

## EPSRC POLICY FRAMEWORK ON RESEARCH DATA MANAGEMENT

### A Summary

#### Introduction

In May 2011, the EPSRC (Engineering and Physical Sciences Research Council) set out their expectations concerning the management and provision of access to EPSRC funded research data. A range of institutional policies and practices can satisfy the expectations. They encourage approaches aligned to EPSRC expectations which are appropriate to the institution's structures and cultures.

The expectations arise from seven core principles - aligned with the RCUK principles on data sharing. The following are regarded as particularly important:

Publicly funded research data should be made as widely and freely available as possible in timely and responsible manner

research process should not be damaged by the inappropriate release of such data

#### Definitions

**Research data** are recorded factual material commonly retained by and accepted in the scientific community as necessary to validate research findings (irrespective of the format in which it is created).<sup>1</sup>

**Digital curation** is about maintaining and adding value to a trusted body of digital information for both current and future use: in other words, it is active management and appraisal of digital information over its entire life cycle.<sup>2</sup>

#### Principles

The following principles (aligned with RCUK principles on sharing of research data) should inform all decisions relating to the management of research data.<sup>3</sup>

1. EPSRC-funded research data is a public good produced in the public interest and should be made freely and openly available with as few restrictions as possible in a timely and responsible manner.
2. EPSRC recognises that there are legal, ethical and commercial constraints on release of research data. To ensure that the research process (including the collaborative research process) is not damaged by inappropriate release of data, research organisation policies and practices should ensure that these constraints are considered at all stages in the research process
3. Sharing research data is an important contributor to the impact of publicly funded research. To recognise the intellectual contributions of researchers who generate, preserve and share key

---

<sup>1</sup> <http://www.epsrc.ac.uk/about/standards/researchdata/Pages/scope.aspx>

<sup>2</sup> [http://www.ukoln.ac.uk/ukoln/staff/m.pennock/publications/docs/lib-arch\\_curation.pdf](http://www.ukoln.ac.uk/ukoln/staff/m.pennock/publications/docs/lib-arch_curation.pdf)

<sup>3</sup> <http://www.epsrc.ac.uk/about/standards/researchdata/Pages/principles.aspx>

research datasets, all users of research data should acknowledge the sources of their data and abide by the terms and conditions under which they are accessed.

4. EPSRC-funded researchers should be entitled to a limited period of privileged access to the data they collect to allow them to work on and publish their results. The length of this period will depend on the scientific discipline and the nature of the research.
5. Institutional and project specific data management policies and plans should be in accordance with relevant standards and community best practice and should exist for all data. Data with acknowledged long term value should be preserved and remain accessible and useable for future research.
6. Sufficient metadata should be recorded and made openly available to enable other researchers to understand the potential for further research and re-use of the data. Published results should always include information on how to access the supporting data.
7. It is appropriate to use public funds to support the preservation and management of publicly-funded research data. To maximise the scientific benefit which can be gained from limited budgets, the mechanisms for managing and providing access to research data should be both efficient and cost-effective in the use of such funds.

#### **Expectations - with UEA Action Points**

There are nine separate points raised by EPSRC<sup>4</sup>:

EPSRC has the following clear expectations of organisations in receipt of EPSRC research funding:

- i. Research organisations will promote internal awareness of these principles and expectations and ensure that their researchers and research students have a general awareness of the regulatory environment and of the available exemptions which may be used, should the need arise, to justify the withholding of research data;

**ACTION:** Identify the mechanisms by which the data management policy and the regulatory environment can be incorporated into information and training for EPSRC funded researchers and others.

- ii. Published research papers should include a short statement describing how and on what terms any supporting research data may be accessed.

**ACTION:** Incorporate information about this in the existing funding award process, and provide researchers with the support they need to take responsibility for including the statement.

---

<sup>4</sup> <http://www.epsrc.ac.uk/about/standards/researchdata/Pages/expectations.aspx>

- iii. Each research organisation will have specific policies and associated processes to maintain effective internal awareness of their publicly-funded research data holdings and of requests by third parties to access such data; all of their researchers or research students funded by EPSRC will be required to comply with research organisation policies in this area or, in exceptional circumstances, to provide justification of why this is not possible.

**ACTION:** Prepare a data management policy and associated processes to implement this, including a Data Management Plan template, and guidelines on metadata.

- iv. Publicly-funded research data that is not generated in digital format will be stored in a manner to facilitate it being shared in the event of a valid request for access to the data being received (this expectation could be satisfied by implementing a policy to convert and store such data in digital format in a timely manner);

**ACTION:** Work with researchers in SCI to establish what kinds of data are currently not generated in digital format. Consider options for digitisation (including feasibility and cost), and when this should be carried out.

- v. Research organisations will ensure that appropriately structured metadata describing the research data they hold is published (normally within 12 months of the data being generated) and made freely accessible on the internet; in each case the metadata must be sufficient to allow others to understand what research data exists, why, when and how it was generated, and how to access it. Where the research data referred to in the metadata is a digital object it is expected that the metadata will include use of a robust digital object identifier (For example as available through the DataCite organisation - <http://datacite.org>).

**ACTION:** Work with researchers in SCI to find out what metadata they currently produce, and how it is used. Investigate standards required for metadata, and whether this varies by discipline. Provide necessary training to enable researchers to produce robust metadata to the required standard. Implement a storage and retrieval system to publish the metadata.

- vi. Where access to the data is restricted the published metadata should also give the reason and summarise the conditions which must be satisfied for access to be granted. For example 'commercially confidential' data, in which a business organisation has a legitimate interest, might be made available to others subject to a suitable legally enforceable non-disclosure agreement.

**ACTION:** Incorporate guidance and work processes for restriction of access to data into the current provisions for FOIA/EIR requests.

- vii. Research organisations will ensure that EPSRC-funded research data is securely preserved for a minimum of 10-years from the date that any researcher 'privileged access' period expires or, if others have accessed the data, from last date on which access to the data was requested by a third party; all reasonable steps will be taken to ensure that publicly-funded data is not held in any jurisdiction where the available legal safeguards provide lower levels of protection than are available in the UK

**COMMENTS:** This is the main area of concern for ITCS. The second part of the statement refers to the use of data repositories or cloud services outside the EU (or countries without a safe harbour agreement). EPSRC do not provide a data repository, but there are many others available which could be used. UEA currently does not have one – does the EPSRC Framework support a case to develop one?

Are we likely to want to meet storage requirements in house? We do not know what those storage requirements are, either in the long (10 year) or short term, and how expectations such as (iv) (digitisation) are going to affect those requirements.

**ACTION:** To collect information from a range of researchers in the Faculty of Science about the type and volume of data they produce, what data they retain or discard, how it is currently stored, and what they would require of a data storage system as potential clients.

- viii. Research organisations will ensure that effective data curation is provided throughout the full data lifecycle, with ‘data curation’ and ‘data lifecycle’ being as defined by the Digital Curation Centre. The full range of responsibilities associated with data curation over the data lifecycle will be clearly allocated within the research organisation, and where research data is subject to restricted access the research organisation will implement and manage appropriate security controls; research organisations will particularly ensure that the quality assurance of their data curation processes is a specifically assigned responsibility;

**ACTION:** A long term storage solution is required, which could be delivered locally or via sector or discipline specific data repositories hosted outside the organisation. This/these will need to be linked with the data access request process so that the data retention period is maintained (see vii above).

These decisions need to be informed by the availability of existing resources within UEA and externally, and the proposals of other organisations in the sector and specific fields.

- ix. Research organisations will ensure adequate resources are provided to support the curation of publicly-funded research data; these resources will be allocated from within their existing public funding streams, whether received from Research Councils as direct or indirect support for specific projects or from higher education Funding Councils as block grants.

**ACTION:** As the roadmap is developed the costs of various options need to be developed and options for resourcing them discussed and prioritised.

### **Impact, timescales and support**

We are expected to review our research data management practices and capabilities. Changes could be to clarify and formalise existing processes and policies, or to create new posts, and invest in new infrastructure. Some actions may be completed in the medium to long term.

Therefore there is a need to **develop a roadmap** to 'align our policies and processes with EPSRC's expectations **by 1 May 2012**' and **be fully compliant by 1 May 2015**. EPSRC will monitor progress and compliance, and investigate non-compliance.

Research projects are going to need to cost storage into grants for both working data for the life of the project, and also archive data for the period after the project finishes. This will require a cost effective and sustainable storage infrastructure. ISD are the process of discussing with their storage suppliers the wider IT research storage requirements particularly with EPSRC in mind.

Support is available to those working to comply with the framework. Advice is available from JISC, the Digital Curation Centre, and the Information Commissioner's Office.<sup>5</sup>

### General points

The following general points about the EPSRC expectations arose as a result of this consideration:

- **Other funders' requirements.** Other funders might adopt the EPSRC principles and approach to RDM, and meeting these expectations could put the institution in a strong position generally across all sources of funding.

BBSRC already has a policy on Data Sharing. All applications seeking research grant funding from BBSRC must submit a statement on data sharing. This should include concise plans for data management and sharing as part of research grant proposal or provide explicit reasons why data sharing is not possible or appropriate.<sup>6</sup>

- **Research concentrated in an elite.** This could be one of the means of marking out a difference between the small number of elite research institutions and the rest for whom research activity will progressively diminish. Only the best facilitated institutions will be able to meet the expectations and apply successfully for funding.
- **DCC help.** The DCC may be looking into developing more structured ways of developing a roadmap to meet the EPSRC expectations, and we will continue to work with the DCC to ensure we can make full use of any tools which become available.
- **Evaluation.** The project should incorporate a proportionate system of ongoing evaluation of efficacy, impact and cost.

---

<sup>5</sup> <http://www.epsrc.ac.uk/about/standards/researchdata/Pages/impact.aspx>

<sup>6</sup> <http://www.bbsrc.ac.uk/web/FILES/Policies/data-sharing-policy.pdf>

## EPSRC RDM roadmap

Author: Raymond Scott (ISD), Sue Steel (REN)

Date: 2 March 2012

Version: 0.2

**This document is a roadmap indicating how the University plans to meet the research data management expectations of the EPSRC by 1 May 2015.**

Processes will be developed which will lead to a University Research Data Management Policy, which will also satisfy the EPSRC expectations. The resulting systems should be compatible with existing good practice, and will be capable of University-wide implementation.

### Version history

Version	Date	Note
0.2	2/3/2012	Second draft

### Introduction

In May 2011, the EPSRC (Engineering and Physical Sciences Research Council) set out their expectations<sup>1</sup> concerning the management and provision of access to EPSRC funded research data. A roadmap indicating what the University plans to do to meet the expectations is required by 1 May 2012, and the implementation of those actions should be complete by 1 May 2015. The expectations arise from seven core principles - aligned with the RCUK principles on data sharing<sup>2</sup>.

More detailed information about the EPSRC requirements is contained in the accompanying document EPSRC Data Management Summary v2.

### Definitions

- **Research data** are recorded factual material commonly retained by and accepted in the scientific community as necessary to validate research findings (irrespective of the format in which it is created).<sup>3</sup>
- **Digital curation** is about maintaining and adding value to a trusted body of digital information for both current and future use: in other words, it is active management and appraisal of digital information over its entire life cycle.<sup>4</sup>
- A **curation lifecycle model** documents the relationships between all the stages in the existence of digital information, to enable active management of the resource over time thus maintaining accessibility and usability. (Typically lifecycle runs: conceptualise > creation > active use > appraisal and selection > disposal > transfer > storage and preservation > access and reuse)<sup>5</sup>

<sup>1</sup> <http://www.epsrc.ac.uk/about/standards/researchdata/Pages/expectations.aspx>

<sup>2</sup> <http://www.epsrc.ac.uk/about/standards/researchdata/Pages/principles.aspx>

<sup>3</sup> <http://www.epsrc.ac.uk/about/standards/researchdata/Pages/scope.aspx>

<sup>4</sup> [http://www.ukoln.ac.uk/ukoln/staff/m.pennock/publications/docs/lib-arch\\_curation.pdf](http://www.ukoln.ac.uk/ukoln/staff/m.pennock/publications/docs/lib-arch_curation.pdf)

<sup>5</sup> <http://www.dcc.ac.uk/digital-curation/what-digital-curation>

## Abbreviations

- **EPSRC** Engineering and Physical Sciences Research Council
- **FOIA** Freedom of Information Act 2000
- **EIR** Environmental Information Regulations 2004
- **DMP** Data Management Plan
- **RDM** Research Data Management
- **RRS** Records Retention Schedule

## Costs

Systems and infrastructure for storing and managing access to research data have costs. It is reasonable and appropriate 'to use public funds to also fund the associated data management costs'. EPSRC expects us to make provision from within public research funding received (both direct and indirect funding). A more efficient approach would be to appoint a 3rd party or collaborate with other institutions on a shared service. EPSRC says this would be 'entirely acceptable'.<sup>6</sup>

## Evaluation

The project should incorporate a proportionate system of ongoing evaluation of efficacy, impact and cost.

---

<sup>6</sup> <http://www.epsrc.ac.uk/about/standards/researchdata/Pages/responsibility.aspx>

**Actions**

<b>Area</b>	<b>Action</b>	<b>Responsibility</b>	<b>Due Date</b>	<b>EPSRC Exp.</b>
Communications	Use a range of methods to work with the research community at UEA. Investigate internal and external requirements for research data management, and establish the work needed to provide a comprehensive system.	ISD/REN	October 2012	N/A
Communications	Include references to obligations under legislation in annual reminder to all staff	VCO/ISD	October 2012 (and annually)	(i)
IT systems	Determine research data storage needs, including consultation with researchers and identification of existing external repositories	ISD/PI	July 2013	(viii)
IT systems	Develop and implement storage solution to meet the requirements for research data storage including migration of existing research data (both short term during the life of the project, and long term in a data repository)	ISD	Dec 2014	(viii)
IT systems	Develop a business case for the creation of an institutional data repository service (system, storage, processes and staffing) so that research data can be managed and preserved independently of researchers where no other suitable facility exists	ISD	Dec 2013	(vii)
IT systems	Ensure that the institutional digital repository (EPrints) is able to hold and link research papers and associated data papers	ISD/REN	July 2013	(v)
IT systems	Investigate tools to help with metadata capture and creation	ISD/SCI Schools	Dec 2013	(v)
Policy	In tandem with the consultation with researchers, review existing policies and procedures.	ISD/REN	October 2012	N/A
Policy	Create and seek approval on a RDM policy which includes the expectation that researchers produce a data management plan, produce a metadata paper, include data access arrangements in published papers, and which states where responsibilities for data curation (researcher) and the QA of data curation (research director) lie	ISD	January 2013	(ii), (iii), (v), (viii)
Policy	Complete consultations and install final Research Data Management Policy.	ISD	July 2013	(iii)
Research management	Set up a working group to represent all interests across the project, including academic and research staff, post graduate researchers, technical and administrative staff. This will identify and allocate specific tasks, receive feedback and report in the first instance to Research Executive and ISD Research Board.	REN	July 2012	N/A
Research management	Identify all costs associated with data curation and determine the levels required in those identified areas year on year	ISD/REN	July 2013	(ix)

Area	Action	Responsibility	Due Date	EPSRC Exp.
Research management	Ensure funding streams are identified to maintain systems, storage and staffing to support the data curation lifecycle	ISD, REN, FIN	July 2012	(ix)
Research management	Include costs relating to the curation of the research data (short-term and long-term storage) in research grant applications and establish processes to charge for UEA facilities	REN, ISD, FIN	Dec 2012	(vii), (viii), (ix)
Research management	Clarify with EPSRC and internally questions around which funding streams will cover the costs of the digitisation of data	REN	Dec 2012	(iv)
Research management	Clarify with EPSRC the requirement that the metadata paper is available publically 12 months after data generation rather than at the same time as the publication of the data	REN	July 2012	(v)
Research management	Establish processes for the storage of DMPs	REN	Dec 2012	(iii)
Research management	Create an RRS for EPSRC research data reflecting the requirement to keep the data for 10 years after last access	REN/ISD	July 2013	(vii)
Research management	Develop QA system and processes for data storage (based on any relevant existing standards)	REN/ISD	Dec 2014	(viii)
Research management	Review and design processes for handling requests for access to research data. Processes should include the need to reset the preservation time clock after each access	REN/ISD	Dec 2013	(iii), (vii)
Research management	Implement processes for handling requests for access to research data	REN/ISD	Dec 2014	(iii)
Research management	Establish processes for the digitisation of non-digital research data	ISD (LIB/PRG)/REN	July 2013	(iv)
Research practice	Add information on access to data to published papers	PIs	Dec 2012	(ii)
Research practice	Add information about research data formats and locations to DMPs	PIs	Dec 2012	(iv)
Research practice	Produce metadata papers including both high level repository metadata and low level storage metadata	PIs	July 2014	(v)
Support	Process to handle requests for access to data as FOIA/EIR requests	ISD	Complete	(vi)
Support	Creation of web pages for researchers describing EPSRC expectations	REN	August 2012	(i)
Support	Provide advice to researchers on IPR and confidentiality	REN	Complete	(vi)
Training	Develop training to encourage and equip researchers with skills to curate their research data	ISD (LIB)	July 2013	(viii)

<b>Area</b>	<b>Action</b>	<b>Responsibility</b>	<b>Due Date</b>	<b>EPSRC Exp.</b>
Training	Develop training to encourage and equip researchers with skills to produce research data metadata papers (including whether or not access to the data is restricted)	ISD (LIB)/SCI Schools	July 2013	(v), (vi)
Training	Incorporation of research data management processes into REN training modules	REN	August 2012	(i)
Training	Add advice on EPSRC expectations (inc. the need to include details on access to data to published papers) in REN training modules	REN	August 2012	(i), (ii)
Training	Incorporation of EPSRC expectations into SCI PPD and CSED training modules	REN/ISD working with SCI PPD / CSED	August 2012	(i)