



Citation for published version:

Ball, A 2013, 'An Engineering Perspective on Good Practice', Supporting Data Management Planning, Loughborough, UK United Kingdom, 24/04/13.

Publication date:
2013

Document Version
Early version, also known as pre-print

[Link to publication](#)

Publisher Rights
CC BY

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.



D | C | C

because good research needs good data

An Engineering Perspective on Good Practice

Alex Ball

DCC/UKOLN, University of Bath

24 April 2013

Loughborough University



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

Funded by **JISC**



[Basic Search](#) | [Advanced Search](#) | [Last Search Results](#) | [Previous Searches](#) | [Edit Search](#)

Find results...

... with **all** of the words:

Perx Search

... with the **exact phrase**:... with **at least one** of the words:... **without** the words:

In Collections...

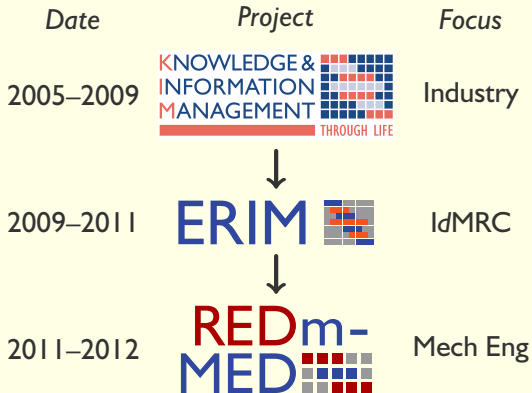
[View Category List](#) - View Alphabetical List of Collections[Check All](#) - [Clear All](#)

- [Australasian Digital Thesis \(ADT\)](#) [description website](#)
- [Aeronautical Research Council Technical Reports](#) [description website](#)
- [Australasian Research Repositories Online to the World \(ARROW\)](#) [description website](#)
- [arXiv - ePrint Archive](#) [description website](#)
- [Caltech Earthquake Engineering Research Laboratory Technical Reports](#) [description website](#)
- [CISTI - Canada Institute for Scientific and Technical Information](#) [description website](#)
- [Copac - Library Catalogues](#) [description website](#)
- [CSA Discovery Guides](#) [description website](#)
- [DIVA - Academic Archive On-Line](#) [description website](#)

EPSRC Expectations

1. **Research organisations (ROs)** to raise awareness of data sharing responsibilities and issues.
2. Publications should link to underlying data.
3. **ROs** must keep track of their research datasets and requests for them.
4. Born-analogue data must also be shareable on request.
5. **ROs** must provide open, online catalogues of their data; digital data must be given a robust ID.
6. Access restrictions should be clear and justified.
7. **ROs** must provide access to data for 10 years from last access.
8. **ROs** must curate their research data.
9. **ROs** must pay for this from their existing public funding streams.

Data management at Mech Eng, Bath



Key ERIM Research Findings

1. Poor framework for
 - ▶ pre-project considerations of data management;
 - ▶ data management during the research;
 - ▶ during-project data management for post-project re-use.
2. Poor knowledge of context in which data were generated:
 - ▶ engineering research data is very diverse;
 - ▶ large number of diverse research data records;
 - ▶ relations between data records complex.
3. Knowing the context is vital for understanding data.

Slide: Mansur Darlington

Engineering RDM Planning Guidance

Principles of Engineering Information Management

<http://tinyurl.com/KIM-PEIM>

Principles of Engineering Information Management

1. Parsimony
2. Granularity
3. Identity
4. Uniqueness
5. Usability
6. Reusability
7. Evaluation
8. Portability
9. Robustness
10. Discovery
11. Design

Engineering RDM Planning Guidance

Principles of Engineering Information Management

<http://tinyurl.com/KIM-PEIM>

Engineering RDM Planning Guidance

Principles of Engineering
Information Management

<http://tinyurl.com/KIM-PEIM>

Principles of Engineering
Research Data Management

<http://opus.bath.ac.uk/22201>

Thematic Analysis of DMP
Tools and Exemplars

<http://opus.bath.ac.uk/21278>

Engineering Research DMP
Requirement Specification

<http://opus.bath.ac.uk/21280>

Draft IdMRC Projects
Data Management Plan

<http://opus.bath.ac.uk/22200>

DMP Template for
IdMRC Projects

<http://opus.bath.ac.uk/23279>

Principles of Engineering Research Data Management

1. See the Principles of Engineering Information Management.
2. See the DCC Charter and Statement of Principles.
3. Data processing should be **reproducible**.
4. Use **generic/standard tools** where possible.
5. DMPs should support repurposing and help to support data reuse.
6. Treat records of a research activity **together as a set**.
7. Make **context/associations** between records explicit.
8. Methods of recording context should also be documented.
9. Confidentiality agreements should be as permissive as ethically possible.
10. RDM should be **costed** into each research proposal.
11. Supporting reuse/repurposing should **not get in the way** of use.
12. Any RDM tools should be **simple, engaging & easy to access**.

Engineering RDM Planning Guidance

Principles of Engineering
Information Management

<http://tinyurl.com/KIM-PEIM>

Principles of Engineering
Research Data Management

<http://opus.bath.ac.uk/22201>

Thematic Analysis of DMP
Tools and Exemplars

<http://opus.bath.ac.uk/21278>

Engineering Research DMP
Requirement Specification

<http://opus.bath.ac.uk/21280>

Draft IdMRC Projects
Data Management Plan

<http://opus.bath.ac.uk/22200>

DMP Template for
IdMRC Projects

<http://opus.bath.ac.uk/23279>

Engineering Research DMP Requirement Specification

Infrastructure

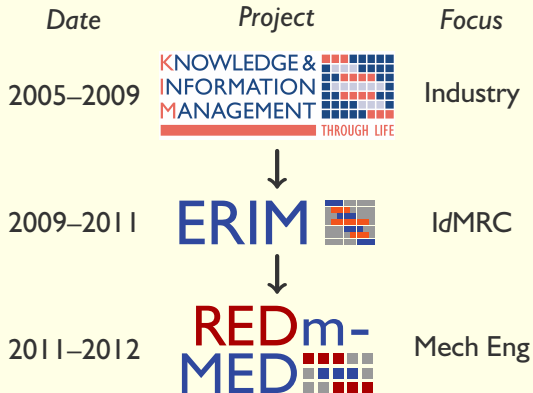
1. Links from DMP to key docs
2. Links from key docs to DMP
3. Purpose of DMP
4. Roles and responsibilities
5. Review and adherence
6. Version control
7. Budget
8. Storage, backup and security
9. Receiving repository

DMP contents

1. Summary of activity

2. Reuse of existing data
3. Fitting in with existing data
4. Preparing for expected reuse
5. Record manifest
6. Data generation and manipulation
7. Data organisation
8. Quality assurance
9. Data structures and formats
10. Data semantics

Data management at Mech Eng, Bath



Engineering RDM Planning Guidance

Principles of Engineering
Information Management

<http://tinyurl.com/KIM-PEIM>

Principles of Engineering
Research Data Management

<http://opus.bath.ac.uk/22201>

Thematic Analysis of DMP
Tools and Exemplars

<http://opus.bath.ac.uk/21278>

Engineering Research DMP
Requirement Specification

<http://opus.bath.ac.uk/21280>

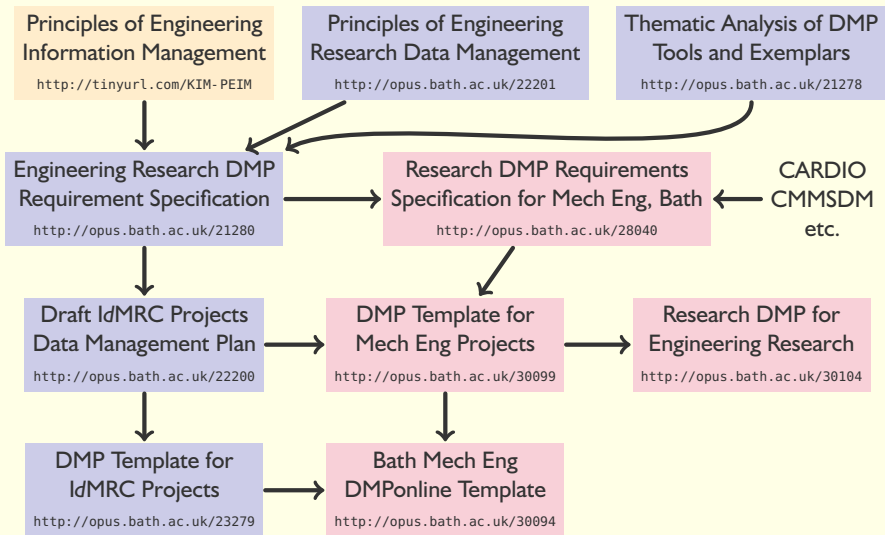
Draft IdMRC Projects
Data Management Plan

<http://opus.bath.ac.uk/22200>

DMP Template for
IdMRC Projects

<http://opus.bath.ac.uk/23279>

Engineering RDM Planning Guidance



Research DMP Requirements Specification

- ▶ Requirement
- ▶ Rationale
- ▶ Validation (DCC Checklist, applicability to research and data lifecycles)
- ▶ Role supported by requirement
- ▶ Responsibility – institution, department or project?
- ▶ Information/resource needed

Key documentation

Should be able to find quickly:

- ▶ Data management plan/record
- ▶ Project proposal (pre-award)
- ▶ Detailed project plan (post-award)
- ▶ Project record manifest
- ▶ Confidentiality agreements
- ▶ IPR statements, licences

Departmental RDM Wiki



 Edit  Share  Add  Tools 

Added by [Alex Ball](#), last edited by [Alex Ball](#) on May 22, 2012 ([view change](#)) [show comment](#)

Principal Investigator: Chris McMahon

Location on X drive: <file://localhost/X:/MechEng/Research/General/CAM/CAM-0007/>

Data Management Plan: n/a (but if we had one it would be attached)

Links

- [Project web site](#)
- [Project blog](#)

Description

The Research Data Management for Mechanical Engineering Departments (REDm-MED) Project is one of a number of projects being funded through Phase 2 of the JISC Managing Research Data Programme. The project will scope, specify, design and implement a research data management plan suited especially to the needs of the Department of Mechanical Engineering at the University of Bath. This work will build upon the team's work investigating the research data management needs of the Innovative Design and Manufacturing Research Centre (IdMRC) at the University of Bath during the JISC-funded ERIM Project and other recent work by the Managing Research Data Community.

Work completed in the ERIM Project in specifying a Research Activity Information Development (RAID) Associational Tool will be continued by the development of a prototype tool for integration within the specified research data management infrastructure. The presumption is made that the research data management needs of the participating department will intersect closely with those of similar departments in other research institutions, not only of higher education but in the wider commercial world. Thus, the aim of the research will be to implement effective and practical research data management where there is little or none and that can be adopted easily elsewhere.

The work will take place in conjunction with the DCC at Bath, and with the support of the University of Bath's Department of Mechanical Engineering and the Bath University Computer Service (BUCS). The REDm-MED project will be complementary to the University of Bath's JISC-funded Research 360 Project.



Project Record Manifest template

Project Data Record Manifest Template for IdMRC Projects

The **Project Data Record Manifest (PDRM)** constitutes the principal conduit through which the records relating to a research project may be identified and retrieved. It must be located in a publicly accessible and searchable place. The default location is an anonymous log-in page of the research project web.

The Project Data Management Plan and the Project Data Record Manifest should be considered a pair, and should be co-located.

The PDRM should be 'read-only', editing rights being limited to members of the originating research project team and by other nominated individuals such as the data manager. A versioning system must be in force.

Whilst the PDRM will be globally available, there will be some records associated with the research project which are confidential or sensitive. Access to records of this nature must be limited by placing the records in appropriately password-protected locations, this could be BCCS file spaces or within the research project web or other web space. If in doubt, the advice of the data manager (or failing that, the project PI) should be sought.

Summary of Research Activity

Project name
e.g. Long And Technical Textual Evaluation (LATTE)
Period of Project
e.g. October 2009 – March 2011
Lead and partner organizations
e.g. University of Bath (lead), University of Cambridge, University of Leeds
Principal Investigator (name and contact details)
Name:
Contact details:
Data access summary
Data access refers to the physical means by which access to records is constrained. The overarching data access provisions for this research project are recorded in the CMP associated with this PDRM. For details of confidentiality status of individual records see the Project Data Record List below. As a guide, data access should be either consistent with or more restrictive than the confidentiality status.
Receiving repository
e.g. The data from this Research Activity will be deposited according to the IdMRC CMP (see below).
or
The data from this research activity will be deposited in ____.
Related documentation
<ul style="list-style-type: none">• RCUK Policy and Code of Conduct on the Governance of Good Research Conduct• The University of Bath Good Practice Guide for Research• Engineering Research Data Management Plan Specification• IdMRC Projects Data Management Plan

Project Management Documentation

Note that some of these records may need to be placed in a password-protected storage area.

- Project Data Record Manifest: [\[wiki link\]](#)
- Project Proposal: [\[wiki link\]](#)
- Project Plan: [\[wiki link\]](#)
- Confidentiality agreement with [name]: [\[wiki link\]](#); note if this agreement is itself confidential it should be placed in an appropriately protected location
- Participant consent forms: [\[wiki link\]](#); [\[physical location/contact name/contact details\]](#)
- Ethics forms(s): [\[wiki link\]](#); [\[physical location/contact name/contact details\]](#)
- IRB Statement: [\[wiki link\]](#); [\[physical location/contact name/contact details\]](#)
- UK Data Archive deposit requirements: [\[wiki link\]](#)

Project Data Management Documentation

- Project Data Management Plan [\[wiki link\]](#) (this will be a reciprocal association, since the PDRM will identify the Project Data Record Manifest)
- RAD records: [\[wiki link\]](#) or
- Other data record associative documents [\[wiki link\]](#)

Project Data Record List

Every project data record should be listed in the table below in the form: Title, file name, record type, location, owner and contact details, confidentiality status

Record Type (for both electronic and physical records)

Every data record will be one of the following: research data record, content data record, associative data record, research object data record, experimental apparatus data record

Location

If all the files are archived in a single, central location, the location need be identified for the set of records (the Data Cases) only. For electronic records it is expected that a hyperlink or filepath to the location is recorded. For physical records the location should be described.

Owner

The 'owner' is the person currently responsible for the management of the record, and who is in a position to consider matters such as ownership and security. Ownership does not imply any rights to use or disposal. During the period that the research project is under way it is likely that the owner will be a research officer or an individual in a supervisory role. As project end the ownership should be transferred to an appropriate individual, such as the project PI or the data manager responsible. In many cases it will be appropriate for a research officer to retain ownership.

Confidentiality Status

Confidentiality status indicates what classes of people and what automated information-gathering systems may have right of the data record. It does not provide information about how such records are protected. It is likely that the confidentiality status will change during the life-cycle of the data record, in which case the status must be updated. Access is either free or limited. If access is free, then the term 'public domain' should be used. If this access is limited, then the entities who are permitted to see the data should be identified either by naming groups or individuals.

Record Title	File Name	Owner	Contact Details	Data Record Type	Confidentiality Status
Example					
IdMRC Research Project Data Record Manifest	idmrcm1/ID7/np	Manuel Davidson	emw@bath.ac.uk	associative data record	public domain

History of this PDRM



Project Record Manifest template

Every data record will be one of the following: *research data record, context data record, associative data record, research object data record, experimental apparatus data record.*

Location

If all the files are archived in a single, central location, the location need be identified for the set of records (the Data Case) only. For electronic records it is expected that a hyperlink or filepath to the location is recorded. For physical records the location should be described.

Owner

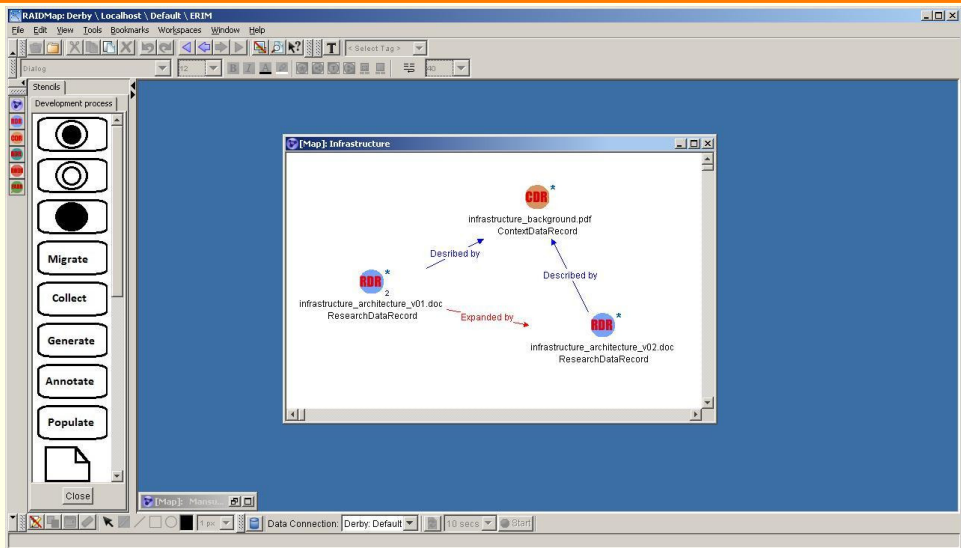
The 'owner' is the person currently responsible for the management of the record, and who is in a position to consider matters such as shareability and security. Ownership does not imply any rights to use or disposal. During the period that the research project is under way it is likely that the owner will be a research officer or an individual in a supervisory rôle. At project end the ownership should be transferred to an appropriate individual, such as the project PI or the data manager responsible. In many cases it will be appropriate for a research officer to retain ownership.

Confidentiality Status

Confidentiality status indicates what classes of people and what automated information-gathering systems may have sight of the data record; it does not provide information about how such records are protected. It is likely that the confidentiality status will change during the life-cycle of the data record, in which case the status *must* be updated. Access is either free or limited. If access is free, then the term 'public domain' should be used. If the access is limited, then the entities who are permitted to see this data should be identified either by naming groups or individuals.

Record Title	File Name	Owner	Contact Details	Data Record Type	Confidentiality Status
<i>Example:</i>					
<i>IdMRC Research Project Data Record Manifest</i>	<i>erim6man110217mjd</i>	<i>Mansur Darlington</i>	<i>ensmd@bath.ac.uk</i>	<i>associative data record</i>	<i>public domain</i>

Using RAIDmap



Rolling your own DMP regime?

- ▶ What do you want to achieve?
- ▶ What can be done by the institution/department, and what is left to researchers?
- ▶ What can be done with what you already have, and what do you need to improve?
- ▶ Can you introduce research data management so it decreases the burden on researchers?

Thank you for your attention

DCC Website: <http://www.dcc.ac.uk/>

Alex Ball: <http://www.ukoln.ac.uk/ukoln/staff/a.ball/>

ERIM <http://www.ukoln.ac.uk/projects/erim/>

REDm-MED <http://www.ukoln.ac.uk/projects/redm-med/>

Research360 <http://blogs.bath.ac.uk/research360/>