University of Glasgow

# Guidance notes - Data Management Plan template for PGR students

A data management plan or DMP is a document that outlines how data will be handled both during a research project, and after the project is completed. The goal of a data management plan is to consider many aspects of data management before the project begins, ensuring that data are well-managed in the present and prepared for preservation in the future.

If a student will not be generating data, using secondary data, developing software or undertaking qualitative or quantitative analysis (e.g. Pure Mathematics) then they need to make a statement to this effect in the first part of section 2 and leave the remainder blank. Please note that such a student may still find it useful to know and record their funder and award number, as asked for in section 1.

Please note that if you receive external funding to support your research, you should find out if your funder has any requirements for RDM or data sharing. Some funders have strict requirements for data sharing for funded work and others (notably ESRC and NERC) provide datacentres for long-term archiving of their funded work.

If you need further help when completing your data management plan, there is support available:

- You can attend a workshop on data management planning https://www.gla.ac.uk/myglasgow/datamanagement/training/
- You can look at our webpages https://www.gla.ac.uk/myglasgow/datamanagement/
- You can look at example data management plans for inspiration http://www.dcc.ac.uk/resources/data-management-plans/guidance-examples
- You can contact the Research Data Management Service research-datamanagement@glasgow.ac.uk

| 1. Overview | |
|---|---|
| Researcher name | Fill in the name of the PGR student conducting the research |
| Supervisor name | Fill in the name or names of the student's supervisor(s) |
| Project title | Fill in the title of the project |
| Funder & award number | If the project receives any funding (internal or external), give the name of the funder and the award number. If the award number is not known, please check with your local Grad School administrator to find it out as this number will also be required when you publish papers based on your research and submit your PhD thesis. |
| Project Summary | Give a brief overview of your research |

| 2. Data |
|---|
| What types of data will be collected or created? |
| List the broad categories of data that will be collected or created during the project.<br>Examples might include:<br>• Semi-structured interview recordings |

- Transcripts of focus group sessions
- Relational database of testaments found in archival collections
- Experimental measurements
- Metabolomic data
- Gel images
- Photographs / micrographs
- Design data for sensors and test setup
- Custom written code

What formats will you use?

Identify the file formats that will be used for each category of data in the section above. If the data will be collected in one format, but stored in another, record both.

Examples might include:

- interview recordings will be in MPEG-4 format
- transcripts will be in .rtf format
- data in spreadsheets will be in .csv format
- image data will be captured in the manufacturer's format but will be stored as .tif files
- design data will be in .igs and .gds format and also stored as .pdf
- code will be written in python

How much data will you collect?

For each of the categories of data identified in the first part of section 2, try to estimate the volume of data to be stored (in Mb / Tb). It can be helpful to identify the size of a single file of the appropriate type and then multiply up based on the number of each file type you anticipate collecting / creating. Please note that estimations of dataset size become more important with increasing size, to ensure sufficient storage is available.

## 3. Documentation

How will the data be documented and described?

To ensure future understanding of the research data and to maximise the potential for re-use of data, contextual information or documentation (metadata) should accompany research data. This provides a secondary user (or the original creator at a point in the future) with any necessary details on the origin or manipulation of the data in order to prevent any misuse, misinterpretation or confusion. Recognised standards for metadata (documentation) should be used where possible.

In this section, list any documentation you will record to contextualise the data files.

Are there any standards for this in your field of research?

Some areas of research have established standards for documentation and metadata. If these exist in your field of research, please list the standards to which you will work.

Examples might include:

- MIAME standards for microarray data
- CIF standards in crystallography
- DDI standards for social, behavioural and economic data
- FITS standards for astronomical data
- PDBx exchange dictionary for 3D protein structure data
- SDAC standard for documentation of astronomical catalogues
- SDMX standards for statistical data and metadata exchange

## 4. Ethics and Intellectual Property

### Who owns the data in your project?

The answer to this question depends on your individual circumstances and whether you have assigned any of your rights to the University, a funder or another party.

The University's policy is that PGR students who are not employed by the University own their IP unless this is governed by a third-party agreement (e.g. funding or sponsorship) or other factors which confer an interest in the IP. Students may also choose to assign their IPR to the University. Any student who assigns any IPR created during their research degree to the University would then be granted the same rights as any employee inventor under the University's Intellectual Property & Commercialisation Policy.

It is advisable to check the terms of your studentship to determine if your IP is governed by a third-party agreement.

https://www.gla.ac.uk/research/ourresearchenvironment/prs/intellectualproperty/ipfaqs/

### Detail any ethical, legal or commercial considerations relating to your research data

List any ethical, legal or commercial issues that affect your research data

Examples might include:

- I anticipate patenting outputs from my research, so the data must be kept confidential
- Third-party datasets used in my research are subject to copyright or licence restrictions
- My project will collect data relating to human participants
- Participants will be potentially identifiable in interview recordings

### How will these concerns be dealt with?

For each of the legal, ethical or commercial issues identified, detail how you will proceed, to minimise the impact on your research and the dissemination of your research. Please note that these issues may only apply to a subset of your total data, in which case it should be made clear exactly which data subsets are affected.

Examples might include:

- Permission will be sought from secondary sources to share the findings of the research.
- Ethical approval for this research has been sought / gained from XXX ethics committee and data will be processed in line with legal requirements (eg data protection / GDPR).
- Sharing of data and other outputs from this research will be postponed, allowing patents to be acquired. Once patents have been awarded, datasets from this project will be shared.
- Transcripts of original recordings will be used for archiving and sharing purposes to reduce the possibility of identification of participants

## 5. Storage and organisation

### How will the data be named, organised and structured?

For each category of data, the researcher should outline how they will organise their folders and files, the file-naming convention that they will use and how they will keep track of versions. The actual file-naming conventions need not be detailed, but it should be indicated that detailed information will be included with project documentation.

| How will the data be stored for the duration of the project? |
| --- |
| Please indicate which University-provided storage options will be used for your research data. The University of Glasgow provides three main storage options which are suitable for research data:<br>• J: drive space – access to this can be arranged through your supervisor.<br>• OneDrive for Business https://www.gla.ac.uk/myglasgow/it/office365/<br>• OwnCloud https://www.gla.ac.uk/myglasgow/it/filesharing/owncloud/ - access to this storage is also accessed via your supervisor.<br><br>Please note that external storage providers who are not under contract to the University of Glasgow (such as DropBox, Google Drive, Box etc) should not be used for the storage of research data.<br><br>If you need advice about storing research data during your project, you can contact your local IT support (link below) for advice.<br><br>If your school or research group has opted out from central IT provision, advice on data storage should be sought from local IT support (link below).<br><br>https://www.gla.ac.uk/myglasgow/it/localitsupport/ |
| How will the data be backed up during the project? |
| State how often the data will be backed up and to which locations. How many copies are being made? Storing data on laptops, computer hard drives or external storage devices alone is very risky. The use of robust, managed storage provided by university IT is preferred. Similarly, it is normally better to use automatic backup services provided by IT Services than rely on manual processes.<br>https://www.gla.ac.uk/myglasgow/it/informationsecurity/backups/ |
| Does access to the data need to be controlled for the duration of the project? |
| List any datasets which you collect or create which contain confidential or personal information. For each of these datasets, state how the data will be stored and how security or confidentiality will be maintained.<br>This might include encryption of mobile devices and storage media and password protecting files or folders, but might also include physical measures such as locking your screen when you're away from your computer and locking your office door when you leave the office.<br>This could also include details of how the folder structure and file naming convention will identify and segregate data that needs to be secure.<br>Lastly, consider how you might transfer files containing confidential information – will you send encrypted files via email, use the University file transfer service (https://transfer.gla.ac.uk/) or transfer files using shared network drives or OneDrive?<br><br>https://www.gla.ac.uk/myglasgow/it/informationsecurity/confidentialdata/ |
| Who has the right to access the data during the project? |
| List the people (either by name or by position held) who should have access to your research data during your project. It may be that different people need access to different subsets of your research data. Bear in mind that if you anticipate archiving your research data (as per University requirements), members of the Research Information Management service will need to be able to handle the data. |

| 6. Deposit and long-term preservation |
| --- |
| Which data should be retained long-term? |

Not all research data has long-term value (defined by the University as underpinning a research publication, thesis or funding application). Indicate which subsets of the data you will collect or create will have long-term value and will be archived for long-term preservation.

How long will data be retained for?

Indicate how long your archival dataset(s) will be retained for. This could be based on any obligations to retain certain data, the potential reuse value, what is economically viable to keep, and any additional effort required to prepare the data for data sharing and preservation.
The University of Glasgow requires that data of long-term value (defined above) should be retained for a minimum of 10 years from the end of the project.

Please keep in mind that regardless of ownership of the IPRs, the University requires access to all data generated using University resources and will retain the original data when a student leaves the University. This is because the University has obligations placed on it by funders and the government to make publicly funded research data accessible for reuse and increasing its potential impact is a key objective for most UK research councils and funding bodies.

Where will the data be archived at the end of the project?

Indicate where you will deposit your archival data at the end of your project.

Options might include:

- a funder repository as required by your funder (e.g. the UK Data Archive)
- a subject repository (e.g. the Cambridge Crystallographic Data Centre)
- Enlighten: Research Data, the University of Glasgow's institutional data repository

A searchable directory of research data repositories is available:
https://www.re3data.org/browse/by-subject/

If you do not propose to use an established repository, the data management plan should demonstrate that resources and systems will be in place to enable the data to be curated effectively beyond the lifetime of the grant.

What formats will the data be archived in?

Which file formats will be used for your archival dataset(s)? These might be the same as the working formats identified in Section 2, but it is often advisable to archive data in alternative formats which will still be readable many years in the future.

The Library of Congress maintain the most comprehensive list of recommended archival file formats for different data types: https://www.loc.gov/preservation/resources/rfs/TOC.html

## 7. Data sharing

Is any of the data suitable for sharing?

Indicate which of your datasets will be made available for future reuse (either as publicly available downloadable data or on a restricted basis).

The University is committed to ensuring that data derived from publicly funded research is made available to other organisations and individuals. Once results have been published, the University expects researchers to deposit the data available in a trusted repository (for example, 'Enlighten: Research Data' or a funder or subject repository) for long-term safekeeping and, subject to any restrictions by the funder or due to legal, ethical or commercial sensitivity, make the data openly available.

Most funders expect that you will take steps to ensure that your data can be shared, including seeking consent for sharing (anonymised) data from participants and project partners. Data which is truly unsuitable for sharing is expected to be in the minority if sufficient attention is paid to this at the planning stages.

How will the data be shared?

Please indicate which mechanism will be used for sharing archival data.
Examples might include:

- Data will be shared via the repository deposit (this is the preferred option)
- Data will be shared via personal communication (this option should be a last resort as contact details and availability cannot be guaranteed over the necessary period)
- Data will be published in a data journal or as supplementary information to an article

Who should be able to access and use the shared data?

If any of the datasets identified in the first part of this section can be shared, indicate here who should have access.
Examples might include:

- Suitable for open public access
- Bonafide researchers only
- Signatories of end-user agreements only

## 8. Implementation

Who is responsible for implementing this plan?

Indicate here: who will be responsible for each data management activity?

How will this plan be kept up-to-date?

The researcher should provide a statement about how regularly the plan will be reviewed, and with whom. If the project is part of a wider programme of research, then consideration should be made about how the PGR DMP aligns with a wider programme DMP.

What actions are necessary to implement this plan?

What do you need to do to make this plan work?
Examples might include:

- Contact local IT support to ensure storage provision is adequate
- Contact preferred data repository to check ingest criteria and required formats
- Ensure consent for data archiving and sharing are included in ethics documentation

What training or further information are needed to implement this plan?

The researcher should identify any further training or support that they might need to implement the DMP. For example, they might need to attend one of our workshops on managing research data, contact the Data Protection team for advice on GDPR or attend a UK Data Service webinar on using ReShare.