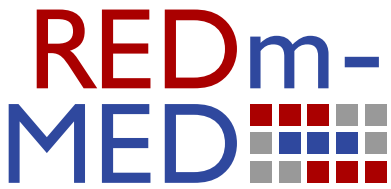

DMP Online Template for the Department of Mechanical Engineering at the University of Bath

Alex Ball, Mansur Darlington

redm5too120504ab10.pdf

Issue date: 25th May 2012

Catalogue Entry

Title	DMP Online Template for the Department of Mechanical Engineering at the University of Bath
Creator	Alex Ball (author)
Contributor	Mansur Darlington
Subject	research data management; data management plans
Description	This document provides a template for data management plans suitable for use with the DCC's DMP Online tool. It is based on the specification contained within *A Research Data Management Plan for the Department of Mechanical Engineering, University of Bath* for the contents of Project Data Management Plans produced within the Department.
Publisher	University of Bath
Date	4th May 2012 (creation)
Version	1.0
Type	Text
Format	Portable Document Format version 1.5
Resource Identifier	redm5too120504ab10
Language	English
Rights	© 2012 University of Bath

Citation Guidelines

Alex Ball, Mansur Darlington. (2012). *DMP Online Template for the Department of Mechanical Engineering at the University of Bath* (version 1.0). REDm-MED Project Document redm5too120504ab10. Bath, UK: University of Bath.

CONTENTS

1 INTRODUCTION

This document provides a template for data management plans suitable for use with the DCC's DMP Online tool.[1] It is based on the specification contained within *A Research Data Management Plan for the Department of Mechanical Engineering, University of Bath* [DB12] for the contents of Project Data Management Plans produced within the Department.

2 DATA MANAGEMENT PLAN TEMPLATE

2.1 Summary of Research Activity

Q 1.1.1 or DMP metadata. *Project name*

[No guidance.]

DMP metadata. *Start date*

[No guidance.]

DMP metadata. *End date*

[No guidance.]

Q 1.1.2. *Funding body (or bodies)*

[No guidance.]

Q 1.1.5 or DMP metadata. *Lead partner organisation*

[No guidance.]

Q 1.1.6 or DMP metadata. *Other partner organisation*

[No guidance.]

Q 4.1.1. *Are you under obligation or do you have plans to share all or part of the data you create/capture?*

Answer 'yes' if there are no impediments to making the data publicly available. Otherwise answer 'no'.

Q 4.1.2. *If you answered No to DCC 4.1.1, why will you not share your data?*

Access restrictions should be stated here in the simplest, most concise form (e.g. Dataset X may only be viewed by Y and Z); where important complexities are glossed, prompts to consult the full documentation should be given.

1. DMP Online, URL: http://dmponline.dcc.ac.uk/

Q 3.2.1. *Will the dataset(s) be covered by copyright or the Database Right? If so give details in DCC 3.2.2, below.*

[Default guidance.] Guidance:

- DCC Legal Watch Paper on the Database Right

Q 3.2.2. *If you answered Yes to DCC 3.2.1, Who owns the copyright and other Intellectual Property?*

Ownership of information should be stated here in the simplest, most concise form; where important complexities are glossed, prompts to consult the full documentation should be given.

Q 6.1. *What is the long-term strategy for maintaining, curating and archiving the data?*

State where the data will be kept in the long term. Unless a suitable disciplinary data archive is identified, this will probably be the University of Bath X Drive.

Q 1.3.1. *Funding body requirements relating to the creation of a data management plan*

Provide a link to the RCUK Policy and Code of Conduct on the Governance of Good Research Conduct. For EPSRC-funded research, cite the EPSRC Expectations of Funded Institutions. Consider using a service like WebCite to provide a snapshot from the time of writing.

Q 1.3.2. *Institutional or research group guidelines*

Provide links to:

- the University of Bath Code of Good Practice in Research;
- the most recent version of the Research Data Management Plan for the Department of Mechanical Engineering, University of Bath (e.g. `http://opus.bath.ac.uk/30099/`);
- where the locations of project documentation (project plan, confidentiality agreements, consent forms, etc.) can be found, e.g. the Departmental Research Data Management wiki or, if non-confidential, the Project Record Manifest.

2.2 Existing Data Re-use

Q 2.2.1. *Have you reviewed existing data, in your own institution and from third parties, to confirm that new data creation is necessary?*

In the interests of economy, consideration should be given, if possible during formulation of the project proposal, to whether the research activity's data requirements could be met in whole or in part by existing data.

A Research Activity can either mine existing data for new results, add to an existing body of data (to fine tune, generalize or place limits on previous results), or create an entirely new body of data. When planning a new Research Activity, researchers should be able to justify taking one of these three approaches.

Q 2.2.2. *What existing datasets could you use or build upon?*

[Default guidance.] If none, enter 'n/a'

Q 2.2.3. *Describe any access issues pertaining to the pertinent, existing data*

Typical access issues might include access that is contingent on successful application; unclear data licensing; and cost.

Q 2.3.1. *Why do you need to capture/create new data?*

Typical reasons for not re-using data include conducting a search and finding that there has been no similar previous research; conducting research on an object that has not previously been studied; operating in an area where all research is covered by strict confidentiality agreements. Typical reasons for generating new data might include performing a comparison over time; extending existing research to cover new areas.

2.3 Relating New Data to Existing Data

Q 2.4.1. *What is the relationship between the new dataset(s) and existing data?*

[Default guidance.] This is concerned less with existing data that may be used in the Research Activity, but rather with the disciplinary context. A typical answer might identify a body of data with which it would be helpful to harmonise newly generated data, or from which methodologies might be drawn, e.g. ISO standard materials testing data, time/motion studies data.

Q 2.4.2. *How will you manage integration between the data being gathered in the project and pre-existing data sources?*

State the measures that will be/have been taken to ensure integrability between newly generated data and existing data. The following are possible issues to consider. Only very brief answers are required here: full details should be given in corresponding sections later in the DMP:

- Method of assuring data quality.
- Method of recording provenance.
- Mechanisms for ensuring trustworthiness of data.
- Choice of standard formats, ontologies, conventions, etc. for the data.
- Choice of standard formats, ontologies, conventions, etc. for the metadata.

A typical way in which data are re-used is in combination with similar data. This is considerably easier if compatibility issues are addressed in the planning stages of a Research Activity.

2.4 Future Use of the Data

Q 4.3.1. *Which groups or organisations are likely to be interested in the data that you will create/capture?*

If the future uses for research data are known or can be predicted at the outset, special provisions can be made during the research that increase the compatibility of the data with that future use. Explicitly stating where this has been done can help Data Librarians/Managers continue this work in the preservation stage.

To assist future use, list any bodies/groups which might be interested in the data, and the foreseeable contemporary or future uses to which they might put the data.

It is acceptable to define groups based on discipline, research interest or specific research topic. It is acceptable to list bodies or groups without reference to uses, and foreseeable uses without reference to specific groups, if appropriate."

Q 4.3.2. *How do you anticipate your new data being reused?*

State the measures that will be/have been taken to prepare the data for these bodies/ groups/uses.

The following are possible issues to consider. Only brief answers are required here: full details should be given in corresponding sections later in the DMP.

- Forms of data organization.
- Choice of standard formats, ontologies, conventions, etc. for the data.
- Choice of standard formats, ontologies, conventions, etc. for the metadata.

2.5 The Project Record Manifest

Q 10.3. *Other annexes as required*

Provide here the location of the Project Record Manifest (PRM), which may be a wiki page or a standalone document. This is the document that records what Data Records are included in a Data Case, how they came about and what relationships exist between them. It helps future researchers to understand the data, assess their suitability and re-use them for new research; it also satisfies some users' requirements for provenance information.

The PRM should be written according to the Departmental template, and ideally the Data Case it documents should be represented using a Research Activity Information Development (RAID) diagram. If it has not been possible to use the RAID association method, an acceptable alternative is to present an annotated list of Data Records showing associative information. If the latter method is used, specify here the procedure for keeping the list up to date.

2.6 Data Generation, Manipulation and Organization

Q 2.3.2. *Describe the process by which you will capture/create new data*

Give a detailed account of how the data will be/have been generated and manipulated, including the methods, technology, conventions, coding schemes, etc. that were used. For the sake of future re-usability and re-purposing, the use of robust, generic or standard data generation and manipulation tools is preferred.

When writing a DMP in retrospect, it is acceptable to cite a journal/conference paper containing the information, provided it is detailed enough, and that a pre- or post-print is available in case of access difficulties. In the normal course of events, the information should be provided here first and then adapted for use in a journal/conference paper. It may be helpful to provide this information in the form of a commentary on a RAID diagram.

Describe how the data will be/have been organized. This refers equally to how data are organized within Data Records, how Data Records are organized within the Data Case, and how project management records are organized. Providing this information makes it easier for others involved in the later management of research data (for example, by a project or Data Manager) or re-use of the data to navigate the Data Records and find specific parts. In addition, such information will aid in the efficient response to FOI requests. It can also help you to check that all the Data Records have been included.

The basic method recommended for organization of project documents (both within the Data Case and project management information) is by using a file naming convention such as that recommended by the Department.

Q 2.5.3. *How will you create or capture these metadata?*

It is strongly encouraged for management of Departmental research data and associated records, that full use is made of the metadata recording facilities provided by the electronic records-handling methods that are used. An example of this is the 'properties box' provided for individual files in all Microsoft Office applications in which can be recorded such things as document title, author(s), and so on.

Where possible, additional metadata description of the research data or Research Data Records should be provided. It is recommended at present that the DataCite metadata schema is adopted for this.

Q 5.1.3. *How will you transfer/transmit the data, if this is required?*

All research data and other data and document records associated with a research activity should be assembled into a Data Case at completion of the research activity for ease of long-term curation and management. The Data Case should be packaged in a format that will promote its long-term potential for re-use. In particular, where special arrangements for archiving are mandated (e.g. by the funder) the packaging should follow the guidance provided which will ensure it is acceptable to the Data Librarian(s) taking custody of it. In the absence of such mandates, the BagIt packaging format is recommended: at their simplest such packages consist of a directory containing a file identifying the BagIt version, another listing the files in the package along with checksums, and a data directory containing the files and folders themselves. It is the responsibility of the PI that this is done and that all data is collected from independent file storage (independent and removable media) and from the personal (H-drive) storage space of any researcher involved in the project. At this juncture the data case records should be associated by bringing the Project Record Manifest up to date.

2.7 Data Quality

Q 2.3.4. *What criteria will you use for Quality Assurance/Management?*

The quality assurance procedures and standards should be recorded here that will be/have been used for collection, generation and manipulation of the research data. If any data quality issues were encountered, list them and describe what was done to resolve them. Quality management mechanisms may include: documentation, calibration, validation, monitoring, transcription metadata, peer-review.

2.8 Data Structures and Formats

Q 2.3.3. *Which file formats will you use, and why?*

At the project planning stage, the hardware and software environment in which the Research Activity will be conducted should be specified to the extent that it can be known. Where possible use off-the-shelf software and software that is supported by BUCS or a more local IT support service.

Indicate the formats to be used, and explain why these have been selected for use. If the choice of formats has been justified elsewhere in the DMP, readers may be directed to those sections in place of a recapitulation here.

Also indicate at this stage how this section will be completed during the course of the Research Activity.

Once Data Records have been made, start by specifying the hardware and software environment in which the data were generated or manipulated, and then consider alternative environments, tools and libraries that might support the data. If specialist tools were used, consider installing them on a virtual machine; in which case, provide details here of how to run the virtual machine.

Specify the information, tools or resources that would be needed to manipulate or make your Data Records human readable. If available/known, cite here format specification documents for all data formats used.

2.9 Data Semantics

Q 2.5.4. *What form will the metadata take?*

Data cannot be re-used if their meaning is not properly understood.

At the project planning stage, provide if possible a general statement about the conventions that will be used to allow interpretation of data (such as data dictionaries, coding schemes, ontologies, and so on). Once Data Records have been made, confirm or update the list as necessary. Indicate how this answer will be updated during the course of the Research Activity.

Q 2.5.5. *Why have you chosen particular standards and approaches for metadata and contextual documentation?*

[Default guidance.] Decisions relating to metadata standards may be made with recourse to: staff expertise, a preference for Open standards, or widespread usage with a given community.

Guidance:

- DCC Briefing Paper on Metadata Standards

Q 2.5.1. *Are the datasets which you will be capturing/creating self- explanatory, or understandable in isolation?*

[Default guidance.] You may wish to consider this from the perspective of a typical reader of a journal for your discipline.

Q 2.5.2. *If you answered No to DCC 2.5.1, what contextual details are needed to make the data you capture or collect meaningful?*

At the project planning stage, it is acceptable to say 'Unknown' here. Once Data Records have been made, provide any additional information that would be needed by an interested reader to understand the Data Records. As an example, tabular data can have terse column headings; fuller explanations of what a column represents can be given here. The information can be given directly in the DMP, or instructions can be given on how to look up the information for each Data Record. An example of the latter could be a direction to consult the RAID diagram for the project, since the RAIDmap Application may be used to associate explanatory documents (Context Data Records) with Data Records.

REFERENCES

[DB12] M Darlington & A Ball (2012-05-23). *A Research Data Management Plan for the Department of Mechanical Engineering, University of Bath.* redm1rep120207mjd12. REDm-MED Project Document. Version 1.2. University of Bath: Bath, UK. URL: http://opus.bath.ac.uk/30099/.