

LJMU Research Data Policy: information and guidance

Prof. Andrew J Young

Director of Research

April 2013

Aims

This document outlines the University policy and provides advice on the treatment, storage and sharing of research data for members of staff and research students of the University.

In doing so it addresses the following elements that, increasingly, are a prerequisite of research grant funding by gathering together aspects of best practice from the sector:

- Managing Research Data
- Data Management Planning
- Ethics and Freedom of Information
- Commercialisation
- Practical Aspects for Data Storage
- Sharing Research Data and Open Access
- Institutional Digital Repository

This document considers each of the above elements and will be reviewed on at least an annual basis to ensure alignment with both national requirements and to take into account the University IT strategy and associated developments. The policy and guidance notes will be posted on the webpages of Research and Innovation Services.

CONTENTS

- MANAGING YOUR RESEARCH DATA **3**
- DATA MANAGEMENT PLANNING **6**
- ETHICS AND FOI **8**
- COMMERCIALISATION **9**
- PRACTICAL ASPECTS FOR DATA STORAGE **9**
- SHARING YOUR DATA – OPEN ACCESS **11**
- INSTITUTIONAL DIGITAL REPOSITORY **12**

MANAGING YOUR RESEARCH DATA

Why manage your research data?

If you are starting a new research project, then you will need to consider the management and possibly subsequent sharing of the research data you generate. The major funders of research in the UK now generally require that you address specific aspects of data management, often requesting a data management plan be developed as part of your grant application.

However, you should consider this data management whether the project is externally funded or not as there are considerable benefits to be gained from considering this at the outset of any project:

- Funding body requirements are met
- Help ensure research data integrity and accuracy
- Research results may be replicated by you and others
- Minimises the risk of data loss or corruption
- Prevent duplication of effort by other researchers
- Compliance with best practice in industry

Data Management Requirements for major research funders

Most research funders have data sharing policies in place obliging award holders not only to publish but to share their data online in an open access repository. You should find out whether raw data forms part of archiving mandates using the [Sherpa JULIET](#) service. Most research funders have some form of policy regarding managing research data, but can vary greatly. Again, check the [Sherpa JULIET](#) website for guidance.

What constitutes research data?

The definition of research data will depend on your discipline. If you are a social scientist, your data may be survey results, interviews or statistics. Equally, scientists may compile 3D models, lab data, field data etc.

The Research Information Network classify research data as follows:

- Observational:** data captured in real-time, usually irreplaceable. For example, sensor data, survey data, sample data, neuroimages.
- Experimental:** Data from lab equipment, often reproducible, but can be instrument specific. For example, gene sequences, chromatograms, toroid magnetic field data.
- Simulation:** data generated from test models where model and metadata are more important than output data. For example, climate models, economic models.
- Derived or compiled:** data is reproducible but expensive. For example, text and data mining, compiled database, 3D models.
- Reference or canonical:** a (static or organic) conglomeration or collection of smaller (peer-reviewed) datasets, most probably published and curated. For example, gene sequence databanks, chemical structures, or spatial data portals.

Research data can take many forms, for example:

- Documents (text, Word), spreadsheets
- Laboratory notebooks, field notebooks, diaries
- Questionnaires, transcripts, codebooks
- Audiotapes, videotapes
- Photographs, films
- Test responses
- Slides, artefacts, specimens, samples
- Digital objects acquired and generated during the process
- Data files
- Database contents (video, audio, text, images)
- Models, algorithms, scripts
- Contents of an application (input, output, logfiles for analysis software,

- simulation software, schemas)
- Methodologies and workflows
- Standard operating procedures (SOPs)

The following **research records** may also be important to manage during and beyond the life of a project:

- Correspondence (electronic mail and paper-based correspondence)
- Project files
- Grant applications
- Ethics applications
- Technical reports
- Research reports
- Master lists
- Signed consent forms

DATA MANAGEMENT PLANNING

Do I require a Data Management Plan as part of my funding application?

An increasing number of research funders formally require a data management plan to be submitted as part of a grant application (see the national [Sherpa JULIET](#) website, provided by the University of Nottingham for guidance from several major funders). In some cases, a full plan is not needed but you may be asked to outline your plans for the treatment and ultimately the sharing of data.

For an example, see EPSRC's expectations on the handling of research data that arises out of studies they fund:

<https://www.epsrc.ac.uk/about/standards/researchdata/>

What is a Data Management Plan?

A data management plan is developed at the commencement of a project. It describes the following:

- What research data will be created.
- What policies (funding, institutional, and legal) apply to the data.
- What data management practices (backups, storage, access control, archiving) will be used.
- What facilities and equipment will be required for data storage and handling (hard-disk space, backup server, repository).
- Who will own and have access to the data over the data lifecycle.
- Who will be responsible for each aspect of the plan.
- How its reuse will be enabled and long-term preservation ensured after the original research is completed.

How do I create a Data Management Plan?

LJMU recommends the use of an on-line tool for management plan development, namely DMP online (<https://dmponline.dcc.ac.uk>). This JISC tool permits researchers preparing grant applications to develop a management plan quickly and efficiently and ensure tailored fit with the requirements of the major funders. On-line registration is free to anyone in the HE sector.



Many funders of research provide their own guidelines for developing data management plans, for example:

NERC: <http://www.nerc.ac.uk/research/sites/data/dmp.asp>

Alternatively, you can adopt the following checklist to ensure your plan addresses the key points needed for any viable plan:

- What type of data will be produced? Will it be reproducible? What would happen if it got lost or became unusable later?
- How much data will be generated and how often will it change?
- Who will be the audience for your data and how will they use it now, and in the long run?
- Who controls it (Principal Investigator, yourself, computing officer in your school)?
- How long should it be retained? For example, 5 years, up to 10 years, or permanently (see below).
- Are tools/ software needed to analyse, process or visualise the data?
- Are there any special privacy or security requirements? For example, personal data, high-security data?

Are there any sharing requirements? e.g. funding body's data sharing policy?

- Are there any other funding body requirements? e.g. data management plan in proposal?
- Is there good project and data documentation?
- What directory and file naming convention will be used?
- What project and data identifiers will be assigned?
- What file formats will be used? Are they long-lived?
- What will be the storage and backup strategy?
- When and where the data will be published?
- Is there a community standard for data sharing/integration?

Are there ethical and legal issues to consider?

Please see: <https://www2.ljmu.ac.uk/RGSO/93042.htm> to learn more about the University's policy on the ethical review of research projects involving human participants.

Personal or sensitive data may **not be suitable for sharing** with others, depending on whether informed consent has been obtained from participants. You may wish to consider anonymisation techniques or data aggregation for numeric data, editing of video or sound recordings, use of pseudonyms in qualitative data etc.

How does Freedom of Information and Data Protection Act affect data management?

For Freedom of Information and Data Protection Act requirements,
Contact the Data Protection Officer
Email: foi@ljmu.ac.uk
Phone: 0151 231 3116

How do I determine whether my data is commercially exploitable?

Access to some research data may be legitimately restricted, e.g., due to commercial sensitivities. You should therefore first determine

whether your research data is commercially exploitable. There may be restrictions or delays on data sharing needed to protect the IP, copyright or patentable data. In the first instance please contact a member of the Knowledge Exchange and Commercialisation Team to discuss this: <https://www2.ljmu.ac.uk/RIS/128814.htm>

Intellectual Property

Having a secure and comprehensive data set is an important element in securing the intellectual property rights for any work (e.g., obtaining and defending patents). However, clearly such data may not be suitable for sharing. For copyright and intellectual property issues, please contact the University IP Officer directly: (e.nolan@ljmu.ac.uk).

How long do I need to retain my data?

You should allow for your data to be securely preserved for **at least 10 years** from the last date when it was viewed by a third party.

As a condition of receiving external research funding from the UK Research Councils and some other sponsors of research, the University is required to support effective data curation throughout the full data lifecycle.

PRACTICAL ASPECTS FOR DATA STORAGE

What are the best formats for data storage and sharing?

For more information on data formats and sizing, see the UK Data Archive page on [data formats and software](#).

What is metadata and why do I need to use it?

Data needs to be properly organised in order to make it useable by yourself **and others**. This should be done on a day-to-day basis as part of your ongoing research.

Metadata (often referred to as 'data about data') can be used to help define the data sets in such a way that large data sets are readily searchable. To do this metadata tags are used to assist the user. These fall into three broad categories:

Descriptive: Commonly used fields such as title, author, abstract and keywords.

Administrative: Preservation, rights management and technical metadata about formats.

Structural: How different components of a set of associated data, such as tables in a database, relate to one another.

Metadata standards and documentation are discipline-specific – please consult the UK Data Archive pages on [Metadata and Documentation](#).

How do I backup my data in the short-term? What about security?

Academic Planning & Information Services (APIS)

(<https://www2.ljmu.ac.uk/APIS/index.htm>) should be your first port of call in the University for storage of 'active' research data. We strongly recommend that you use the University network for storage and in doing so safeguard against data loss in the event of computer failure.

Using the Networked facilities will mean that your data is automatically backed up and is both safe and secure. It is ideal for the long-term storage of University files and both interim (i.e., on-going) and final data sets. The service is free to postgraduate research students and staff. Additional storage space on the University server can be obtained as necessary for research data: see

<https://www2.ljmu.ac.uk/ITHelp/disk-storage/index.htm>

Removable devices such as pen drives, CDs and DVDs are not recommended as backup devices due to the risk of disk failure.

The University plans to make available the institutional digital repository ("LJMU Research Online") for longer-term data storage and sharing. Further details will be made available once this functionality is available.

SHARING YOUR DATA – OPEN ACCESS

How do I disseminate and share my data via open access repositories?

Many funders now stipulate that you must share your raw data as well as your final research outputs in an open access repository of some sort. For research projects funded by the UK Research Councils the normal expectation is that this happens within 12 months of the data being generated.

Data repositories - for a list of discipline-based repositories for sharing data and long-term data curation - see

http://oad.simmons.edu/oadwiki/Data_repositories

Data Centres - some of the UK Research Councils fund data services to curate, disseminate, and preserve data created as part of their funded programmes, e.g., the ESRC's UK Data Archive (ESDS - Economic and Social Data Service) and the designated data centres from the Natural Environment Research Council (NERC).

INSTITUTIONAL DIGITAL REPOSITORY

For details on the University's Digital Repository for research (LJMU Research Online) please go to: <https://www2.ljmu.ac.uk/lea/128045.htm>