Managing Research Data and Data Management Planning

John Southall
Bodleian Data Librarian

Cathy Scutt
Bodleian Education Librarian

Trinity Term
(May 2021)
Overview

Help you think about practical issues to do with creating, working with and securing data

Show how ‘data management’ benefits you

Highlight what support is available

Discuss how to apply these ideas through planning

Link to Open and Reproducible Research
University’s policy on research data management (RDM)

There is a formal policy on research data management (RDM) and records:

“The University of Oxford seeks to promote the highest standards in the management of research data and records as fundamental to both high quality research and academic integrity.”

See Research Data Oxford (RDO)

http://researchdata.ox.ac.uk/
Policy in practice

No one would disagree with the policy as it is worded, but in practice what does it mean?

‘Management’ - Protecting ourselves as projects develop or things go wrong;

Laptop lost or stolen
Hard drive crash
Funding terminated
Team members disperse
Memory fails
Management Policy

Research data and records should be:
   a. Accurate, complete and reliable;
   b. Identifiable, retrievable, and available when needed;
   c. Secure and kept in an appropriate manner
Impact on You

Responsibility is yours as data creators to be aware of policy
Combine with wider research skills development
**Respond** to funder/publisher expectations
Make use of the **support framework** at Oxford
Right from the start
“Overall, doing research robustly and fairly does not necessarily require more money, it simply requires that you think before you start.”

Ben Goldacre, Bad Science (2008)
Some Principles of Data Management

How can RDM help with these concerns?

Data Management Planning and the research lifecycle

All stages of research
- Before – During - After

Data and Metadata

Usable for you now (secure storage)

In the future
- Accessible to you
- Preserved for you
Working with Data?

Typical examples?
- Born digital
- Or digitised
- Used and unused
Digital Media - strengths

Digital – a key factor in RDM

Strengths of digital
- Perfect copies
- Easy to share
- Convenient

But there are weaknesses to be managed
Digital Media - weaknesses

Weaknesses of digital
- **Too** easy to share
- Medium dependent
- Corrupted - Immediate loss
- Inflexible – difficult to repurpose
- Too many copies
- Hardware and software dependent
- Long term use issues - Digital obsolescence and fragility
- Ethical and licensing issues

Curate it
- Collaborate on data management
Management takes Shape

Day to day protocols on collection and use

Disaster planning
  ◦ Multiple storage and backups
  ◦ Data security

Appropriate workflow?

Documentation – Metadata
  ◦ Natural offshoot of literature search/ research diary?

Formalisation of procedures ensures preservation

Preservation as basis of sharing with others
Getting Support for RDM

Important distinctions in managing your data and how others view it

Dynamic or static?

See where and when support fits in

During or after the project?
- One Drive for business (during but not after)
- HFS (during but not after)
- ORA-Data
- RDO on other options
- External archival services
http://researchdata.ox.ac.uk/
Support Frameworks (1/3)

You are not left to figure this all out yourself!

At Oxford:
- The Library and its Subject Consultants
- Departmental level support
- Research Skills training
- Research Data Oxford webpage
- Research Data Oxford email
- ORA / ORA-Data
Support Frameworks (2/3)

Ethical and legal issues
- Creating data – live participants
- Curec
- Collecting data – Licensing

Research Ethics
- Access Restrictions
- Participation/ Confidentiality agreements
- Ethics Committees and Informed Consent
Support Frameworks (3/3)

Outside Oxford:
- Digital Curation Centre
- UK Data Archive/Service
- Edinburgh – MANTRA Course
- Publications (Rice & Southall 2016, Angus Whyte 2014)
- websites
Mapping of RDM Services
Open and Reproducible Research

Key trend in academic research around:
- Preservation of materials
  - Data / Software / Methodology
- Documentation of research
- Transparency in methodology
- Potential for provenance, verification etc.
- Augmentation

All research can be managed and planned
Not all research is Open or Reproducible
RDM Planning

Data Management Plans and Planning (DMP)
Growing popular with funders and publishers

What is it?
About applying RDM principles
Formalising previously informal stages
See chapter 7 (Rice & Southall 2016)
‘Research Data Oxford’ pages
DMP – an outline

Describes the research data being created or collected
Key responsibilities
How the data will be organised
Disaster recovery
Documentation during the collection and analysis phase
Tools
Plan v. Planning - a considered approach
Other Elements of DMP

Policy on data storage and security
What facilities and equipment will be required
How stakeholder requirements being addressed
How / If the data will be preserved
How / If the data will be shared
ORA-Data Pre-deposit checklist

DMPOnline tool
In Conclusion

Beneficial to you
Make more efficient use of data
Protect against common problems of ‘fragile’ digital data
Increase citations/ impact of research
Respond to funder, publisher expectations:
  ◦ Good Data management
  ◦ Open research
  ◦ Reproducible Research