



Citation for published version:

Ball, A, Donnelly, M & Jones, S 2011, 'Research Data Management with the Digital Curation Centre', 22nd Machine Evaluation Workshop, Adelphi Hotel, Liverpool, 29/11/11 - 30/11/11.

Publication date:
2011

[Link to publication](#)

University of Bath

Alternative formats

If you require this document in an alternative format, please contact:
openaccess@bath.ac.uk

General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

Take down policy

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

Research Data Management with the Digital Curation Centre

Alex Ball

Martin Donnelly

Sarah Jones

30 November 2011

Contents

1	Who we are	1
2	What we do	2
3	Data Management Plans	3
4	Genesis of DMP Online	4
5	Demonstration of DMP Online	6

I Who we are

The Digital Curation Centre (DCC) is a collaboration between

- University of Edinburgh
- HATII, University of Glasgow
- UKOLN, University of Bath

Key facts

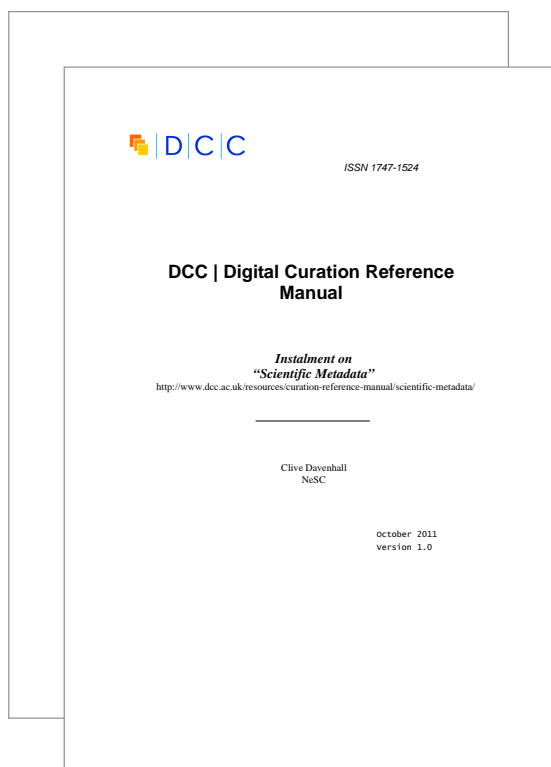
- Funded by JISC
- Started in March 2004
- Hub of expertise in curating digital research data
- Observe, reach out, innovate, support JISC

We identify and collect best practice, and disseminate out to the community through peer-to-peer discussions, training events and guidance materials.

2 What we do

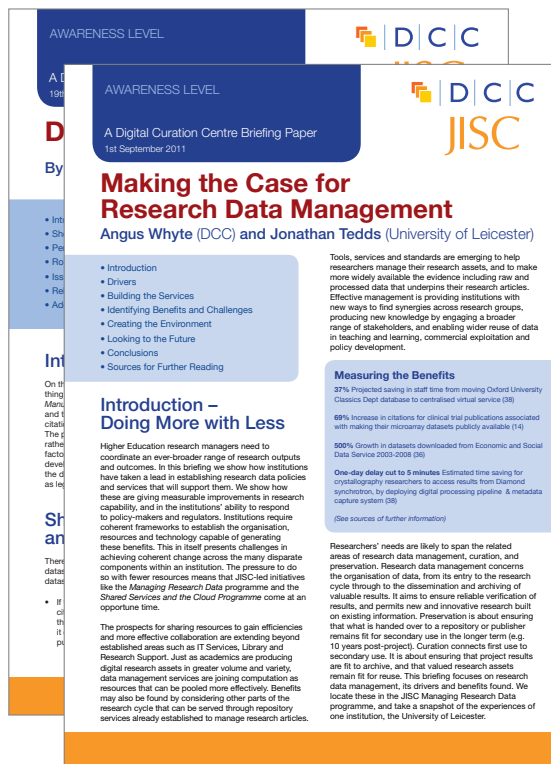
Curation Reference Manual

- Ongoing project to build a comprehensive textbook for digital curation
- Advice, in-depth information and criticism of current techniques and best practice
- Advanced level
- Chapters include:
 - Appraisal and selection
 - File formats
 - Preservation metadata
 - Preservation strategies
 - Scientific metadata



Briefing papers

- Concise overviews of digital curation topics, aimed at managers and absolute beginners
- Awareness level
- Papers include:
 - Making the case for research data management
 - Appraisal and selection
 - Curating eScience data
 - Data citation and linking
 - Data protection
 - Freedom of Information
 - ... and many more

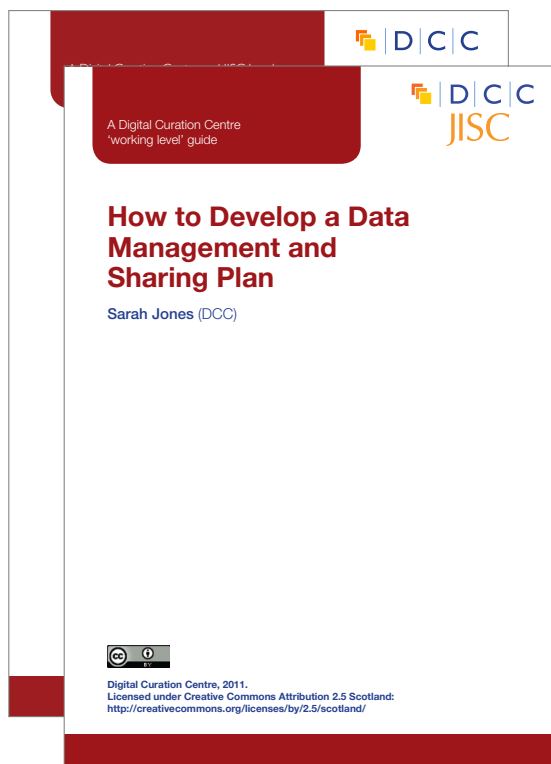


We've been writing briefing papers since the beginning, but a whole load more are on the way now as a result of our involvement with providing support to JISC's Managing Research Data programme.

How-to guides

- Practical introductions for people embarking on unfamiliar digital curation tasks
- Working level
- Guides available:
 - How to appraise and select research data
 - How to cite datasets and link to publications
 - How to develop a data management and sharing plan
 - How to license research data

The third one on the list is about developing a Data Management Plan, and is thus pertinent to our discussions today.

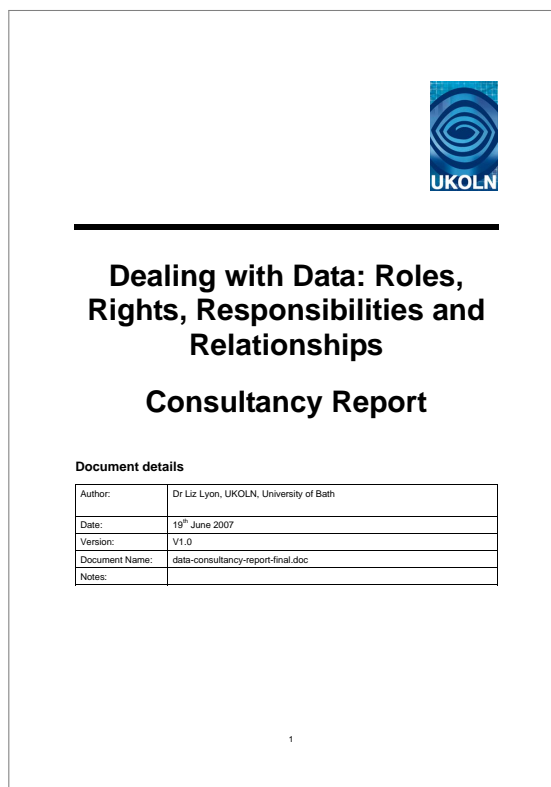


3 Data Management Plans

- 6 **Recommendation 9.** Each funded research project, should submit a structured Data Management Plan for peer-review as an integral part of the application for funding. — Liz Lyon (2007), *Dealing with Data: Roles, Rights, Responsibilities and Relationships* (University of Bath)

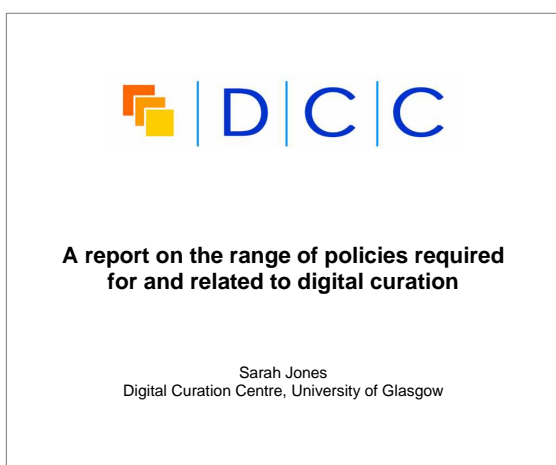
Why? Writing and using a Data Management Plan helps

- to co-ordinate the actions of data stakeholders
- to ensure all necessary tasks are accomplished
- to ensure data are properly curated
- with releasing data in a timely fashion
- with sharing data as openly as possible
- with preserving data for future use



4 Genesis of DMP Online

1. Monitor and analyse data-related policies of major funding bodies



Sarah Jones (2009), *A Report on the Range of Policies Required for and Related to Digital Curation* (Glasgow: DCC)

Funders' data plan requirements

Below is an abstract of UK funders' requirements for data plans:

[Summary of UK research funders' expectations for the content of data management and sharing plans \(PDF\)](#)

The DCC also provides full details on [funders' data policies](#).

Funders' data policies

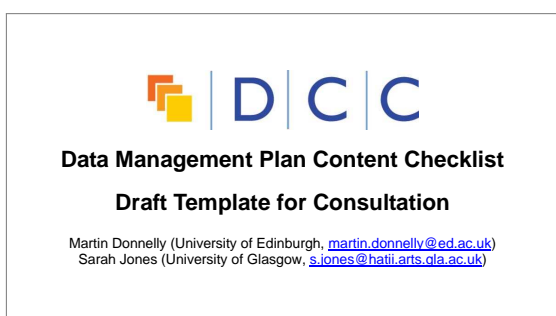
Summaries of UK research funders' data policies are available through the menu on the left.

The summaries provide links to the funder's full policy, note core stipulations and available support e.g. guidance and data centres.

Details are available for:

- [AHRC](#)
- [BBSRC](#)

2. Synthesise funders' requirements for research data management planning



Martin Donnelly and Sarah Jones (2009), *Data Management Plan Content Checklist* version 1.0 (Edinburgh and Glasgow: DCC)



Version 1.0 of the checklist had 51 questions and headings. After a period of consultation, version 2.0 was released, with 115 questions and headings.

3. Map from DCC Checklist to funder requirements

At this stage we were happy we had the checklist about right, so we mapped the DCC questions back to the funder questions. These, of course, had changed in the meantime, so getting the mappings right meant some further tweaking. The result was version 3.0 of the checklist with 118 questions and headings.

How the mapping process works in DMP Online

(N.B. Numbers given are illustrative only, and will change when v3.0 of the Checklist goes live in October 2010.)

<p>Different research funders ask different data-related questions at the funding application stage. The DCC has produced a summary of these requirements, and mapped the funders' questions to its own comprehensive Data Management Planning Checklist.</p> <p>The full DCC Checklist comprises 88 questions (and 27 headings, which may also have guidance attached to them: the system allows funder-specific guidance for both questions and headings). Few users are expected to complete the Checklist in its entirety, whether at the application stage or during the research project's lifetime.</p> <p>In the DMP Online tool, applicants are presented with the funder and DCC questions side-by-side, together with guidance on how to answer each DCC question. The user answers the DCC questions, and by doing so meets the funder's requirements <i>de facto</i>. (N.B. The mappings are not yet endorsed by the funders.)</p> <p>We are working with funders (notably the ESRC/UKDA) to endorse our mappings, and to figure out a way by which DMP Online output will be able to feed directly into the Joint Electronic Submission (Je-S) process used by the seven RCUK councils.</p>		
<p>The Three Stages of a Data Management Plan: Minimal, Core and Full</p>		
<p>APPLICATION STAGE (minimal)</p>	<p>FUNDING AWARDED (core)</p>	<p>LATE-PROJECT / POST-PROJECT (full)</p>
<p>The AHRC asks 11 data-related questions (and 4 non-data-related questions) which map to 23 DCC questions</p> <p>The BBSRC asks 10 data-related questions which map to 15 DCC questions</p> <p>The ESRC asks 5 data-related questions which map to 14 DCC questions</p> <p>The MRC asks 9 data-related questions which map to 24 DCC questions</p> <p>The Wellcome Trust asks 5 data-related questions which map to 14 DCC questions</p> <p>Other funders (e.g. NERC, EPSRC, STFC) do not have explicit (council-wide) requirements at the application stage: applicants to these councils are presented with a 'superset' of the 37 DCC questions that the other councils' requirements map to. (N.B. Users are free to remove irrelevant questions, or add extra ones, as they wish.)</p>	<p>Once funding is in place, all users are presented with the same 'core' set of 54 DCC questions, together with the relevant guidance.</p> <p>This approach facilitates cross-Council comparison of projects during their lifetime, and encourages researchers to engage with data management planning on a deeper level than the bare minimum required by the funders.</p> <p>Again, users are free to add and remove questions as they wish.</p>	<p>There are a further 34 questions covering longer-term aspects that users may wish to include in their data management plans towards the end of the research lifecycle, and after the project has ended. Reaching this stage therefore means the user will have answered between 55 and 88 questions in all.</p> <p>This stage differs from the previous one in that users have to 'opt in' to include these additional questions.</p> <p>Users may also append Annexes covering issues more detailed than the tool's scope covers.</p>

That is, you'll appreciate, rather a lot of questions. And while it would be useful for researchers to answer most or all of them when embarking on their project, it's much more than is needed at the grant application stage, and probably too much to deal with in one go. So what we felt was needed was a software tool that could present researchers with only those questions they needed to answer at the application stage, but which would let them come back and fill out the rest once the project was funded. So that is how we came to develop DMP Online.

4. DMP Online



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

DMP Online has four functions:

1. Create, store and update Data Management Plans
2. Meet funders' specific data-related requirements — CAUTION: awaiting validation
3. Provide specific guidance from funders and institutions
4. Export Data Management Plans in various formats

5 Demonstration of DMP Online



We begin by logging in. . .

. . . and selecting 'Start a new plan'.

The tool asks us at what stage we are (application stage or post-award). . .

. . . and the type of funder (RCUK, non-RCUK, US, etc.).

Finally, we choose the actual funder from a list.

Now we can start filling in the details of our project. Name, budget, etc.

We now get a summary screen showing several plan sections, how many questions are in each and how many we've filled in. We click on 'Edit' to start answering the questions.

In each section we see the funders original requirement on the left, how it breaks down to DCC questions in the middle, and we fill in the boxes on the right.

As we type, help text pops up on the far right. This could be a single line or several paragraphs, and may contain links to external resources.

We can save the plan and return to it later. Once we're happy with it, we can export it. There's a choice of simple or advanced export. The difference is that the advanced export lets us filter out questions and change the order in which they appear.

Either way, there's a choice of output formats and presentation styles.

Alex Ball. DCC/UKOLN, University of Bath. <http://www.ukoln.ac.uk/ukoln/staff/a.ball/>



Except where otherwise stated, this work is licensed under Creative Commons Attribution 2.5 Scotland: <http://creativecommons.org/licenses/by/2.5/scotland/>



The DCC is funded by JISC.

For more information, please visit <http://www.dcc.ac.uk/>