Research Data Management

CoSS Ethics Information and Discussion Session, 2019-12-16

Matt Mahon, Research Data Management Service
research-datamanagement@glasgow.ac.uk

Research data is any information (digital or physical) required to underpin research

All images free to share, from Google and Flickr
CREATING YOUR DATA

Key Principles

UK Research and Innovation

‘Publicly funded research data are a public good, produced in the public interest, which should be made openly available with as few restrictions as possible in a timely and responsible manner that does not harm intellectual property.’

Create Organise Keep Find & Share

Data management plan required as part of application

- policy is mandatory for funded research staff
- policy is unclear about situation for students
- work with their funded researchers to support compliance
- have the potential to resort to financial sanctions

Data should normally be publicly available at the point of publication. Include a data citation in your publication

Data must be deposited in an appropriate repository (normally the UK Data Archive) within 3 months of the end of the award. The data centre will ensure secure preservation of data.

Describe your data using appropriate metadata to enable other researchers to find, cite and understand it

Funding is available for preparing data for deposit.

The Code of Good Practice in Research

Section 3.1 Compliance with policies:
- staff and students should be familiar with University Policies
- university expects researchers to observe policies of funders and professional bodies

Section 3.6 Documenting Results, Storing Data and its Future Use:
- keep clear and accurate records of research
- prepare a data management plan for projects which will generate data
- contextualise research data by keeping sufficient metadata
- store data of long-term value for at least 10 years post project
- deposit a copy of data generated using University resources when member of staff / student leaves.
- deposit data underpinning publications in a trusted repository
- when possible, make deposited data openly available

Section 3.7 Publishing Research
- include a data citation in published work

http://www.gla.ac.uk/media/media_227599_en.pdf
PGR Data Management Plans

All PGR students entering doctoral training programs in academic year 2018-19 onwards must prepare a data management plan for their research projects as part of their submission for progression from year 1 to year 2 of their doctoral studies.

Students should use the PGR template regardless of their funder!

UofG Data Management Plan template: [http://eprints.gla.ac.uk/179057/](http://eprints.gla.ac.uk/179057/)
UofG Data Management Plan guidance: [http://eprints.gla.ac.uk/179058/](http://eprints.gla.ac.uk/179058/)

Templates and guidance are also available in DMPOnline!

If the student will not be generating data, using secondary data, developing software or undertaking qualitative or quantitative analysis (e.g. pure mathematics) then they can make a statement to this effect in the first part of section 2 and leave the remainder blank.

DATA MANAGEMENT PLANS (DMPs)

Data management plans are now a common feature of grant applications. Funders are placing more emphasis on DMPs and we’ve recently seen some applications returned for revision due to weak plans.

What is included in a DMP?

- What is the data that will be collected?
- How will it be documented?
- Ethics and Intellectual Property
- Storage and organisation of the data
- Deposit and long-term preservation
- Plans for data sharing*
- Responsibility for implementation

*Sharing ≠ open access
Where does a DMP fit with other processes?

1. Use the DMP to plan **what you want to do** with the data (while satisfying University and funder requirements).

2. Complete the appropriate Data Protection paperwork you establish **what you are legally allowed to do** with the data.

3. Seek **ethical approval** for all your planned uses of the data.

Full workflow diagram, with links to relevant support and documents:

10.36399/gla.pubs.202746
**Ethics, Consent and GDPR**

Any research involving human participants, material or data requires approval from an appropriate Ethics Committee (College, University or NHS).

Consent from participants is required for ethical purposes, and GDPR requirements should be dealt with separately.

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

The UoG online [Introduction to GDPR training course](#) is a requirement for all University of Glasgow staff and PGR students.

**DP/FOI Office**

*Tel*: 0141 330 6494  
*Email*: dp@gla.ac.uk  
*Web*: [www.gla.ac.uk/services/dpfoioffice](http://www.gla.ac.uk/services/dpfoioffice)

**Employee and Organisational Development**

[www.gla.ac.uk/services/humanresources/employeeandorganisationaldevelopment/learningcoursesandresources/dataprotectionfreedomofinformation/](http://www.gla.ac.uk/services/humanresources/employeeandorganisationaldevelopment/learningcoursesandresources/dataprotectionfreedomofinformation/)

---

**Support for planning and ethics**

**Grant application advice and support**

Review service for data management plans

Advice on costing RDM activities (eg storage: £180/TB/y but falling)

Contact research-datamanagement@glasgow.ac.uk

**Ethics advice and support**

Contact your College or School ethics officer in the first instance

[http://www.gla.ac.uk/services/rsio/researchstrategypolicies/ourpolicies/committeestructure/](http://www.gla.ac.uk/services/rsio/researchstrategypolicies/ourpolicies/committeestructure/)
ORGANISING YOUR DATA

File names allow you to identify a precise experiment or dataset from the name. Choose a format for naming your files and use it consistently.

Good file names are relatively short, but rich in information. Good file names consist of information elements separated by underscores ‘_’

Examples of elements:
- Project or experiment name of acronym
- Location / spatial coordinates
- Researcher name / initials
- Date or date range of data acquisition (use YYYYMMDD)
- Type of data
- Sensitivity of data
- Version number of file
- File extension

Top tips:
- Avoid special characters
- Do not use spaces
- Use an underscore as an element delimiter
- Use a hyphen or capital to delimit words within an element
- Think about the order of elements
- Names should be family name first followed by initials.
- Abbreviate where possible

Commonly seen filenames
- Plan_FinalFINAL.doc
- Test_data_2013
- Second_test
- Meeting Notes Oct 23

Filenames with a naming convention:
- 20130503_DOEProject_DesignDocument_Smith_v2-01.docx
- 20130825_DOEProject_Ex1Test1_Data_Gonzalez_v3-03.xlsx
- 20131002_DOEProject_Ex1Test1_Documentation_Gonzalez_v3-03.xlsx
- 20141023_DOEProject_ProjectMeetingNotes_Kramer_v1-00.docx

Explain your file-naming convention in your ReadMe file!
File naming standard for oral history interviews:
interviewer_interviewee_[status]_[part #]_[date].format

Examples:
boyd_johnson_pres_01_20120801.wav
macdowell_benberry_mez_02_20120801.h264

[status]
Raw = capture format
Pres = preservation copy
Mez = mezzanine copy (working copy)
Ed01 = edited version #1
Dis = dissemination copy, e.g. for web

Readme file
1. What is data?
2. How was data acquired?
3. What software is needed to open?
4. What is the file-naming convention?
5. Anything else useful? eg sample prep

Readme file
1. Where has data come from?
2. How has data been processed?
3. What software is needed to open?
4. What is the file-naming convention?
5. Anything else useful?
KEEPING YOUR DATA SAFE

Restricting access to your data

Use your University’s authenticated storage spaces – not Dropbox, Google Drive, iCloud etc!

Security measures should be proportionate!

- Consider your ethical approval
- Comply with GDPR requirements: don’t share data without adequate safeguarding!
- Who will be able to access the materials?
- Encrypt your files/devices if necessary
- Think about what will happen to your data in transit – is it secure?

General Data Protection Regulation Article 45: Personal data should not be shared outside the EU without explicit consent and adequate safeguarding.

Where should the data reside?

<table>
<thead>
<tr>
<th></th>
<th>OwnCloud</th>
<th>T expiration</th>
<th>T</th>
<th>T</th>
</tr>
</thead>
<tbody>
<tr>
<td>Data must stay in the EU</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✗</td>
</tr>
<tr>
<td>Data must stay in the UK</td>
<td>✓</td>
<td>✓</td>
<td>✗</td>
<td>✗</td>
</tr>
<tr>
<td>Data must stay on campus</td>
<td>✓</td>
<td>✓</td>
<td>✗</td>
<td>✗</td>
</tr>
<tr>
<td>Storage managed or contracted by the University</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✗</td>
</tr>
</tbody>
</table>

General Data Protection Regulation Article 45: Personal data should not be shared outside the EU without explicit consent and adequate safeguarding.

Who can access the data?

<table>
<thead>
<tr>
<th></th>
<th>OwnCloud</th>
<th>T expiration</th>
<th>T</th>
<th>T</th>
</tr>
</thead>
<tbody>
<tr>
<td>Data is sensitive and needs to be restricted</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✗</td>
</tr>
<tr>
<td>Share with colleagues at the University of Glasgow</td>
<td>✓</td>
<td>✓</td>
<td>✗</td>
<td>✗</td>
</tr>
<tr>
<td>Share with collaborators at another university</td>
<td>✓</td>
<td>✓</td>
<td>✗</td>
<td>✗</td>
</tr>
<tr>
<td>Share with collaborators without university logins</td>
<td>✓</td>
<td>✓</td>
<td>✗</td>
<td>✗</td>
</tr>
</tbody>
</table>

OwnCloud: [https://owncloud.gla.ac.uk](https://owncloud.gla.ac.uk)
MS OneDrive for Business: [https://office365.gla.ac.uk](https://office365.gla.ac.uk)
Shared network drive: (usually J:\)
IT Services Guide: [https://www.gla.ac.uk/media/media_477731_en.pdf](https://www.gla.ac.uk/media/media_477731_en.pdf)
Information security guidance

Best practice for handling confidential data

- Use central filestores on secure servers maintained in secure physical environments
- Confidential data should not be held on local disk storage (eg 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
- Stored on portable media (eg memory stick)
- Exchanged with external organisations and individuals

https://www.gla.ac.uk/myglasgow/it/informationsecurity

SHARING YOUR DATA AND FINDING OTHER DATA
Why deposit data?

1. Demonstrate the integrity of your research by depositing all data of long term value
2. Share some of that data with other researchers

Deposit your data in a repository:

- Somewhere that stores and manages files
- Usually accepts a definitive version of files, and
- Usually ‘open’
- Where other people can find and potentially access and reuse our files

Persistent URLs: Digital Object Identifiers

A DOI is a unique alphanumeric string assigned by the International DOI Foundation to identify content and provide a persistent link to its location on the Internet. All DOI numbers begin with a 10 and contain a prefix and a suffix separated by a slash:

A dataset in Enlighten: 10.5525/gla.researchdata.191

A paper in the Journal of the National Cancer Institute: 10.1093/jnci/djv204

You need a DOI to cite your data properly. Always use DOIs in your citations where available.
Choosing a licence for your data

What does your funder/the University/project partner require?

What does your chosen repository accept?

Open access licences:
Creative Commons licence suite
Open Data/Public Domain

The best approach is to divide your data into groups:
• Data that can be shared under an open access licence, and
• Data that should be restricted or closed.

Describe your data licensing plans in your Data Management Plan.

Choosing a repository

There are a number of generic and subject-specific repositories you could use:
• Generic: Figshare
• Funder-funded: NERC, ESRC (UK Data Service)
• Subject specific: Dryad for Ecology, ADS for Archaeology
• Institutional: Enlighten: Research Data
• For code: Github, SourceForge, BitBucket
• Can use re3data to find suitable repository
Further training options for staff and students

**For Staff – book on CoreHR**
- RDCR  Managing your research data
- RIDMP  Understanding Data Management Plans

**For PGR Students – book on MyCampus**
- RSDC 6024  Research Data Management for CoSS
- RSDC 6033  How to organise, store and share your research findings for Arts and Humanities
- RSDC 6025  Research Data Management for MVLS and CoSE
- RSDC 6030  How to write you Data Management Plan

**Bespoke Training**
We can provide tailored training on any aspect of Research Data Management, for any staff or student group.

https://www.gla.ac.uk/myglasgow/datamanagement/training/
What do you need to do?

- Write a Data Management Plan
- Implement a good file naming procedure
- Create a logical folder structure
- Keep good documentation (metadata)
- Curate datasets which underpin publications
- Deposit datasets in repositories
- Use data citations (DOIs where available)

Encourage your students / post-docs / colleagues to do the same!

Most importantly: Contact us!

research-datamanagement@glasgow.ac.uk

Assistance and advice for all stages of the data lifecycle

Mary Donaldson
Mick Eadie
Matt Mahon