PRINCIPLES FOR ENGINEERING RESEARCH
DATA MANAGEMENT

MANSUR DARLINGTON, ALEX BALL, TOM HOWARD, STEVE CULLEY,
AND CHRIS McMATHON

erim6rep101028mjd10.pdf

ISSUE DATE: 14 DECEMBER 2010
Catalogue Entry
Title Principles for Engineering Research Data Management
Creator Mansur Darlington, Alex Ball, Tom Howard (authors)
Contributor Steve Culley, Chris McMahon
Subject information management; data management; research data
Description The Principles for Engineering Research Data Management are a set of consistent, high-level ideals that the authors believe should underlie any engineering research data management activity, and by which the quality of such activity should be judged. The principles were developed in the context of the ERIM Project, supporting data management at the IdMRC at the University of Bath, but are offered to the wider engineering research community as general guidance in the field of data management.
Publisher University of Bath
Date 28th October 2010 (creation)
Version 1.0
Type Text
Format Portable Document Format
Resource Identifier erim6rep101028mjd10
Language English
Rights © 2010 University of Bath

Citation Guidelines

Copyright
© 2010 University of Bath. This work is licensed under a Creative Commons Attribution-NonCommercial-ShareAlike 3.0 Unported License: http://creativecommons.org/licenses/by-nc-sa/3.0/
PRINCIPLES FOR ENGINEERING RESEARCH DATA MANAGEMENT

1. INTRODUCTION
The Principles for Engineering Research Data Management are a set of consistent, high-level ideals that the authors believe should underlie any engineering research data management activity, and by which the quality of such activity should be judged. They are consistent insofar as following one principle does not prejudice one’s ability to follow any of the others. They are high level in that, while they provide pragmatic and concrete guidance, they are widely applicable and avoid prescribing particular systems or implementations. They are ideals rather than strict requirements because, while they are to be strived for, it is acknowledged that it may not be possible to apply them fully in all circumstances.

The key motivation for these principles is to provide the basis upon which data management practices can be engaged in which support the three activities of:

1. making existing research data available and fit for a future known research activity (data re-purposing);
2. managing existing research data such that it will be available for a future unknown research activity (supporting data re-use);
3. using research data for a research purpose or activity other than that for which it was intended (data re-use).

The principles were developed in the context of the ERIM Project, supporting data management at the IaMRC at the University of Bath, but are offered to the wider engineering research community as general guidance in the field of data management. They should be of interest to practitioners engaged in data management and those providing support and tools for such management.

2. THE PRINCIPLES OF ENGINEERING RESEARCH DATA MANAGEMENT

1. Engineering research data management should be informed by the more general Principles of Engineering Information Management\(^1\).

2. Engineering research data management should be consistent with the approach outlined in the DCC Charter and Statement of Principles.

3. Research data management planning should reflect the need for and provide guidance to support research data reproducibility. In particular, where appropriate to the data, each stage of processing from first generation to final data should be reproducible from the records in the data case, alongside readily available tools. (Principle inspired by the Steps for Reproducible Research.)

4. The notions of re-usability and re-purposing should be supported by the use, where possible, of generic or standard data generation and manipulation tools. Where specialist tools are used consideration should be given to the use of virtual machines for the later revisiting, validation and reproduction of data.

\(^1\) A list of the Principles for Engineering Information Management can be found on page 4.
5. All research activities should be supported by a data management plan which contains aspects which explicitly set out to help 'supporting data re-use' and re-purposing.

6. The data objects associated with a research activity should be considered to be a mutually explanatory set and should be managed principally as a set.

7. The data objects associated with a research activity should be associated explicitly – preferably in a human understandable form – by the use of contextual information about such things as the relationship holding between data objects (data records), versioning and the research activity itself.

8. In order to support re-purposing the specification of the contextualizing methods should be made available, to aid their interpretation, as part of the data management plan and later as part of the contextual documentation on the data records associated with a research activity.

9. Any agreement on the confidentiality or otherwise of research data should be as permissive as can be negotiated with affected parties, within the bounds of research ethics.

10. Due consideration should be given in research project planning to the role of data management, the importance of which should be reflected in the resources requested and made available for such management work.

11. Requirements specified in data management plans for the activities of 'supporting re-use' and 're-purposing' should hinder use activities as little as possible.

12. The tools put in place to assist in the satisfaction of requirements specified in data management plans should be simple, engaging and easy to access. (Principle inspired by the JISC-funded Incremental Project.)

3. The Principles of Engineering Information Management

1. The Principle of Parsimony: Create, record and retain information only if necessary.

2. The Principle of Granularity: Record information in a storable information object at a granularity appropriate for use and re-use.

3. The Principle of Identity: Give an information object a unique and persistent identifier.

4. The Principle of Uniqueness: Create an information entity once only and explicitly reference it everywhere else.

5. The Principle of Usability: Design an information object explicitly to achieve its intended goals.

6. The Principle of Reusability: Design an information object explicitly to maximise its potential for reuse wherever appropriate.

7. The Principle of Evaluation: Assess and assign the value of an information object throughout its life from creation to disposal.

8. The Principle of Portability: Create an information entity and its annotations systematically using representations supporting perpetual reuse.
9. The Principle of Robustness: Use robust methods to capture, create and manipulate information entities.

10. The Principle of Discovery: Actively employ the information repository as a resource for learning and discovery.

11. The Principle of Design: Design all aspects of information management to satisfy the organisation’s current and future needs.

4. REFERENCES


