Introduction to Data Management Plans and DMPonline

Sarah Jones
Digital Curation Centre, Glasgow
sarah.jones@glasgow.ac.uk
Twitter: @sjDCC

DMPonline workshop, 17 November 2015, EPFL, Lausanne
What is the Digital Curation Centre?

A UK service to support the Higher Education sector with Research Data Management (RDM)

“Helping to build capacity, capability and skills in data management and curation across the UK’s higher education research community.”

DCC Phase 3 Business Plan

Training | Events | Tools | Advocacy | Tailored Support | IJDC | International Conference

www.dcc.ac.uk
What is research data management?

“the active management and appraisal of data over the lifecycle of scholarly and scientific interest”

Data management is part of good research practice

Plan
Create
Document
Use
Publish
Share
What is a DMP and why write one?

What is in this for you?
What is a data management plan?

A brief plan written at the start of a project to define:

• how the data will be created?
• how it will be documented?
• who will access it?
• where it will be stored?
• who will back it up?
• whether (and how) it will be shared & preserved?

DMPs are often submitted as part of grant applications, but are useful whenever researchers are creating data.
How do DMPs help researchers?

Planning is useful whether or not a DMP needs submitted with a grant application, as it can help the researchers to...

1. Make informed decisions to anticipate and avoid problems
2. Avoid duplication, data loss and security breaches
3. Develop procedures early on for consistency
4. Ensure data are accurate, complete, reliable and secure
5. Save time and effort to make their lives easier!
Benefits of DMPs for institutions

- Opportunity to engage with researchers and improve RDM practice
- Raise awareness of support available
- Collate information to inform service delivery
- Ensure the University is not exposed to risk
- Ability to recover costs via grants
Which UK funders require a DMP?

<table>
<thead>
<tr>
<th>Research Funders</th>
<th>Policy Coverage</th>
<th>Policy Stipulations</th>
<th>Support Provided</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Published outputs</td>
<td>Data</td>
<td>Time limits</td>
</tr>
<tr>
<td>AHRC</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>BBSRC</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>CRUK</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>EPSRC</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>ESRC</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>MRC</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>NERC</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>STFC</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Wellcome Trust</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
</tbody>
</table>

n.b. Although the EPSRC doesn’t ask for a DMP in grant applications, it still expects one to be in place for every project

www.dcc.ac.uk/resources/policy-and-legal/overview-funders-data-policies
Some other funders that require DMPs or equivalent
What do research funders want?

• Usually a brief plan submitted in grant applications, and in the case of NERC, a more detailed plan once funded

• The Horizon 2020 pilot is different as DMPs are deliverables to be submitted during the project

• 1-4 sides of A4 as attachment or a section in application form

• Typically a prose statement covering suggested themes

• An outline of data management and sharing plans, justifying decisions and any limitations
Different stages of a DMP

Outline DMP
- Basic, short plan
- Answer a few questions
- Written at grant application stage

Full DMP
- Detailed plan
- Define procedures
- Written in conjunction with data centres

www.nerc.ac.uk/research/sites/data/dmp
DCC Checklist for a DMP

The DCC assessed existing funder requirements, DMP templates and other best practice to see what should be included in plans. This was synthesised down into common themes and questions.

- 13 questions on what’s asked across the board
- Prompts / pointers to help researchers get started
- Guidance on how to answer

Common themes in DMPs

1. Description of data to be collected / created
   (i.e. content, type, format, volume...)

2. Standards / methodologies for data collection & management

3. Ethics and Intellectual Property
   (highlight any restrictions on data sharing e.g. embargoes, confidentiality)

4. Plans for data sharing and access
   (i.e. how, when, to whom)

5. Strategy for long-term preservation
Ultimately funders expect:

- **timely release of data**
  - once patents are filed or on (acceptance for) publication
- **open data sharing**
  - minimal or no restrictions if possible
- **preservation of data**
  - typically 5-10+ years if of long-term value

See the RCUK Common Principles on data policy:
www.rcuk.ac.uk/research/Pages/DataPolicy.aspx

A DMP framework for PhD students courtesy of the Research360 project at the University of Bath

---

**Data Lifecycles & Data Management Plans**

1. What data will you produce?
2. How will you organise the data?
3. Can you/others understand the data?
4. What data will be deposited and where?
5. Who will be interested in re-using the data?
1. What data will you produce?

- What type of data will you produce?
- What types of file format?
- How easy is it to create or reproduce?
- Who owns the data and is responsible for it?
2. How will you look after the data?

- Is your data safe?
- Is your data organised?
- Can you find your data?
- Is the data properly stored and backed up?
- Is everyone who needs to, able to access the data?
3. How will you document the data?

- Do you still understand your older work?
- Is the file structure / naming understandable to others?
- Are you using standards?
- Which data will be kept?
- Which data can be discarded?
4. What data will be deposited and where?

- Are you *expected* or *allowed* to share your data?
- How will you share it?
- Define the core data set of the project
- Which data will be included in your publication / thesis?
5. Preservation and Re-use

- Which data do you need to keep?
- Who will the data be useful to?
- How long will your data be reusable for?
- Do you need to prepare your data for deposit in a long term archive?
DMPonline for writing DMPs

How the tool works and how it can help you

Image ‘tools’ CC-BY by zzpza www.flickr.com/photos/zzpza/3269784239
What is DMPonline?

A web-based tool to help researchers write data management plans

A short history

Launched in April 2010 at the Jisc conference

Released v.2 in March 2011 with extra functionality

Released v.3 in April 2012 with revisions in light of the DMPTool and work from the Jisc MRD programme

Released v.4 in Dec 2013, incorporating major changes from an evaluation and extensive user testing

Subsequent point releases in 2014-2015 to add new features
Main features in DMPonline

• Templates for different requirements (funder or institution)
• Tailored guidance (funder, institutional, discipline-specific etc)
• Ability to provide examples and suggested answers
• Supports multiple phases (e.g. pre-/ during / post-project)
• Granular read / write / share permissions
• Customised exports to a variety of formats
• Shibboleth authentication
How the tool works

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. Economic and Social Research Council
Otherwise leave blank. Not applicable/not listed.

To see institutional questions and/or guidance, select your organisation.
University of Glasgow
You may leave blank or select a different organisation to your own. Not applicable/not listed.

Tick to select any other sources of guidance you wish to see.
DCC guidance

Create plan

Click to write a generic DMP

Or choose your funder to get their specific template

Pick your uni to add local guidance and to get their template if no funder applies

Choose any additional optional guidance
Answering questions

Notes who has answered the question and when
Progress bar updates how many questions remain
Co-writing DMPs

Sections are locked for editing when they’re being worked on by colleagues.
Exporting plans

Choose what to export and pick from a variety of formats: pdf, txt, docx, csv, xml, html, json,
Institutions can customise DMPonline

Organisations can:
- Add templates
- Different types of questions
- Local guidance with links to support and services
- Example or suggested answers
- Add questions & guidance to funder templates
- Include their own logo and text in a banner
- ...
More information

Customising DMPonline
www.dcc.ac.uk/news/customising-dmponline-admin-interface-launches

Screencast on how to use DMPonline
http://www.screenr.com/PJHN

GitHub
Get the code, amend it, run a local instance, flag issues, request features...
https://github.com/DigitalCurationCentre/DMPonline_v4
Welcome.

DMPonline has been developed by the Digital Curation Centre to help you write data management plans.

Screencast on how to use DMPonline

Sign in

Email address *
Password *
Forgot your password?

Remember me
Sign in

Or sign in with your institutional credentials (UK users only)

Sign up

New to DMPonline? Sign up today
Exercise: trialling DMPonline to write your DMP

Build your own DMP
Exercise: using DMPonline

1. Go to the test site and register: [https://dmponline-test.dcc.ac.uk](https://dmponline-test.dcc.ac.uk)

2. Once logged in, create a new DMP (pick any funder or institution as you prefer)

3. Browse the guidance and start to write your plan

4. Share your plan with a colleague and add comments

5. Export the plan in your preferred format
Useful resources

Tips and guidance to takeaway
A useful framework to get started

Framework for Creating a Data Management Plan

This framework can be used as an outline in assembling data management plans to accompany grant applications. Note that some funders have page limits for data management plans—NSF limits plans to two pages.

Elements of a Data Management Plan

This list of elements is informed by a gap analysis that ICPSR conducted of existing recommendations for data management plans and other forms of guidance made available for researchers generating data. The result of the gap analysis was a comparison of existing forms of guidance. Elements that are highly recommended for inclusion in effective data management plans are noted.

See our bibliography for additional readings germane to the elements of a data management plan.

Data Description (Recommended)

Provide a brief description of the information to be gathered -- the nature, scope, and scale of the data that will be generated or collected.

Why this is important

A good description of the data to be collected will help reviewers understand the characteristics of the data, their relationship to existing data.

Example 1:
This project will produce public-use nationally representative survey data for the United States covering Americans' social backgrounds, enduring political predispositions, social and political values, perceptions and evaluations of groups and candidates, opinions on questions of public policy, and participation in political life.

Example 2:
This project will generate data designed to study the prevalence and correlates of DSM-III-R psychiatric disorders and patterns and correlates of non-treatment among a nationally representative sample of over 8000 respondents. The sensitive nature of these data will require that the data be released the

Think about why the questions are being asked – why is it useful to consider that?

Look at examples to help you understand what to write

www.icpsr.umich.edu/icpsrweb/content/datamanagement/dmp/framework.html
Tips for writing DMPs

• Seek advice - consult and collaborate

• Find out what help is available from IT support, library, data centres etc

• Base plans on available skills & support

• Consider good practice for your field

• Use standards where available

• Make sure implementation is feasible
Example DMP templates

DCC Checklist (generic template)
www.dcc.ac.uk/sites/default/files/documents/resource/DMP/DMP_Che...3.pdf

Bath postgraduate template
http://opus.bath.ac.uk/30772

DataTrain template
www.lib.cam.ac.uk/preservation/datatrain/documents.html

20 questions for RDM
https://datamanagementplanning.wordpress.com/2012/03/07/twenty-questions-for-research-data-management

MRC template
www.mrc.ac.uk/research/research-policy-ethics/data-sharing/data-management-plans
Example plans

108 DMPs from the National Endowment for the Humanities (USA)

20+ scientific DMPs submitted to the NSF (USA) provided by UCSD
http://libraries.ucsd.edu/services/data-curation/data-management/dmp-samples.html

2 Rural Economy & Land Use (RELU) programme examples (UK)
http://relu.data-archive.ac.uk/data-sharing/planning/examples

Further examples:
www.dcc.ac.uk/resources/data-management-plans/guidance-examples
DCC support on DMPs

- Webinars and training materials
- How-to guides and other advisory documents
- Checklist on what to cover in DMPs
- Example DMPs
- DMPonline

www.dcc.ac.uk/resources/data-management-plans
MANTRA online training for RDM

MANTRA is a free online course for those who manage digital data as part of their research project.

http://datalib.edina.ac.uk/mantra
Thanks for listening

DCC resources on DMPs:

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

Follow us on twitter:

@digitalcuration and #ukdcc