Open Access Policy for MDC and CEDAR

Proper research data management is integral to good research practice. It ensures that the data generated by MDC and CEDAR members are stored securely, will be reusable in the future, and can be shared easily amongst collaborators. Moreover, it is an increasingly important part of funder and/or institutional requirements regarding open access (cf. UoE policy PGRs).

Below you will find guidelines detailing how to manage data, and assigning roles and responsibilities. They are largely based on the UK Data Archive's best practices, but tailored as much as possible to the needs of MDC and CEDAR.

1. Brief background to Open Access and Data Management:

Research councils and other public funders of research wish to make research outputs publically available to ensure that research is open to scrutiny, to enhance scientific discovery, and because research output is a public 'product' (i.e., a product that is paid for by tax payers). For example, by making data publically available researchers are able to build on each others' work, conduct secondary analysis, and not repeat null effects. The University has created Open Research Exeter (ORE) to provide a publically available repository of publications and data unpinning publications. Data which is stored elsewhere (e.g. data which is not available publically) can also be registered in ORE. For the purpose of this policy PGR theses are considered publications. This policy affects all new research studies/publications conducted by staff as of **immediate effect**. This policy will become active from **October 2013 for PGR students**.

2. Role and responsibilities:

2.1. PI-led research projects

The PI of the research project will be responsible for:
-   Access to live data and completed research data for external parties, and researchers in other groups within University in accordance with funder and University policy and respecting commercial and confidentiality issues.
-   Setting up new projects. This involves:
    o   Discussing the data management procedures with new staff and students at the beginning of the project.
    o   Setting up file structure and naming conventions.
-   Deletion of any raw data in accordance with funder and university policy.
-   Closing live projects. This involves:
    o   Reviewing folder contents.
    o   Transferring the data to the University ORE or other funder-approved repository, or storing the data and ensuring that an appropriate meta-data record is set with the University ORE

The experimenters (whether they be post-docs or RAs) will be responsible for:
-   Storing the data in the right format (as per section 3).
-   Create data documentation files and update these when required.
-   Assisting in closing live projects. This involves:

o Reviewing the folder contents, ensure no duplication, and ensure that the documentation and software are such that the data can be used in the future.

2.2. PGR research projects

- The lead PGR Supervisor (PI) and the PGR student are jointly responsible for the tasks discussed above in accordance with funder and University policy.
- The lead PGR Supervisor (PI) and the PGR student should discuss and review research data management annually.
- The lead PGR Supervisor is also responsible for advising the PGR student on good practice in research data management.

2.3. Research papers

- Lead authors are responsible for complying with funder and University policy on Open Access research papers.

3. **Data Management Plan:**

- Costs of Data Management can be included in grant application (especially if there is a charge for data archiving)
- All studies require a Data Management Plan (MDC/CEDAR can develop a shared library of data management plans; help and guidance can be provided by rdm@exeter.ac.uk)
- The Data Management Plan should include information on which data will be stored, how data will be stored (including how to clearly label electronic files etc), and where the data will be stored (during the project and post-project). If the researchers intend not to make some of their data publically available, then a rationale needs to be included in the Data Management Plan. If appropriate, a plan for data anonymisation also needs to be included.
- The Data Management Plan should be documented and monitored by each Principal Investigator/ PGR research supervisor
- Research ethics will need to be informed of the Data Management Plan
- Participant information sheets will need to contain information about data storage and access
- Participant consent forms will need to include a statement on data storage and access

4. **Archiving data at the end of a study:**

Currently, there are 2 options for archiving data in ORE:

1. Open Access data (available publically over the internet)
2. Metadata-only record (information about the data is available publically, but the data itself is under embargo or stored elsewhere). The data is only available to ORE administrators.

- You should plan to make data available publically at the end of a study according to funder and University policy whilst respecting legal, ethical and commercial considerations.
- Personal data should be treated according to the Data Protection Act and ethical approval.

- If feasible, confidential and sensitive data should be anonymised to allow for the sharing of the anonymised data set publically.
- All data unpinning publications should be made publically available when legally, commercially, and ethically appropriate (this includes data underpinning PGR theses which are 'published' via ORE).
- Published research papers should include a short statement describing how and on what terms any supporting research data may be accessed.
- Data can also be stored outside of the ORE archive (e.g., NHS servers; research team servers) but should be registered in ORE with a metadata record which includes a link or information about where the data is stored and details about conditions for access. Each metadata record is assigned a 'handle' or permanent identifier which can be linked to from other sources (e.g. from a research paper).
- Researchers/PGRs are not obligated to share confidential/sensitive or commercial data, and PIs/PGR supervisors can be the gate-keepers of sharing data (e.g., decide who to share the data with and for what purpose, post-project)
- It is likely that publically available datasets will be included in the next REF evaluation
- The current ORE plan is to archive data for at least 10 years after the last 3rd party access. If data is stored elsewhere and accessed via a metadata record, then the PI/PGR supervisor will need to keep a record of access requests and inform the ORE

5. **Publications:**
   - Copies of peer reviewed publications are to be stored on ORE
   - If funded by RCUK/Wellcome please refer to funder policy on research outputs (publications). For example, CC-BY licence policy.
   - Researchers (with support from the library) should check the publisher's policy on open access. A helpful website is SHERPA/RoMEO. Researchers might only be able to make the manuscript (or post-print version) available (i.e., not the final published version). It is therefore good practice to keep the post-print version of your paper.
   - Researchers should upload their publications via Symplectic (see Sympletic Guide). PGR students need to upload publications directly via ORE (see OA Guide for PGRs).
   - RCUK and Wellcome Trust provide funding to pay the publisher for open access. See How to Apply for Open Access Funds.
   - PGR theses are to be stored in ORE (on open access when appropriate)
   - For PGR theses where data have not been published prior to submission of the thesis, then an embargo can be set for 18-months or 5-years (and can be renewed with supervisor permission) , during this time the thesis is stored on ORE but is not publically available (a metadata only record)

EXAMPLE CONTENTS FOR DATA MANAGEMENT PLAN

A Data Management Plan could include the following information (for an example, see attached document produced by the CAL research group)

- File structure and documentation (with examples of file labels) procedures
- Data formatting and version control procedures
- Data back-up procedures
- Data storage/access procedures

List the data to be stored	How the data will be accessed (open archive, dark archive, meta-data record)?	Where the original data will be stored (ORE open, ORE metadata record, NHS, research server, paper in locked cabinet)?	How long will the data be stored?
Raw data from cognitive assessments	Meta-data record	research server	10 years after last 3rd party access
Date of birth of participants	Meta-data record	research server	10 years after last 3rd party access
Date of assessments	Meta-data record	research server	10 years after last 3rd party access
Sex of participants	Meta-data record	research server	10 years after last 3rd party access
Patient vs. Control status of participants	Meta-data record	research server	10 years after last 3rd party access
Name of participants	Meta-data record	Paper in locked cabinet; research server if consent to volunteer panel	Panel membership consent will be reviewed every 5-years. If not on the panel, data will be destroyed upon completion of the study.
Contact details of participants	Meta-data record	Paper in locked cabinet research server if consent to volunteer panel	Panel membership consent will be reviewed every 5-years. If not on the panel, data will be destroyed upon completion of the study.

Some additional resources:

1. RCUK Common Principles on Data Policy in 1.1 :
   http://www.rcuk.ac.uk/research/Pages/DataPolicy.aspx

2. Open Access Research and Research Data Management Policy (May 2013)
   http://hdl.handle.net/10036/4280

3. Open Access Research and Research Data Management Policy for PGR Students (January 2013)
   http://hdl.handle.net/10036/4279

4. UoE RDM pages
   http://as.exeter.ac.uk/library/resources/rdm/

5. Sherpa/FACT
   http://www.sherpa.ac.uk/fact/

6. Finch Report
   http://www.researchinfonet.org/wp-content/uploads/2012/06/Finch-Group-report-FINAL-VERSION.pdf