Managing your Research Data

Introduction:

- Data Management plans
  - Data formatting
  - File management
  - Data sharing & backup

And what is 'Data' ?

- **Raw instrument readings**
- **Processed data**
- **Analysed data**
- **Genomic data**
- **Microscopic photos, Western blot images and measurement**
- **Spreadsheets**
- **Data Formats/Metadata**
- **Videos**
- **Surveys and interviews**
- **Field notes**
- **Maps**
- **Lab books**
- **Physical samples**
- **Protocols**
- **Software**
- **Graphs/Figures**

It's basically anything you produce in the course of your research and is the 'bed-rock' of your findings!

What is a Data Management Plan?

A short document (typically 2-3 pages) which contains information about:

- Data sources
- Data input methods (validation)
- Back up strategies (Consider & select option(s))
- Plans for data sharing (Suitable repository & appropriate access controls – FAIR data principles)
- Ethical/legal restrictions (e.g. Patient Data)
- Who is responsible for the data

Researcher needs to write a DMP before the project begins

Why do I need a Data Plan?

- To protect your science
- Re-usable data
- Reproducible better science
- To help your future you and your current colleagues
- Because major funders require it
- University policy
- Legal requirements e.g. use of personal data or confidentiality agreements

http://www.dcc.ac.uk/resources/data-management-plans

University of Cambridge

Research Data Management Policy Framework

www.data.cam.ac.uk/DMPsupport

www.data.cam.ac.uk/funders

University Staff and Students:

13. Retain intellectual property rights where they arise or the right to apply for such rights arise from the results of activities undertaken by University staff in the course of their employment by the University and by students in the course of their study at the University in accordance with Chapter XIII of the University's Statutes and Ordinances on Finance and Property, subsection Intellectual Property Rights[1]
14. Should be aware when considering whether they may want to commercialise the results of their research that, in respect of patents and similar rights in inventions and new technology, protection for and subsequent commercialisation of such inventions may be jeopardised if information about the inventions is publicly disclosed before all relevant applications for protection have been lodged. A public disclosure would include depositing research data in a publicly accessible discipline-based or institutional repository.
15. Are encouraged to do the following at the design stage of a research project:
    a. Prepare a Data Management Plan, in accordance with guidance provided by the University of Cambridge[7] and the Digital Curation Centre[8] (DCC). If funders require a Data Management Plan, such plan needs to be prepared according to funders' requirements[4].
    b. Ensure that legal, ethical and commercial constraints on release of research data are considered at the initiation of the research process and throughout both the research and data life cycles, which shall be described in the data management plan.
    c. Allocate appropriate resources (time and financial resources) for data management in their grant proposal.

https://www.data.cam.ac.uk/university-policy

Data Protection issues

The legislation imposes requirements for use of personal data, but most research will be subject to one of two exemptions:

1) Research purposes – used where:

- The standard provisions would seriously impair research
- No damage or distress to data subjects
- No individual decision-making about data subjects
- Safeguards are in place

2) Academic expression - used where:

- Complying with the standard provisions would be incompatible with the academic purpose
- There will be a publication in the public interest

Patient identifiable Data (has more requirements)

- Even more sensitive
- Protects patients living and deceased unlike Data Protection Act which focuses upon living.
- Very seldom allowed to breach except in interests of patients health and well-being

https://digital.nhs.uk/data-and-information/looking-after-information

Data protection requirements upon researchers

- Know and communicate your legal basis (**research**)
- Be transparent with data subjects (**research**)
- Process accurately and only what you need (**research**)
- Keep personal data secure (**research**)
- Process fairly, considering any ethical risks to the data subject (**research**)
- Comply with institutional accountability processes, e.g. ethical review (**research** and **academic expression**)

Storage of Personal data

- Personal data may only be used under the research exemption if appropriate security measures are taken.
- Ideally this will be prompt anonymization or pseudonymisation.
- Ensure that portable systems or devices are fully encrypted.
- Ensure physical security for hard copy data.

(More about storage and backup later in course)

Handling Personal data

- If possible, don't collect it!
- Gain informed, preferably open and written, consent. See UK Data Service for examples.
- Anonymise data
  Remove identifiers, Aggregate results, Generalise a variable, Remove outliers
- Use managed access repositories

Where to go for further help and advice on your Data

Your local IT Support desk

OSC
Office of Scholarly Communication

Tweets as @CamOpenData

https://www.data.cam.ac.uk

NB I have used materials from OSC in this slidedeck

UIS

https://www.uis.cam.ac.uk

Data Champions

https://www.data.cam.ac.uk/intro-data-champions

Guidance

Institutional website, e.g:

University Research Ethics website:

https://www.research-integrity.admin.cam.ac.uk/research-ethics

NHS Health Research Authority

https://www.hra.nhs.uk/hra-guidance-general-data-protection-regulation/

UK Data Service FAQs

https://www.ukdataservice.ac.uk/media/621794/gdpr-faqs.pdf

Online courses:

- Mantra from Edinburgh Uni

https://mantra.edina.ac.uk/

- Foster Open Science

https://www.fosteropenscience.eu/

- UK Data Service

https://www.ukdataservice.ac.uk/manage-data/training