Introduction to Research
Data Management

Eli Harriss
Knowledge Centre Manager
& Outreach Librarian
AIMS FOR TODAY

• Understand key principles of RDM and organising your data effectively
• Understand common dangers and pitfalls of digital data
• Introduce the University’s research data policy and outline the practical impact this will have on your work
• Understand institutional, funder and publisher requirements
• Be able to draft a data management plan
• Understand issues around preserving data and cybersecurity
• Be aware of ORA-Data, Github and other preservation services
• Share thoughts and insights about the potential of data management in your own field
• Introduce the services available at Oxford to assist you
Media carriers

All of these different types of media carrier are present in Bodleian Libraries' archives. Can you identify any of them?
What is research data management?

- Storage
- Organizing
- Structuring
- Choosing technology
- Preservation
- Versioning
- Documenting
- Sharing
- Curation
- Security
- Backing up

Research data management – a very brief introduction
Typical questions for us to pose

Organization and storage
• Can you find what you need, when you need it?
• Is your data stored securely? Do you have a back-up plan in place?

Documentation and record keeping
• Is it clear what everything is, and what’s been done to it?
• Will information remain intelligible if you return to it later?
DataONE Data Lifecycle
DDI version 3.0 Combined Life Cycle Model
Why?
Manchester cancer hospital fire 'may have destroyed vital research'

Cancer Research UK institute likely to have lost millions of pounds of life-saving equipment in blaze, says its director
File naming conventions

TILS Document Naming Convention

Document naming for the TILS Division should follow this convention:

GDL_TILSDocNaming_V1_20090612.docx

- **A prefix** shows the document type
- **The document title** describes the content
- **The version number**
- **The date** in the format yyyyMMdd

http://www.data.cam.ac.uk/files/gdl_tilsdocnaming_v1_20090612.pdf
## File naming conventions

<table>
<thead>
<tr>
<th>Folder Path</th>
<th>File Name</th>
</tr>
</thead>
<tbody>
<tr>
<td>20160706_Presentations</td>
<td>20160706_Presentations</td>
</tr>
<tr>
<td>20160726_RDMworkshopForLibrarians_MillLane</td>
<td>20160726_RDMworkshopForLibrarians_MillLane</td>
</tr>
<tr>
<td>20160802_Altmetrics</td>
<td>20160802_Altmetrics</td>
</tr>
<tr>
<td>20160804_Bibliometrics</td>
<td>20160804_Bibliometrics</td>
</tr>
<tr>
<td>20160810_RDMworkshopForLibrarians_MilsteinRoom</td>
<td>20160810_RDMworkshopForLibrarians_MilsteinRoom</td>
</tr>
<tr>
<td>GDL_DMP_V8_2016511.docx</td>
<td>GDL_DMP_V8_2016511.docx</td>
</tr>
<tr>
<td>GDL_DMPForLibrarians_V9_20160726.docx</td>
<td>GDL_DMPForLibrarians_V9_20160726.docx</td>
</tr>
<tr>
<td>GDL_ExampleDMP_V9_20160511.docx</td>
<td>GDL_ExampleDMP_V9_20160511.docx</td>
</tr>
<tr>
<td>LST_RDMforLibsAttendees_V1_20160808.pdf</td>
<td>LST_RDMforLibsAttendees_V1_20160808.pdf</td>
</tr>
<tr>
<td>MEM_RDMforLibs2Feedback_V1_20160811.docx</td>
<td>MEM_RDMforLibs2Feedback_V1_20160811.docx</td>
</tr>
<tr>
<td>PRE_DataLossScenarios_V5_20160726.docx</td>
<td>PRE_DataLossScenarios_V5_20160726.docx</td>
</tr>
<tr>
<td>PRE_RDMforLibsSigns_V1_20160808.pptx</td>
<td>PRE_RDMforLibsSigns_V1_20160808.pptx</td>
</tr>
<tr>
<td>PRE_RDMWorkshopForLibrarians_V3_20160726.pptx</td>
<td>PRE_RDMWorkshopForLibrarians_V3_20160726.pptx</td>
</tr>
<tr>
<td>PRE_RDMWorkshopForLibrarians_V4_20160810.pdf</td>
<td>PRE_RDMWorkshopForLibrarians_V4_20160810.pdf</td>
</tr>
<tr>
<td>PRE_RDMWorkshopForLibrarians_V4_20160810.pptx</td>
<td>PRE_RDMWorkshopForLibrarians_V4_20160810.pptx</td>
</tr>
<tr>
<td>20160914_RDMWorkshopforGLS</td>
<td>20160914_RDMWorkshopforGLS</td>
</tr>
<tr>
<td>20161004_PhDTraining</td>
<td>20161004_PhDTraining</td>
</tr>
</tbody>
</table>
Exercise 1

Chat to someone else about these problems and the solutions
The University’s research data policy and the practical impact this will have on your work...

Researchers are responsible for:
1) Developing and documenting clear data management procedures
2) Planning for the ongoing custodianship of the data
3) Ensuring that legal, ethical, and funding body requirements are met

The University is responsible for:
1) Providing access to the services, facilities, and support needed to allow researchers to comply with the policy
University of Oxford Policy

In December 2018, the University of Oxford’s Research and Innovation committee approved a new Policy on the Management of Data Supporting Research Outputs. This superseded the Policy on the Management of Research Data and Records, approved by Council in July 2012.

Summary of key points:

- The research data covered by the policy is the information needed ‘to support or validate a research project’s observations, findings or outputs’
- Research data should be:
  - Accurate, complete, identifiable, retrievable, and securely stored
  - Where possible, able to be made available to others
- Research data should be retained for ‘as long as it has continuing value’ – but a minimum of three years after publication of research
  - Specific requirements from funders take precedence
- The policy applies to all members of the University who are engaged in research
- Researchers are responsible for:
  - Developing and documenting clear data management procedures
  - Planning for the ongoing custodianship of their data
  - Ensuring that legal, ethical, and funding body requirements are met
- The University is responsible for providing access to the services, facilities, and support needed to allow researchers to comply with the policy

http://researchdata.ox.ac.uk/home/introduction-to-rdm/university-of-oxford-policy/
http://researchdata.ox.ac.uk/funder-requirements/
Curate it

Draft a data management plan

• Day to day protocols on collection and use
• Disaster planning!
  o Multiple storage and backups
  o Data security
• Appropriate workflow?
• Documentation and metadata
• Formalisation of procedures ensures preservation
• Preservation as a basis of sharing
DMP

Describes the research data being created or collected
Key responsibilities
How the data will be organized
Disaster recovery
Documentation during the collection and analysis phases
Tools

Plan.. Versus Planning = a living document
DMP

Policy on data storage and security
What facilities and equipment will be required
How stakeholder requirements will be addressed
How or if the data will be preserved
How or if the data will be shared

DMPOnline tool
http://www.dcc.ac.uk/resources/data-management-plans
Benefits to you

• Make more efficient use of data
• Protect against common problems of ‘fragile’ digital data
• You keep control of your data
• .. Now and in the future
• Increase citations and the impact of your research!
What does a DMP cover?

- Type of data collected
- Data storage
- Data back-up
- Metadata standards
- Sensitive information and how this will be handled
- Plans for data sharing
http://researchdata.ox.ac.uk/home/managing-your-data-at-oxford/data-management-planning/

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


Exercise 2

Write a Data Management Plan for your research project;

talk to someone else about it.
Decide how to manage securing your research data. Consider the importance of password safety guidelines and how to encrypt and destroy sensitive data when required.
ORA-DATA
https://ora.ox.ac.uk/deposit

GIT HUB
https://github.com/

Clinical Trial Data Repositories!
https://edctpknowledgethehub.tghn.org/data-sharing-toolkit/repository-finder/
Exercise 3

Talk to someone else – someone new? – about the potential of data management in your field, for your own research
Tools, services, and resources

- **DMP Online** – templates for data management planning
- **HFS** – central back-up service
- **ORA-Data** – Oxford’s institutional data archive
- Visit [Research Data Oxford](https://researchdata.ox.ac.uk) for more
https://libguides.bodleian.ox.ac.uk/workshops/workshopsbydate

Research data management

- Bodleian iSkills: Managing research data and Data Management Planning (DMPs) (Next session TBA) more...
- Bodleian iSkills: Data sources for research - discovery, access & use (Tues 5 Nov 10.00-12.00) more...
- Bodleian iSkills: Introducing MANTRA for Research Data Management (Tues 12 Nov 10.00-11.30) more...
- Bodleian iSkills: Working with sensitive research data in the Social Sciences and Humanities (Mon 18 Nov 14.00-16.00) more...
- Bodleian iSkills: Working with sensitive research data in the Sciences and Medical Sciences (Tues 26-Nov 14.00-16.00) more...
- Bodleian iSkills for the Medical Sciences Division and OUH staff: Introduction to Research Data Management (Tues 12 Nov 14.00-15.30; repeated Mon 25 Nov 14.00-15.30) more...
Research data management plans: How to write one
Data management plans (DMPs) are becoming an increasingly important aspect of research - many funding bodies now require one as part of a grant application. This course provides an overview of the basics of data management planning, plus a chance to try some online tools for building your own DMP.
Objectives
Create data management plans
Explore tools for building a DMP
Understand research data management
Date 04/02/2019 15:00 - 16:30
Cost Staff £15/Student £7.50 [Book and pay]
Research data: What you need to know
Taking good care of your data is an essential part of doing good research. This course gives an overview of some key areas to think about, tips for avoiding common pitfalls, and information about the tools and support services that are available to help.
Objectives
Understand the key issues to consider when managing research data
Learn about University and funding body policies, and the responsibilities these create
Discover what tools, resources, and support services are available, within the University and beyond
Date 11/02/2019 12:30 - 13:30
Last words

Ask yourself…
Will I want to revisit this data in ten years’ time?
Will it be intellectually possible? Is there metadata to help?
Will it be technically feasible? Will the files be findable, openable?
Is it ethically and legally possible?
What am I expected to do?
What do I think is appropriate?
Do I have documentation? Do I have a Data Management Plan?
This is what new DPhils and early career researchers need to know:

- How to develop a DMP: 5 questions, (keep it simple). Eg where is your documentation about your data? Back up and storage plans? Thought about data from start to end – are you collecting the right data to answer the question? Are your variables named clearly, using a controlled vocabulary?

- Data curation – types of errors, detecting anomalies (simple techniques, software tools and statistical methods), fixing anomalies: logic check! If x chemical was used, and it can never go above a certain temperature, then if the temperature for x chemical is above that temperature, that’s bad data. Build into your data collection ways to fix the anomalies: use your time well, because errors will always be an issue which especially will eat into the analysis stage. Take time at the start to plan data collection, curation, storage etc.

- Data storage, backup and disaster recovery – simple!

- Archiving and preservation of data


- Basics of data de-identification and anonymization – to maintain the usefulness of the data// this is a whole other iSkills session.
BEFORE
Your Research

- Create a Data Management Plan (DMP)
- Create Data
- Make your Data FAIR

DURING
Your Research

- Store & Secure your Data
- Create Metadata
- Gain Informed Consent

AFTER
Your Research

- Share your Research Data
- Preserve your Research Data
- Licence your Data

http://libguides.ucd.ie/data
Feedback

Help

researchdata@ox.ac.uk
http://researchdata.ox.ac.uk/healthcheck/
eli.harriss@bodleian.ox.ac.uk