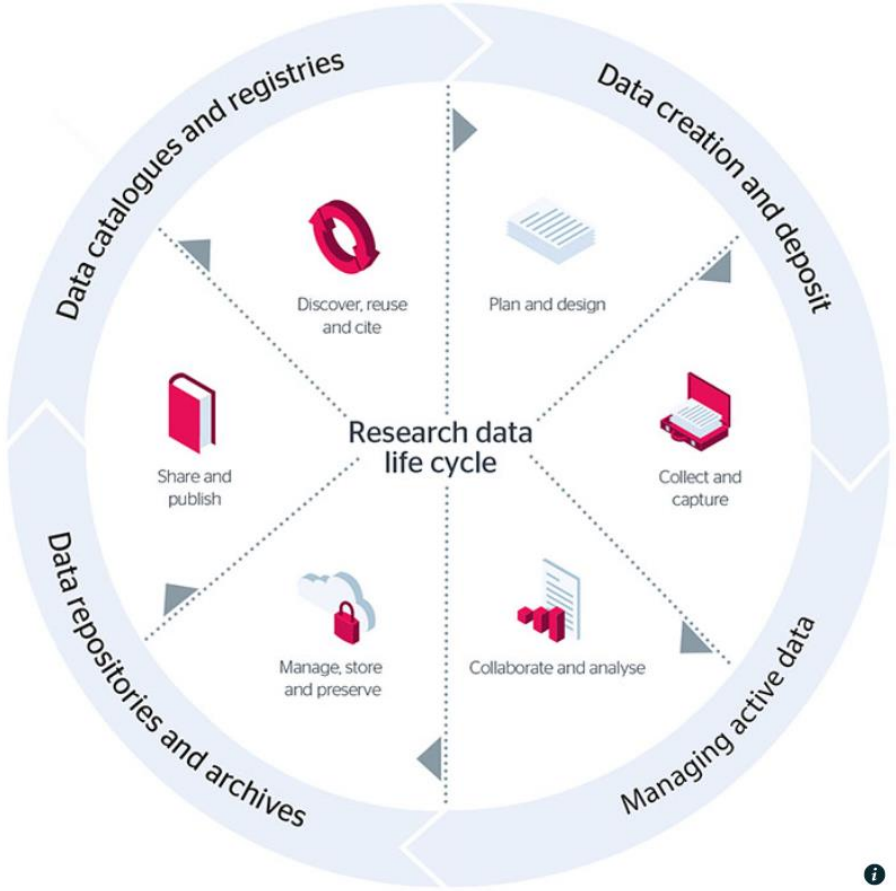
Funders expect researchers to manage research data responsibly and to archive it in a suitable data repository

Funders acknowledge that there may be a valid legal, ethical or commercial reason for restricting access

The following elements are common to many funders' policies:

- Applicants for grant funding are expected to submit a data management plan (DMP) as part of their application
- Data supporting the research outcomes should be preserved and shared on completion of the project and publication of findings by being deposited in a suitable public data repository wherever possible
- Effective data management may require the use of paid-for resources/services, e.g. for storage and archiving. It is appropriate to seek funds for data management activities and costs, and these should be priced into project budgets where possible

Research Data Management in Practice



1. Beginning a project (plan and design; Data Management Plan)

2. Working on a project (organise and describe; manage and store; keep data safe; sensitive data)

3. Finishing a project (select a data repository; publish and share on GRO)

1. Beginning a project

(plan and design; Data Management Plan)

1. Beginning a project: writing a DMP

A Data Management Plan helps you consider how you will organise your data, files, research notes and other documentation throughout the length of your project and to support reuse beyond the life of a project

The aim is to help you find your data easily, keep them safe and have sufficient documentation to be able to re-use them throughout your research and beyond

A DMP should be an evolving document in which information can be added as the project progresses, and when significant changes occur. It is good practice to establish a schedule for reviewing and updating a DMP in combination with project events e.g. funding approval, periodic reviews

Although there is no mandate at Goldsmiths to complete a DMP, it is good research practice to complete one at the start of a project

Research funders increasingly require grant holders to develop DMPs as part of their application, to explain how research data will be managed

1. Beginning a project: writing a DMP

Data Management Plans (DMP) can cover https://goldmine.gold.ac.uk/AdviceInformation/Pages/Writing-a-DMP.aspx:

- **Data Collection** - a description of the type and quantity of data outputs that will be generated by the research
- **Documentation and Metadata** - the procedures that will be used for data collection, processing, documentation, and quality control
- **Ethics and Legal Compliance** - consideration of legal and ethical aspects of data management, including compliance with data protection and research ethics obligations. Information about who will have ownership of the data/how intellectual property in the data will be managed.
- **Storage and Backup** - information about how data will be stored, backed up and shared between members of the team during the project.
- **Selection and Preservation** - information about where data will be archived for long-term preservation and access
- **Data Sharing** - information about when data will be made accessible. Indication of whether access to the data will be subject to any restrictions due to legal, ethical or commercial reasons
- **Responsibilities and Resources** – consider the resources required to deliver your DMP. Examples include specialist training, hardware or software not currently provided by Goldsmiths, charges applied by external data repositories, costs of staff time and resources or services to prepare and upload your data for long term storage

1. Beginning a project: writing a DMP (funder expectations)

Most research funders will either supply a DMP template that applicants must complete or ask a set of questions about the proposed data management plan as part of the funding application form

Applicants are asked to describe any research data they intend to create, how data will be stored during the lifetime of a project and how valuable data will be shared after the project has ended

Applicants are also asked to say how much research data management will cost UKDS costing data management for your research

1. Beginning a project: writing a DMP (funder expectations)

In your application form you need to convince the assessment panel that you:

- Understand the principles of good data management
- Have carefully considered how you will store, share and enable others to access your research data
- Have considered relevant legal and ethical issues
- Have provided costs for the resources needed to implement good data practices e.g. people's time, equipment, infrastructure, training and tools
- Can meet all specific requirements set out by the research funder

The best way to get started writing a DMP is to begin with a suitable template. We recommend using DMPOnline an online tool which provides templates for all UK research funders who require a plan https://dmponline.dcc.ac.uk/

1. Beginning a project: writing a DMP (funder expectations)

Funder DMP templates on DMPOnline

- AHRC https://dmponline.dcc.ac.uk/public_templates?page=1&search=ahrc
- EPSRC https://dmponline.dcc.ac.uk/public_templates?page=1&search=epsrc
- ESRC https://dmponline.dcc.ac.uk/public_templates?page=1&search=esrc
- Horizon Europe https://dmponline.dcc.ac.uk/public_templates?page=1&search=horizon+Europe
- Wellcome Trust https://dmponline.dcc.ac.uk/public_templates?page=1&search=wellcome

1. Writing a DMP (Summary)

Detailed guidance on what to include in a Data Management Plan is available on Goldmine: https://goldmine.gold.ac.uk/AdviceInformation/Pages/Research-Data-Management-.aspx

We recommend using DMPOnline an online tool from the Digital Curation Centre (DCC)

You can access examples of DMP's that have been completed for other research projects via the library of public DMP's created using DMPOnline

The UK Data Service has a data management costing tool and checklist UKDS costing data management for your research

2. Working on a project

(organise and describe; manage and store; keep data safe; sensitive data)

2. Working: Organising/describing data

Use a logical, hierarchical **folder structure** to store your files UK Data Service Filing System

Intelligent use of **file naming** enables you and others to easily identify the contents of a file, and can be used to organize and version-control files UK Data Service File Names

The **file formats** you use for your data may affect what you can do with the data and how effectively they can be preserved and shared UK Data Service Recommended Formats

Effective **documentation and metadata** are crucial in making data findable and reusable in the future, by both the creators and others UK Data Service: Document your data

2. Working: Storing live data

“As a general principle data should be stored on secure Goldsmiths computers/servers”.
https://goldmine.gold.ac.uk/AdviceInformation/Pages/Writing-a-DMP.aspx

IT&IS information on file storage for Goldsmiths research
https://goldmine.gold.ac.uk/NewsEvents/Pages/ITIS-File-storage.aspx

2. Working: Storing live data

Keeping copies of research data on personal laptops, external hard drives and USB sticks should not be relied upon as a primary method of storage. This is because:

- data stored on USBs or hard drives can easily become corrupted;
- personal laptops or hard drives are at risk of being lost or stolen;
- portable storage devices can break or become faulty with use.

Only use portable storage as a secondary storage option in conjunction with Goldsmiths storage provision, and securely encrypt any data that is being kept on a personal device.

2. Working: Storing live data

Commercial cloud storage services (Google Drive, Dropbox etc.) can seem convenient but there are several factors to consider before using them.

- No guarantee that a commercial service will not be withdrawn or terminated.
- It may not be apparent where your data is being stored (UK, EU, beyond?)
- Backups may occur infrequently; different companies have different policies on how often they back up data.
- It is not always clear who can view and access the data.

Use Goldsmiths maintained services rather than commercial cloud providers for the storage of research data. Commercial cloud storage should not be used to store sensitive or confidential data.

2. Working: Keeping data safe

You should always ensure that you back up copies of important data in multiple locations so that you can restore it quickly in case of data loss UK Data Service guidance on backup

Save three copies of your data (original copy and two backup copies) with at least one in a different geographical location

Version your files UK Data Service guidance on versioning

Encrypt your data UK Data Service guidance on data encryption

2. Working: Sensitive data

Sensitive data can refer to:

- any data that could be used to identify an individual
- confidential data, including commercially sensitive data
- data that, if released, is likely to cause harm to any individual or community or will have significant negative public impact

Guidance on ethics and legal compliance is available on Goldmine

UK Data Service have advice on data protection

3. Finishing a project

(select a data repository; publish and share on GRO)

3. Finishing a project: Using a data repository

At the end of a research project there are three options for your data:

- Sharing in a data repository such as GRO
- Retaining it in on a secure server UK Data Service Guidance on Secure Storage
- Destroying the data in a secure manner UK Data Service Guidance on Disposal

3. Finishing a project: Using a data repository

A data repository is a service that exists to preserve and provide access to research data

The best way to ensure the long-term preservation, access and reuse of data is to deposit with a trusted data repository

The advantages of depositing your data with a repository include:

- Relieves the researcher of the long-term management of the data
- A permanent public record will be created to enhance discoverability
- Data will be assigned a persistent identifier (e.g. DOI) making it easier for you and others to cite your data in publications
- Compliance with funder and publisher data policies

3. Finishing a project: Using a data repository

Datasets can be deposited in our institutional repository, Goldsmiths Research Online - GRO (http://research.gold.ac.uk/)

Discipline-specific repositories also exist, for example the UK Data Service (economic, social and population data)

Examples of general-purpose data sharing services, include figshare and Zenodo

Your funder may recommend a particular repository

- ESRC recommends the UK Data Service
- Wellcome maintains a list of approved data repositories

Some publishers provide lists of recommended data repositories

3. Finishing a project: Using a data repository

Before you can go about preserving and sharing your data, you will need to identify what needs to be preserved

You are unlikely to need to preserve all the data you collect or create in the course of your research

It is important to consider the potential re-use value of your data, as well as any legal or policy issues. You will need to select data of value, and dispose of the remainder

Digital Curation Centre: How to Appraise and Select Research Data for Curation

3. Finishing a project: Using a data repository

Not all data can be made publicly available. Appropriate safeguards need to be in place before data that contain sensitive or confidential information can be shared

If your data is too sensitive to share publicly, you could deposit a summarised or anonymised version

You can deposit datasets under an embargo. This will ensure that you can deposit the data at convenient point in your research process and have it automatically made public at a future date. Many funders will accept short embargoes but it is important to check your funders' policy

3. Finishing a project: Using a data repository

Licensing is recommended when publishing research data. Allows you to share your data while stating clearly what others can do with the data.

Creative Commons provides a way to licence the use of material you create and share. Creators retain copyright while allowing others to copy, distribute, and make some uses of their work.

Six main licences https://creativecommons.org/licenses/. Creators can select a license which suits their needs and authorise the appropriate use of their work. The more letters in the licence, the more restrictive it is.

Check the policy of your funder. Horizon Europe require data to be licenced CC BY or Creative Commons Public Domain Dedication (CC0).

3. Finishing a project: Using GRO

- All researchers can deposit their research outputs and datasets to Goldsmiths Research Online - GRO (http://research.gold.ac.uk/)
- Provides long-term hosting and secure digital preservation for research data
- Aligns with best practice and standards, ensuring information is preserved, accessible and usable for future research
- Full text material in GRO is licensed with a Creative Commons license
- A Digital Object Identifier (DOI) will be created to allow citations

3. Finishing a project: Using GRO

Examples of datasets on GRO:

McGrath, Sean. 2022. *Evaluating Colour in Concept Diagrams Dataset for Diagrams 2022* https://research.gold.ac.uk/id/eprint/31899/

Barber, Ros. 2020. *BDNE Zeta Dataset* https://research.gold.ac.uk/id/eprint/28390/

Nolas, Sevasti-Melissa; Varvantakis, Christos; Aspa, Chalkidou and Apgar, Marina. 2020. *to archeio/the archive project: Greek Crisis Literature database* https://research.gold.ac.uk/id/eprint/28062/

Orgs, Guido. 2017. *Joint Action Aesthetics Dataset* https://research.gold.ac.uk/id/eprint/20364/

Miller, Rebecca. 2016. *Bodymap Workshop 1* https://research.gold.ac.uk/id/eprint/20788/

Phillips, Andrea and Cruz, Carla. 2014. *Tagore, Pedagogy and Contemporary Visual Cultures* https://research.gold.ac.uk/id/eprint/20909/ , https://research.gold.ac.uk/id/eprint/20910/

McVeigh, Simon. 2014. *Calendar of London Concerts 1750-1800* https://research.gold.ac.uk/id/eprint/10342/

3. Finishing a project: Data access statements

Data access statements, are included in publications to describe where the data associated with the paper is available, and under what conditions the data can be accessed

UKRI Open Access policy - a Data Access Statement needs to be included in research articles covered by the policy

Your journal guidance for authors should indicate the format and placement of a data access statement

Your statement should typically include:

- where the data can be accessed (preferably a data repository)
- a persistent identifier, such as a Digital Object Identifier (DOI) or a link to a permanent record for the dataset
- details of any restrictions on accessing the data and a justifiable explanation (e.g. for ethical, legal or commercial reasons)

Guidance on writing data access statements University of Manchester how to write a data access statement

Some useful resources

Research Data Management Guidance on Goldmine https://goldmine.gold.ac.uk/AdviceInformation/Pages/Research-Data-Management-.aspx

DMPonline https://dmponline.dcc.ac.uk/

UK Data Service: Research Data Management Learning Hub https://ukdataservice.ac.uk/learning-hub/research-data-management/

JISC: Research Data Management Toolkit https://www.jisc.ac.uk/guides/rdm-toolkit

CESSDA Data Management Expert Guide https://www.cessda.eu/Training/Training-Resources/Library/Data-Management-Expert-Guide

Digital Curation Centre https://www.dcc.ac.uk/guidance

Key messages

Research Data Management makes life easier: well-organised data management increases your efficiency, and saves time and effort in the long run

Importance of research data will vary depending on the nature of your research but all researchers should think about Research Data Management at the start of a project

A Data Management Plan helps you consider how you will organise your data throughout the length of your project and to support reuse beyond the life of a project

Remember to

- Save work in a format that can be re-opened easily
- Name and organise files so they can be found quickly
- Back up valuable data
- Control who has access to data
- Share data via a repository such as GRO to help maximise research impact

Further help

Research and Knowledge Exchange offers support on funder expectations for research data management res-services@gold.ac.uk

If you need guidance on using GRO to deposit research data, or advice on using an alternative data repository contact gro@gold.ac.uk

Contact IT&IS for provision of secure storage facilities during data collection and throughout the data lifecycle https://www.gold.ac.uk/it/

Contact the Data Protection Officer for advice on data protection and GDPR dp@gold.ac.uk