



because good research needs good data

Data Management Planning

Monica Duke
Digital Curation Centre
M.Duke@ukoln.ac.uk

Funded by:



This work is licensed under the Creative Commons Attribution-NonCommercial-ShareAlike 2.5 UK: Scotland License. To view a copy of this license, visit <http://creativecommons.org/licenses/by-nc-sa/2.5/scotland/>; or, (b) send a letter to Creative Commons, 543 Howard Street, 5th Floor, San Francisco, California, 94105, USA.

DCC East Midlands Roadshow, Loughborough 7-8 February, 2012 #dcc_lboro



D | C | C

because good research needs good data

Funders have DMP requirements

Research Funders	Policy Coverage		Policy Stipulations					Support Provided			
	Published outputs	Data	Time limits	Data plan	Access/sharing	Long-term curation	Monitoring	Guidance	Repository	Data centre	Costs
AHRC	●	●	●	●	●	◐	○	●	○	◐	○
BBSRC	●	●	●	●	●	●	●	●	●	◐	●
CRUK	●	●	●	●	●	●	●	◐	●	○	○
EPSRC	●	●	●	○	●	●	●	◐	○	○	●
ESRC	●	●	●	●	●	●	●	●	●	●	◐
MRC	●	●	●	●	●	●	○	◐	●	○	◐
NERC	●	●	●	●	●	●	●	●	●	●	◐
STFC	●	●	●	●	●	●	●	◐	●	◐	○
Wellcome Trust	●	●	●	●	●	●	●	●	●	◐	●

www.dcc.ac.uk/resources/policy-and-legal/overview-funders-data-policies



because good research needs good data

Why are data plans important?

As a funder, we have tried to encourage this by ensuring all researchers think about data at the application stage, and have a plan for sharing any data that will be generated.

We particularly welcome the DCC's work to develop data management planning tool – which we think could be a great aid to researchers and institutions in planning and resourcing data management and sharing as an integral part of the research process.

David Lynn, Head of Strategic Planning and Policy at the Wellcome Trust

<http://www.dcc.ac.uk/news/idcc11-preview-interview-david-lynn>



D | C | C

... because good research needs good data

Background and context

Liz Lyon (DCC Bath) 'Dealing With Data' (2007) consultancy report...

REC 9. Each funded research project should submit a structured Data Management Plan for peer-review as an integral part of the application for funding. (1, 2)

<http://www.ukoln.ac.uk/ukoln/staff/e.j.lyon/publications.html#2007-06-19>



Dealing with Data: Roles, Rights, Responsibilities and Relationships Consultancy Report

Document details

Author:	Dr Liz Lyon, UKOLN, University of Bath
Date:	19 th June 2007
Version:	V1.0
Document Name:	data-consultancy-report-final.doc
Notes:	



D | C | C

... because good research needs good data

Requirements: 5 common themes

1. Describe **the data** (i.e. content, type, format & volume)
2. **Standards & methodologies for data collection & management** (including metadata & documentation)
3. **Plans for data sharing and access** (i.e. how, when, to whom)
4. **Outline ethical & Intellectual Property (IP) concerns** (highlight any restrictions on data sharing e.g. embargoes, confidentiality)
5. **Plans for archiving & long-term preservation**

Funders all ask for different things

<http://www.dcc.ac.uk/resources/data-management-plans/funders-requirements>

DCC Data Management Roadshow, Cambridge, 11 November 2011



because good research needs good data

A Digital Curation Centre
'working level' guide



How to Develop a Data Management and Sharing Plan

Sarah Jones (DCC)

Practical Help

- Show awareness
- Comply with requirements
- Use existing support
- Justify decisions
- Plan for implementation
- Think in terms of the themes

www.dcc.ac.uk/resources/how-guides/develop-data-plan



Digital Curation Centre, 2011.
Licensed under Creative Commons Attribution 2.5 Scotland:
<http://creativecommons.org/licenses/by/2.5/scotland/>



| D | C | C

... because good research needs good data

ICPSR Framework

ICPSR | INTER-UNIVERSITY CONSORTIUM FOR
POLITICAL AND SOCIAL RESEARCH

Deposit Data & Findings

How to Deposit

Data Deposit Form

Data Preparation Guide

Data Management Plans

Preserving Respondent
Privacy/Confidentiality

Suggest Data to Archive

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, and any disclosure risks that may apply.

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 service utilization for these disorders in a nationally representative sample of over 8000 respondents. The sensitive nature of these data will require that the data be released through a restricted use contract.

[+ Hide example\(s\)](#)

Access and Sharing (Recommended)

Indicate how you intend to archive and share your data and why you have chosen that particular option. Possible mechanisms for archiving and sharing include:

- Domain repository like ICPSR (social science)
- Self-dissemination through a dedicated Web site that the research team will create and maintain. *If this option is chosen, it is recommended that the data producer arrange for eventual archiving of the data after the self-dissemination period terminates and specify the schedule for data sharing in the grant application.*
- Preservation with delayed dissemination. Under such an agreement the data producer makes an arrangement with a public data repository for archival preservation of the data with dissemination to occur at a later date, usually within a year.

- Why is this important
- Examples

Digital video data files generated will be processed and submitted to the [repository] in MPEG-4 (.mp4) format.

Staff time has been allocated in the proposed budget to cover the costs of preparing data and documentation for archiving. The [repository] has estimated their additional cost to archive the data is [insert amount]. This fee appears in the budget for this application as well.

<http://www.icpsr.umich.edu/icpsrweb/ICPSR/dmp/index.jsp>

DCC Data Management Roadshow, Cambridge, 11 November 2011



because good research needs good data

Examples

UCSD (NSF grants)

b

ICPSR links


<https://wiki.ucar.edu/display/cgdccr/CCSM+Data+Management+Plan>

www.glasgow.ac.uk/datamanagement

The screenshot shows the University of Glasgow website. At the top, there is a navigation bar with the University of Glasgow logo and a search bar. Below this is a header for 'Data management support for researchers'. The main content area is titled 'Data planning' and includes several links: 'What is data planning?', 'What do research funders expect?', 'What should I cover in my data plan?', and 'Who can help with data planning at the University of Glasgow?'. There is a section titled 'What is data planning?' which describes it as a process of making decisions at the outset of research. It lists three key considerations: which software to use, how to organise, store and manage data, and what to include in consent agreements. It also mentions that these factors affect what can be done with data in the future. Another section, 'What do research funders expect?', states that most funders expect a short statement outlining plans for data management and sharing, and lists five themes: 1. What data will be created? 2. How will the data be documented and described? 3. How you will manage ethics and IPR? 4. What are the plans for data sharing and access? 5. What is the strategy for long-term preservation and sustainability? The page concludes by stating that the DCC has collated specific details of UK research funders' data management plan requirements and provides a link to the 'Data Management Plans' page. A 'Back to the top' link is also present.

DCC guidance

Accessibility | Glossary | Sitemap | RSS feeds

 because good research needs good data

Search

Home | Digital Curation | About Us | News | Events | Resources | Training | Projects | Community | Contact Us

[Home](#) > [Resources for digital curators](#) > [Data management plans](#) > [Data plan guidance and examples](#)

Data plan guidance and examples

In this section

[Curation Reference Manual](#)

[Curation Lifecycle Model](#)

[Policy and legal](#)

[Data management plans](#)

[DMP Online](#)

[Funders' requirements](#)

[Guidance and examples](#)

[Case studies](#)

[Tools and applications](#)

[Briefing Papers](#)

[How-to Guides](#)

[Standards](#)

[Publications](#)

[External resources](#)

[Roles](#)

[Curation journals](#)

[Informatics research](#)

Lots of guidance and examples are available to help with data plans.

A summary of some useful resources is provided below and practical guidance is available in the DCC guide:

[How to develop a data management and sharing plan](#)

[ARL Resources for Data Management Planning](#)

Lists compiled by the US Association of Research Libraries covering data planning guidance webpages, data repositories and tools

[ANU data management planning website](#) -

Resources by the Australian National University, including a template plan, course slides and the very comprehensive [data management manual](#)

[ESRC Data Management Plan: Guidance for Peer-Reviewers](#)

An explanation of what the ESRC is looking for in plans

[MIT Libraries data planning checklist](#)

A short list of questions with links to useful clarifications and guidance

[ICPSR Guidelines for effective data management plans](#)

Suggestions of what elements to include in data management plans and a useful list of example plans and guidance resources

DMP Online



The beta version of the DMP Online tool launched in March 2010 and was showcased at the JISC Conference in April. The DCC's data management planning tool draws on funders' data-related requirements to help researchers create, develop and export customisable data management plans at the grant application stage and throughout the project's lifetime.

[Read more](#)



because good research needs good data

What is DMP Online?

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

- Create, store and update multiple versions of Data Management Plans at the application and in-project stages
- Meet funders' specific data-related requirements*
- Get funder- and institution-specific guidance on best practice and helpful contacts
- Customise and export DMPs in a variety of formats
- A mapped checklist underlies the tool

<http://dmponline.dcc.ac.uk/>

* Disclaimer: mappings are not yet endorsed by funders



Type of Funder

e.g. UK Research Council



Select Funder

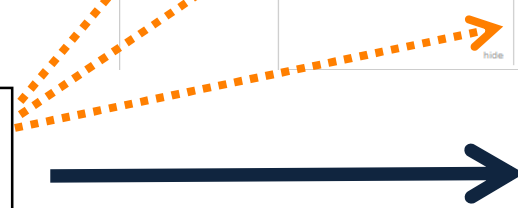
e.g. BBSRC, ESRC, MRC



Funder question mapped to checklist

3.1 Data volume and data type, e.g. qualitative or quantitative data	2.1 Give a short description of the data being generated or reused in this research	<input type="text"/>
ADD ANOTHER QUESTION		
3.2 Data quality, formats, standards documentation and metadata	2.3.3 Which file formats will you use, and why?	<input type="text"/>
	2.3.4 What criteria will you use for Quality Assurance/Management?	<input type="text"/>

Here you should outline and justify your choice of format, e.g. Microsoft Excel for recording measurements or SPSS for analysis, as these are in widespread use; the University has the relevant software licences or they're accepted standards in your field, etc. Decisions relating to file formats may also be made with recourse to staff expertise, a preference for Open formats, accepted standards, or widespread usage with a given community.
Guidance: - UKDA guidance on Formatting Your Data



Pointers to answers

...

10 sections of the checklist

- Introduction and requirements
- Data types, Formats, Standards and Capture Methods
- Ethics and Intellectual Property
- Access, Data Sharing and re-use
- Short-term storage and data management
- Deposit and long-term preservation
- Resourcing
- Adherence and review
- Agreement/ratification by stakeholders
- Annexes



because good research needs good data

Demo.....

A screenshot of the DMPonline website. The page title is 'New plan [MRC]'. The user is logged in as 'martin2'. The form contains the following fields: 'Name of project' (My project), 'Currency' (£), 'Budget' (1000), 'Starting date' (2011-03-24), 'Ending date' (2011-03-31), 'Lead organisation' (empty), and 'Other partners' (empty). Below the form is a question 'Has your application been funded?' with radio buttons for 'Yes' and 'No'. A 'Next' button is at the bottom. A large teal watermark 'www.screenr.com/Syo' is overlaid on the bottom half of the screenshot. The screenr player interface at the bottom shows a progress bar, a timestamp of 1:52 / 4:08, and the screenr logo.

DMPOnline at the JISC MRD workshop, 1-2 Dec 2011, #jiscmrd



because good research needs good data

Further resources for DMPOnline

Main DCC DMP page (start here)

<http://www.dcc.ac.uk/dmponline>

DMP Online tool (includes instructions, description, future plans)

<https://dmponline.dcc.ac.uk/>

Supporting documents (e.g. most up to date Checklist, worked examples)

<https://dmponline.dcc.ac.uk/documents>

Funders' requirements, Screencast demo, DMP overview presentation (DCC Cambridge Roadshow, November 2011)

DMP overview presentation with a different exercise (SeIUCCR Summer School, Abingdon, September 2011)

http://www.ngs.ac.uk/sites/default/files/file/Seiuccr%20Presentations/MD-abingdon-slides.ppt_0.pdf

DMP tutorial from IASSIST 2011, Vancouver

<http://www.dcc.ac.uk/events/workshops/iassist-dmp-online-workshop>



because good research needs good data

Exercise: 4 case studies

Exercise 2: How to support researchers with data management planning

Read the short case studies and discuss the merits of each approach:

- Which approach do you favour?
- What may be feasible to implement at your institution?
- Are there other types of support you would consider providing?



D | C | C

because good research needs good data

Acknowledgments....

Sarah Jones (DCC)

Martin Donnelly (DCC)

martin.donnelly@ed.ac.uk

info@dcc.ac.uk

THANK YOU