The University of Exeter's Roadmap for EPSRC's Research Data Management Expectations

Live document: Version 1.3 published on 28.02.2013

Contents

Context

1. EPSRC requirements

The open scholarship and open access agenda is fast moving, with RCUK mandating open access publication of research outputs and an announcement on open access to research data expected soon. Many of the research councils, notably the EPSRC and the MRC, already mandate compliance with the seven core RCUK fundamental principles on data sharing.[1] EPSRC has developed 9 expectations for institutions in receipt of EPSRC funding, which arise from the RCUK principles and uphold the primary concepts that:

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

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

EPSRC will monitor progress against their expectations in accordance with the timescales presented in the letter from the Director Communications, Information and Strategy, EPSRC, which is included for reference in appendix A. The major milestones are:

- 1st May 2012: "institutions should have a clear roadmap in place to align policies and processes with EPSRC's expectations";

- 1st May 2015: "institutions should be fully compliant with EPSRC's expectations".

The full expectations of EPSRC are listed for in appendix B. The current status of internal developments leading to compliance with these milestones is outlined below.

2. Overview of Current Status

The University of Exeter has successfully applied for external funding for the initial development of research data management infrastructure – the ERDMS project and Open Exeter. £800k from the Science Strategy has funded the ERDMS [Exeter Research Data Management Service] project which has established a 'deep archive' for research data and £250k has been received from JISC for the Open Exeter project to embed the institutional repository and appropriate sharing [based on ethical, legal and commercial considerations] of research data into routine practice.

The Open Exeter project is led by Academic Services [Library; Exeter IT] in partnership with Research and Knowledge Transfer (RKT), with Professor Nick Talbot as project executive and institutional sponsor. Open Exeter is developing the skills, training, policy, governance and cost recovery models, informed by a Data

---

[1] The RCUK Common Principles on Data Policy are available at
http://www.rcuk.ac.uk/research/Pages/DataPolicy.aspx

Asset Framework [DAF] survey to audit the current research data management practice and complex storage behaviour within our active research community. The DAF has been noted by the DCC as an example of best practice, and has been supplemented by detailed interviews for a full understanding of researchers' processes enabling tailored training and guidance for data management and storage to be developed.[2]

A Task and Finish Group has been established to draft and oversee ratification of institutional policy and governance for research data management and Open Access. Its remit includes information security and compliance with FOI and DPA requirements and liaison with other groups as necessary.

The measures currently being taken, or planned, to ensure compliance with EPSRC and other funders' policies include:

- The establishment of governance structures to provide continuing leadership and coordination of data management and Open Access issues.

- Development and ratification of an institutional research data management policy and associated procedures.

- Associated with the above, clarification of organisational and individual responsibilities and obligations.

- Provision of a robust, extensible, actively-curated research data repository available to all University of Exeter researchers for preservation and dissemination of their research data.

- Quality control mechanism for reviewing and approving submissions to the repository.

- Option to submit metadata-only data records where placing data on Open Access is inappropriate (where the Data Protection Act applies, or for reasons of commercial sensitivity, for example).

- Upskilling those currently providing advice to researchers to ensure research data management knowledge and skills are widely dispersed (for example, Research and Knowledge Transfer staff, Subject Librarians).

- Ongoing programme of research data management training delivered via a number of options, aimed at increasing the knowledge and skills of researchers and PGRs.

- Development of a range of discipline-specific tools and resources for embedding good practice in management of research data and its associated metadata and documentation in the research lifecycle.

More detailed information about these measures can be found in the next section "EPSRC Expectations".

3. Document Status and Timeframe

---

[2] The Summary Findings of the Open Exeter Data Asset Framework (DAF) Survey are available at: http://hdl.handle.net/10036/3689

This is a living document which will be sent for informational purposes to the Research and Knowledge Transfer Management Group and the EPSRC Stakeholder Group. It is expected that the document will be updated after each phase indicated below:

Phase 0: Until end of March 2013: Planning, pilot activity, early deliverables.

Phase 1: April 2013 – December 2013: Initial rollout of primary services.

Phase 2: January 2014 – May 2015: Continued rollout, maturation of services.

4. Authors

The roadmap was developed by members of the Open Exeter team, RKT and the Library: Gareth Cole, Jill Evans, Jessica Gardner, Hannah Lloyd-Jones, Stephen Trowell.

EPSRC Expectations

1. 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.

Proposed Actions

Objectives	Actions	Responsible Team(s)	Target date
1.1 Develop the data management skills and knowledge of researchers and PGRs.	Include research data management training in the following courses/sessions as well as develop dedicated training sessions on research data management: • RDP • PCAP • Doctoral Supervision Course • Inductions for PGRs • Inductions for research staff • DoRs/Peer Reviewers	Phase 0: Open Exeter project Phase 1 and 2: Library	Phase 0
1.2 Develop the data management skills and knowledge of relevant professional services staff.	Develop inductions/training for professional services staff including: • Library • RKT • College research support staff and CDOs • Exeter IT	Phase 0: Open Exeter project Phase 1 and 2: Library	Phase 0

	• Legal services • Graduate Faculty Office		
1.3 Develop RDM website as a first point of guidance for researchers and professional services staff.	Guidance website developed and updated.	Phase 0: Open Exeter project Phase 1 and 2: Library	Phase 0
1.4 Develop and implement a communications plan	Develop and implement a communications plan.	Phase 0: Open Exeter project Phase 1 and 2: Library	Phase 1
1.5 Provide clear guidance on the regulatory environment and exemptions which may be used to justify the withholding of research data.	Integrate training on the DPA, FOI requests, commercial and sensitive data into RDM training packages and online guidance.	Phase 0: Open Exeter project Phase 1 and 2: Exeter IT, Library; RKT	Phase 0

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

Proposed Actions

Objectives	Actions	Responsible Team(s)	Timescale
2.1 Advocate the benefits of increasing discoverability of data.	Training and guidance on the need to include statement on accessibility of data and development of model data access statement. Include requirement for this statement in policy.	Phase 0: OA and RDM Task and Finish Group; Open Exeter project Phase 1 and 2: Library	Phase 0
2.2 Enable the linking of data and publications in the University's repository.	Ensure that the data and publications can be linked in the institutional repository.	Phase 0: Open Exeter project Phase 1 and 2: Exeter IT, Library	Phase 0

3. 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.

Proposed Actions

Objectives	Actions	Responsible Team(s)	Timescale
3.1 Develop policies on research data management.	Institutional policies on RDM developed, ratified and updated regularly. Encourage and support the development of College level and research group level policies.	Phase 0: OA and RDM Task and Finish Group; Open Exeter project Phase 1 and 2: Library; Graduate Faculty Board – PGR policy; RKT Management Group – researcher policy	Phase 0
3.2 Track third party requests for access to data.	Develop procedures for tracking third party requests for access to data.	Exeter IT; Library; RKT	Phase 1
3.3. Raise awareness of data management and archiving issues in early stages of drawing up contracts with research partners.	Review and specify RDM requirements in new contracts with research partners to ensure that appropriate legal clauses are in place and maximise accessibility and impact of data outputs where appropriate.	Legal Services; RKT	Phase 1

4. 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).

Proposed Actions

Objectives	Actions	Responsible Team(s)	Timescale
4.1 Investigation of strategies for improving access and visibility of non-digital research data.	Investigate strategies to enhance storage of non-digital data, to include cost estimates. Open Exeter will explore the feasibility of 'live research data' capture using new technologies [e.g. iPads as 'field books' and digital pens]. Investigate strategies for potential digitisation of non-digital research data, to include cost estimates. £5k pilot arising from the DAF studies to investigate the feasibility of auditing, digitising and archiving of analogue research data [e.g. lab books] to inform future decisions over larger scale digitisation of these 'hidden' assets.	Phase 0: Open Exeter project Phase 1 and 2: Exeter IT; Library	Phase 1
4.2 Develop policies on non-digital data.	Include non-digital data in University's OA and RDM policy and clarify that it should be	Phase 0: OA and RDM Task and Finish Group; Open Exeter project	Phase 0

	registered with the University's repository, specifying how it can be accessed.	Phase 1 and 2: Exeter IT; Library	
	Develop specific policy on the storage, digitisation and disposal of non-digital research data.		Phase 1

5. 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).

Proposed Actions

Objectives	Actions	Responsible Team(s)	Timescale
5.1 Seek to increase impact of Exeter research data through increasing the discoverability of data through appropriate metadata.	Training and guidance on how to document research data as well as how to complete the metadata record when uploading data sets to the institutional repository to enhance data discovery. Investigate research data deposit on EDA using PGRs to create the discipline-specific metadata standards as a £5k pilot.	Phase 0: Open Exeter project Phase 1 and 2: Library; RKT	Phase 0
5.2 Include requirement for metadata record in policy even when data not available on open access.	Requirement for metadata record included in policy.	OA and RDM Task and Finish Group; Open Exeter project	Phase 0
5.3 Data repository to assign robust digital object identifiers to digital objects.	Ensure that repository can assign digital object identifiers to digital objects.	Phase 0: Open Exeter project Phase 1 and 2: Exeter IT	Phase 0

6. 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.

Proposed Actions

Objectives	Actions	Responsible Team(s)	Timescale
6.1 Ensure that the conditions for access to data that is restricted are clear.	Develop guidelines for the metadata records of restricted access data e.g. use of embargo periods, commercialisation reasons. Provide training and guidance on how restricted access data may be shared.	Phase 0: Legal Services; Open Exeter project; RKT Phase 1 and 2: Legal Services; Library; RKT	Phase 0

7. 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.

Proposed Actions

Objectives	Actions	Responsible Team(s)	Timescale
7.1 Ensure that research data is securely preserved according to EPSRC expectations.	Update UoE Code of Good Practice in Research[3] according to new OA and RDM policy. Develop standards and training on data retention and disposal. Track third party access to and preservation of EPSRC research data. Ensure that data in the University's repository will be stored and backed-up in the UK.	Exeter IT; Library; RKT	Phase 1

---

[3] Available at http://www.exeter.ac.uk/staff/development/research/support/codeofpractice/goodpracticeinresearch/

8. 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.

Proposed Actions

Objectives	Actions	Responsible Team(s)	Timescale
8.1 Clarify research data management responsibilities.	Define cross-University responsibilities for research data management, including responsibility for the quality assurance of the data curation processes.	Phase 0: Open Exeter project Phase 1 and 2: Exeter IT, Library, RKT	Phase 1
8.2 Sensitive data managed according to standards of good practice.	Training and guidance on the management of sensitive data for researchers and professional services staff.	Phase 0: Open Exeter project Phase 1 and 2: Exeter IT, Library	Phase 0

9. 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.

Proposed Actions

Objectives	Actions	Responsible Team(s)	Timescale
9.1 Include research data management infrastructure requirements in the institutional planning lifecycle.	Ensure that research data management is included in high-level institutional budget planning.	Phase 0: Open Exeter project Phase 1 and 2: Exeter IT, Library, RKT	Phase 0
9.2 Development of business plan for sustainable research data management service.	£800k is included in the IT Infrastructure planning for research systems for a 5+ years for upgrade and replacement of the institutional repository. Sustainability and business plan models for research data storage and curation services are being developed through the Open Exeter project, with the aim to agree a cost recovery model [perhaps via indirects] which retains the principle of free services at point of use.	Phase 0: Open Exeter project Phase 1 and 2: Exeter IT, Library, RKT	Phase 0
9.3 Ensure sustainable funding resources for research.	Identify and apply for further funding opportunities.	Phase 0: Open Exeter project Phase 1 and 2: Exeter IT, Library, RKT	Phase 0

9.4 Encourage researchers to include costs of research data management in data management plans when possible.	Raise awareness about including costs in data management plans in data management plan training.	Phase 0: Open Exeter project Phase 1 and 2: Library	Phase 0

Appendix A – Letter from Attila Emecz about EPSRC expectations

EPSRC
Pioneering research and skills

Our ref: 9558

10th February 2012

Dear Vice Chancellor,

We wrote in April last year to let you know about our Policy Framework on Research Data, which sets out our principles and expectations concerning how the institutions we fund will ensure that the research data generated as a result of our support is appropriately managed and shared. The importance of ensuring data is accessible was recently re-emphasised in the government's Innovation and Research Strategy for Growth1. The Strategy explicitly required Research Councils to ensure there is compliance with current policies on access to research outputs.

I am therefore writing to remind you that the EPSRC policy timetable, whilst acknowledging that institutions would need time to review their current practice, gave two deadlines: by 1st May 2012 each institution should have a clear roadmap in place to align policies and processes with our expectations, and by 1st May 2015 each institution should be fully compliant with our expectations. You may also recall that the EPSRC policy indicates that if, after 1st May 2015, an institution is found to be deliberately obstructing the proper sharing of research data, or otherwise seriously failing to comply with EPSRC's expectations, this will initiate a process that could ultimately lead to it being declared ineligible for EPSRC support.

However, a key aspect of our policy is that it is not prescriptive about how our expectations should be met; it recognises that what may be appropriate for one institution may not be so for another. Since we are not in a position to evaluate the appropriateness of institution-specific plans I wanted to let you know that we will not require roadmaps to be formally submitted to us. We will however ask that institutions self-certify that they have taken and will continue to take the necessary steps to ensure they are able to comply with our expectations from 1st May 2015, and we may ask to see individual roadmaps on a case by case basis.

The full policy framework is available on our website and I would be grateful if you could ensure that the colleagues responsible for this area in your institution are reminded of our requirements. If you or these colleagues have any questions relating to the policy please address them directly to Ben Ryan, Senior Evaluation Manager [ben.ryan@epsrc.ac.uk].

Yours sincerely

Attila Emecz

Director Communications, Information and Strategy

Appendix B - EPSRC Expectations

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;

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

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.

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];

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**].

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.

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 take 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

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;

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.

Appendix C – Letter from Sean Fielding about EPSRC Research Data Policy Framework

UNIVERSITY OF EXETER

RESEARCH & KNOWLEDGE TRANSFER

The Innovation Centre
Rennes Drive
Exeter
EX4 4RN

Telephone +44 (0)1392 723456
Fax   +44 (0)1392 263686
Email   rkt@exeter.ac.uk
Web   www.exeter.ac.uk/reserch/rkt

15 June 2012

Ben Ryan
Senior Evaluation Manager
EPSRC
Polaris House
North Star Avenue
SWINDON, SN2 1ET

Dear Ben

EPSRC Research Data Policy Framework

Further to my email on 31 May 2012 advising that the University had formally approved our roadmap setting out how we will ensure compliance with your expectations on research data management by 1 May 2015, I am writing to confirm that the following measures are currently being undertaken to ensure sustainable compliance with your (and other funders') policies:

- The University of Exeter has an 'Open Exeter' project – a team of staff dedicated to ensuring that data and publications are made available openly. Open Exeter is developing the skills, training, policy, governance and cost recovery models, informed by a Data Asset Framework [DAF] survey to audit the current research data management practice and complex storage behaviour within our active research community. The DAF has been noted by the DCC as an example of best practice, and will be supplemented by detailed interviews for a full understanding of researchers' processes enabling tailored training and guidance for data management and storage to be developed.

- We have taken care to build a high level of understanding of the importance of these issues at an institutional level and this has resulted in allocation of £800k from our Science Strategy Fund to establish a 'deep archive' for research data, known as the Exeter Data Archive [EDA].

- Successful bidding into JISC funding delivered £250k for the Open Exeter project to embed the EDA and ensure that appropriate sharing [based on ethical, legal and commercial considerations] of research data is becoming routine practice.

- Coordinated working is happening across a wide range of professional services and academic staff, underpinned by robust internal governance procedures with strong institutional support led by the DVC for Research. This provides continuing leadership and coordination of data management and Open Access issues and a framework for the development and ratification of an institutional research data management policy and associated procedures.

I trust that these measures will meet with approval, and that you will feel free to contact me should you require any further clarification.

Yours sincerely

S.N. Fielding
<u>Director of Research and Knowledge Transfer</u>