Data Management: The New Frontier for Libraries

Greg Simpson – Research Data Officer
### (potential) library services in RDM

<table>
<thead>
<tr>
<th>Stage</th>
<th>Services</th>
<th>Other stakeholders</th>
</tr>
</thead>
<tbody>
<tr>
<td>DMP Planning</td>
<td>• Assistance with data management planning</td>
<td>• Researchers</td>
</tr>
<tr>
<td></td>
<td>• Training graduate students and researchers about data management</td>
<td>• IT department (if DMP Online)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Administrators</td>
</tr>
<tr>
<td>Collection/production of data</td>
<td>• Help with organizing data, adopting standards and ensuring comprehensive metadata</td>
<td>• Researchers</td>
</tr>
<tr>
<td></td>
<td>• Local storage</td>
<td>• IT department (if big data)</td>
</tr>
<tr>
<td>Analysis</td>
<td>• Tools for active data management (e.g. Open Science Framework or <a href="http://www.eric-project.org">http://www.eric-project.org</a>)</td>
<td>• Researchers</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• IT department (if big data)</td>
</tr>
<tr>
<td>Data dissemination and sharing</td>
<td>• Managing a data repository</td>
<td>• Researchers</td>
</tr>
<tr>
<td></td>
<td>• Minting/assigning DOIs</td>
<td>• IT department</td>
</tr>
<tr>
<td></td>
<td>• Indexing and registering datasets</td>
<td>• Administrators</td>
</tr>
<tr>
<td></td>
<td>• Licensing and privacy issues</td>
<td>• Ethics boards</td>
</tr>
<tr>
<td>Re-use</td>
<td>• Data discovery</td>
<td>• Researchers</td>
</tr>
</tbody>
</table>
Practical steps researchers can take

- Write a data management or a sharing plan

- Make sure data are shareable and can be understood:
  - Obtain consent to share
  - Do not share identities without consent
  - Use open and standard formats
  - Provide context and documentation
  - Protect your data at all stages
The research data lifecycle

**Creating Data:**
- Designing research
- DMPs, planning consent
- Locate existing data
- Data collection and management
- Capturing and creating metadata

**Processing Data:**
- Entering, transcribing
- Checking, validating and cleaning data
- Anonymising data
- Describing data
- Managing and storing data

**Analysing Data:**
- Interpreting, deriving data
- Producing outputs
- Authoring publications
- Preparing for sharing

**Preserving Data:**
- Data storage
- Backup & archiving
- Migrating to best format & medium
- Creating metadata and documentation

**Re-Using Data:**
- Follow-up research
- New research
- Undertake research reviews
- Scrutinising findings
- Teaching & learning

**Access to Data:**
- Distributing data
- Sharing data
- Controlling access
- Establishing copyright
- Promoting data

Ref: UK Data Archive: http://www.data-archive.ac.uk/create-manage/life-cycle
Funding bodies’ requirements

- Funders are increasingly requiring researchers to meet certain data management criteria.

- When applying for funding, you may need to submit a technical or data management plan.

- You may be asked to make your data available through an archive.

- **SHERPA JULIET** provides a useful summary of funding bodies’ policies.
## Data-related policies in the UK

<table>
<thead>
<tr>
<th>Research Funders</th>
<th>Published outputs</th>
<th>Data</th>
<th>Time limits</th>
<th>Data plan</th>
<th>Access/sharing</th>
<th>Long-term curation</th>
<th>Monitoring</th>
<th>Guidance</th>
<th>Repository</th>
<th>Data centre</th>
<th>Costs</th>
</tr>
</thead>
<tbody>
<tr>
<td>AHRC</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
</tr>
<tr>
<td>BBSRC</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
</tr>
<tr>
<td>CRUK</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
</tr>
<tr>
<td>EPSRC</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
</tr>
<tr>
<td>ESRC</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
</tr>
<tr>
<td>MRC</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
</tr>
<tr>
<td>NERC</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
</tr>
<tr>
<td>STFC</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
</tr>
<tr>
<td>Wellcome Trust</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
</tr>
</tbody>
</table>

Common DMP questions

• What data will be created (format, types and size) and how?
• How will the data be documented and described?
• How will you manage ethics and Intellectual Property?
• What are the plans for data sharing and access?
• What is the strategy for long-term preservation?
http://dmponline.dcc.ac.uk

Create a new plan

Please select from the following drop-downs so we can determine what questions and guidance should be displayed in your plan.

If you aren't responding to specific requirements from a funder or an institution, select here to write a generic DMP based on the most common themes.

If applying for funding, select your research funder.

Otherwise leave blank.

Contact us  |  Terms of use  |  DM Pon Online previous version
© 2004 - 2017 Digital Curation Centre (DCC)
What does DMPonline do?

A web-based tool that enables users to...

i. **Create, store and update** multiple versions of Data Management Plans across the research lifecycle

ii. **Meet a variety of specific data-related requirements** (from funders, institutions, publishers, etc.)

iii. **Get tailored guidance** on best practice and helpful contacts, at the point of need

iv. **Customise, export and share DMPs** in a variety of formats in order to facilitate communications within and beyond research projects

* N.B. The templates have varying degrees of endorsement from funders, stakeholder communities, etc. Plans are currently being discussed for UoE to create a customised template for researchers…
DMPonline: possible export formats

For human readership…
- Pleasant formatting
- Editable. Can be used in conjunction with (e.g. MS Sharepoint)
- Removes all formatting

For machine readership…
- Facilitates quick public sharing
- Compatible with API for linking with other systems
- Minimal formatting
Research Data Repositories

• Digital Research data is best preserved and published using a research data repository

• A repository is an online database service, an archive that manages the long-term storage and preservation of digital resources and provides a catalogue for discovery and access

• Most data repositories do not charge to deposit research data, though many require registration
Choosing a repository

- Best repository to choose for your research data will be a national data centre or discipline specialist repository, because they have the expertise and resources to deal with particular types of data.

Data Centres and portals:

- **UK Data Service** - funded by the ESRC
- **NERC** data centres
- **ADS** (Archaeology Data Service)
- **EMBL-EBI** (European Molecular Biology Laboratory - European Bioinformatics Institute)
Choosing a repository (cont.)

Registries of research data repositories:

- Re3Data provides a searchable directory of research data repositories.
- Research Pipeline is a guide to the world's free data.
- Biosharing.org a curated, searchable portal of data standards, databases and policies in the life, environmental and biomedical sciences.
- There are many discipline specific repositories available, whilst some are multi-disciplinary such as:
  - Figshare
  - Zenodo
  - Dryad
NIHR Data Policy

- Revised in April 2014, NIHR policy on Open Access now covers not only published peer-reviewed research articles, but also the research materials – such as data, samples and models – that underpin them.

- The NIHR, along with other UK biomedical research funders, is a partner in the European version of PubMedCentral (Europe PMC).

- Europe PMC provides a permanent and free-to-access online digital archive of the full text, peer reviewed research publications and datasets that arise from research funded by the NIHR.

- All NIHR researchers are now required to prepare and submit a statement on how underlying research materials, such as data, samples or models, can be accessed.

- Although the NIHR expects researchers to consider and plan for data access, the policy does not require that data must be made open, nor does it specify how long data must be retained for.
The guideline states:

Making clinical trial data sets available to investigators beyond the original research team can improve patient care, advance medical knowledge and provide better value for money from health research.

Data sharing achieves many important goals for the scientific community, such as:

• Reinforcing open scientific inquiry.
• Encouraging diversity of analysis and opinion.
• Promoting new research, testing of new or alternative hypotheses and methods of analysis.
• Supporting studies on data collection methods and measurement.
• Facilitating education of new researchers.
Data sharing and the NIHR Journals Library

Your final report should include a statement about your data sharing and accessibility. The statement should provide a clear and positive indication of where and when the data will be shared.

Possible responses might state that all available data:

- Can be obtained from the corresponding author
- Is included as an appendix to the report
- Can be obtained from the corresponding author via the (XXX) repository
The Data Access Statement

All research publications must include a statement on how the underlying data can be accessed (a "data access statement"). This is in line with RCUK and individual funder policies.

A data access statement should include the following key pieces of information:

- **How the data can be accessed**: where it can be downloaded from or who must be contacted to request access.
  
  This should always include either a web link (a DOI or other persistent identifier if possible) or a departmental/group email address (not a personal email address).

- **What conditions use of the data is subject to**: whether a general licence applies to all users, or whether a data sharing agreement must be entered into before access to the data is granted.
Examples of data sharing statements

Some examples of data sharing statements you might include in your report:

• ‘All available data can be obtained by contacting the corresponding author.’

• 'We shall make data available to the scientific community with as few restrictions as feasible, while retaining exclusive use until the publication of major outputs. Anonymised data will be deposited here <link> to encourage wider use.’

• ‘The data will be made available via <link/corresponding author> within <x> months of publication, this is due to <insert reason>.’

• 'Due to <insert reasons>, there is no data that can be shared.’
The Clinical Trials Regulations define the archiving requirements for Clinical Trials of Investigational Medicinal Products (CTIMPs).

All essential documents should be archived and this includes essential documents held by investigators, sponsors and others involved in the conduct of a clinical trial (including services departments such as pharmacy, laboratories and radiology).

Require the sponsor to appoint ‘named individuals’ for archiving to ensure all requirements are met and systems are in place to track and retrieve archived documents.

The Joint Project Notes on Archiving give further information on the storage and destruction of essential documents and the duration/timelines appropriate for archiving: http://www.ct-toolkit.ac.uk/routemap/archiving/downloads/archiving-joint-project-notes.pdf
Depositing in ORE

Recently Uploaded

Anisotropic diffusion of spherical particles in closely confining microchannels
Dettmer, SL; Pagliara, S; Misiunas, K; Keyser, UF (American Physical Society, 2014-06)
We present here the measurement of the diffusivity of spherical particles closely confined by narrow microchannels. Our experiments yield a two-dimensional map of the position-dependent diffusion coefficients parallel and ...

Local mechanical properties of electrospun fibers correlate to their internal nanostructure.
Camposeo, A; Greenfeld, I; Tantussi, F; Pagliara, S; Moffa, M; Fuso, F; Allegrini, M; Zussman, E; Pisignano, D (American Chemical Society, 2013-11-13)
The properties of polymeric nanofibers can be tailored and enhanced by properly managing the structure of the polymer molecules at the nanoscale. Although electrospun polymer fibers are increasingly exploited in many ...

EFL Teachers’ Beliefs and Attitudes towards English Language Assessment in a Saudi University’s English Language Institute
Mansory, Mazin (University of ExeterCollege of Social Sciences and International Studies, 2016-06-30)
State universities in Saudi Arabia have adopted a new educational policy, which made English the medium of instruction for all scientific departments. This has led to establishing a Foundation Year Programme (FYP) in the ...
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>dc:identifier:uri</td>
<td><a href="http://hdl.handle.net/10871/25521">http://hdl.handle.net/10871/25521</a></td>
<td></td>
</tr>
<tr>
<td>dc:language:iso</td>
<td>en</td>
<td>en_US</td>
</tr>
<tr>
<td>dc:publisher</td>
<td>University of Exeter</td>
<td>en_US</td>
</tr>
<tr>
<td>dc:subject</td>
<td>magnetic fields – MHD – methods: numerical – binaries: general – stars: formation</td>
<td></td>
</tr>
<tr>
<td>dc:title</td>
<td>The impact of non-ideal magnetohydrodynamics on binary star formation (Dataset)</td>
<td>en_US</td>
</tr>
<tr>
<td>dc:type</td>
<td>Dataset</td>
<td>en_US</td>
</tr>
</tbody>
</table>
Open access

Accepted for publication?
Ensure your work meets REF and other funders' open access requirements.

Deposit via Symplectic

+Funder requirements

Researchers must be aware of their funder's requirements before submitting their work to a publisher.

All researchers must comply with the Higher Education Funding Council for England (HEFCE) policy as well as their individual funder requirements.

Compare funder requirements >

+REF requirements

For your work to be eligible for the next REF you must deposit:
- In a repository (institutional, shared service or disciplinary).
- No later than three months after acceptance.
- Your final accepted version (following peer review)

Find out more >
Any Questions?

Contact us:
Greg Simpson, Research Data Officer. Open Access and Research Data Management Team
openaccess@exeter.ac.uk
rdm@exeter.ac.uk