University of Edinburgh EPSRC RDM Expectations Awareness Survey Findings

March 2015

Stuart Macdonald

Cuna Ekmekcioglu

Evamaria Krause

BACKGROUND

In alignment with the core RCUK Common Principles on Data Policy[1] the Engineering and Physical Sciences Research Council (EPSRC) set out their Research Data Management (RDM) expectations (see Appendix 1) for institutions in receipt of EPSRC grant funding in April 2011. The expectations recognise that a range of institutional policies and practices should satisfy these expectations and as such encouraged research organisations to develop specific approaches which are appropriate to their own structures and cultures through the development of an institutional 'Roadmap'.

In order to comply with EPSRC expectations and to implement the University's RDM Policy, the University has invested significantly in RDM services, infrastructure (incl. storage and security) and support as detailed in the University of Edinburgh's RDM Roadmap[2].

EPSRC assessment of compliance with these expectations will begin on 1 May 2015 for research outputs published on or after that date.

In an effort to gauge the University of Edinburgh's 'readiness' in relation to EPSRC's RDM expectations, Information Services conducted a short survey (see Appendix 2) of existing EPSRC grant holders. This was based in part on EPSRC's own survey of 2526 research grant and fellowship holders in January 2014.

This survey aimed to find out more about grant holder awareness of those expectations as detailed in the EPSRC Policy Framework on Research Data[3] and to learn more about their understanding of the University's RDM Policy, and the RDM Services offered by Information Services in order to ensure that service development meets with compliance requirements.

The survey also asked for EPSRC grant holders who would be willing to undertake a follow-up interview to talk through these issues in more detail to help shape the development the RDM services at the University of Edinburgh. The EPSRC grant holders' follow-up interview questionnaire is attached as Appendix 3, the outcome of which will be reported separately.

SURVEY

The University of Edinburgh EPSRC RDM Expectations Awareness Survey[4] was produced by Dr Cuna Ekmeckioglu, Evamaria Krause and Stuart Macdonald from Research and Learning Services, Library & University Collections, and delivered using Bristol Online Survey tool. It comprised 13 non-mandatory questions and ran from 3 March until 28 March 2015. Respondents were informed that all data collected in the survey will be held anonymously and securely and that data will be reviewed within this institution only, and aggregated and published anonymously.

Email addresses of EPSRC grant holders were obtained from Edinburgh Research and Innovation (ERI) and an initial email detailing the survey background and a link to the survey was sent out to those grant holders on 3 March. A follow up email was sent on 12 March. The email was also

---

[1] http://www.rcuk.ac.uk/research/datapolicy/
[2] http://www.ed.ac.uk/polopoly_fs/1.116873!/fileManager/UoE-RDM-Roadmap-140106.pdf
[3] http://www.epsrc.ac.uk/about/standards/researchdata/
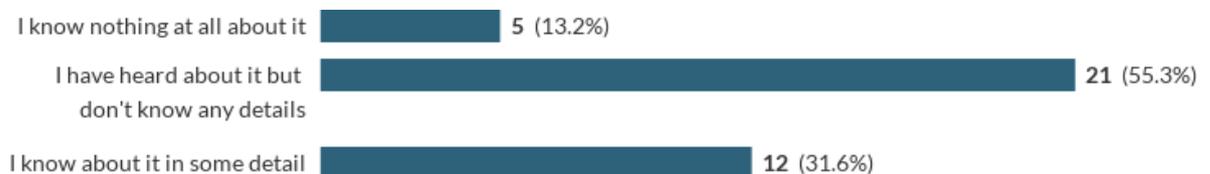[4] https://edinburgh.onlinesurveys.ac.uk/epsrc-expectations-awareness

forwarded to the research administrators of the seven[5] Schools from the College of Science and Engineering, most likely recipients of EPSRC grant funding, for onward dissemination. 98 EPSRC grant holders were identified (these include recent awards and those in receipt of award that go beyond the 1 May 2015) of which 38 responded yielding a response rate of 38.8%.
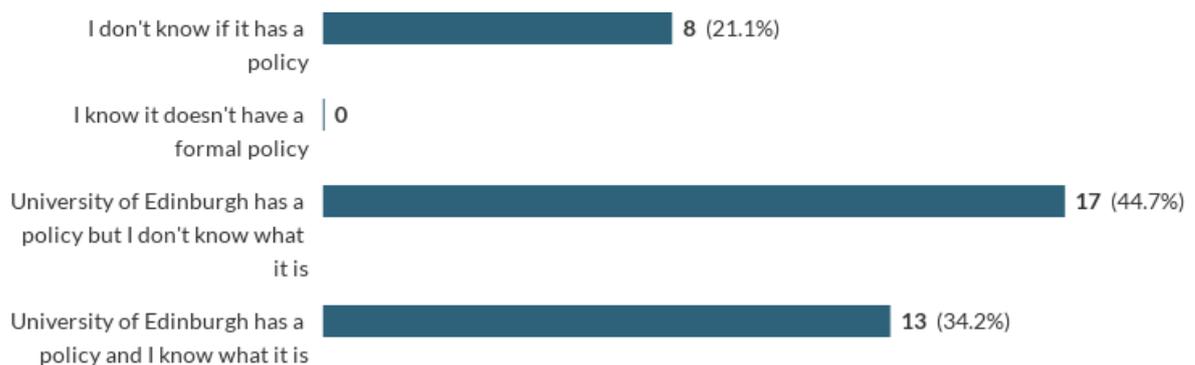
Provisional plans to conduct a survey of EPSRC-funded studentship and Chancellors Fellowship recipients are being discussed.

RESULTS

Q1. Which statement best describes your awareness of EPSRC's expectations on research data?

I know nothing at all about it — 5 (13.2%)
I have heard about it but don't know any details — 21 (55.3%)
I know about it in some detail — 12 (31.6%)

Q2. Which statement below best describes your awareness of University of Edinburgh policy on research data management (http://www.ed.ac.uk/is/research-data-policy)?

I don't know if it has a policy — 8 (21.1%)
I know it doesn't have a formal policy — 0
University of Edinburgh has a policy but I don't know what it is — 17 (44.7%)
University of Edinburgh has a policy and I know what it is — 13 (34.2%)

---

[5] Schools of: Biological Sciences; Chemistry; Engineering; Geosciences; Informatics; Mathematics; Physics and Astronomy

Q3. Are you aware of the following University of Edinburgh (RDM) tools and Services?

Q3.1 DataStore, the new 0.5TB file store

Response	Count
Yes, used	3 (7.9%)
Yes, but not used	15 (39.5%)
Not aware	20 (52.6%)

Q3.2 DMPonline, web-based data management planning tool

Response	Count
Yes, used	2 (5.3%)
Yes, but not used	5 (13.2%)
Not aware	31 (81.6%)

Q3.3 DataShare, University's open access data repository

Response	Count
Yes, used	3 (7.9%)
Yes, but not used	14 (36.8%)
Not aware	21 (55.3%)

Q3.4 MANTRA - RDM training online course

Response	Count
Yes, used	2 (5.3%)
Yes, but not used	4 (10.5%)
Not aware	32 (84.2%)

Q3.5 RDM training workshops and awareness-raising sessions

Response	Count
Yes, used	1 (2.6%)
Yes, but not used	9 (23.7%)
Not aware	28 (73.7%)

Q3.6 Data Asset Register (DAR) using PURE as a catalogue to collect description metadata about research data

Yes, used	5 (13.2%)
Yes, but not used	10 (26.3%)
Not aware	23 (60.5%)

Q4. Do you link to your research data from your publications?

Yes	11 (28.9%)
No	27 (71.1%)

Q5. Do you create or use non-digital data?

Yes	5 (13.2%)
No	33 (86.8%)

Q5a. If you answered 'Yes' will you be able to make your non-digital data available for sharing in the event of a valid request for access to the data?

Yes	4 (80%)
No	1 (20%)
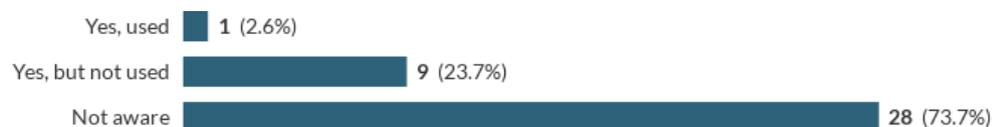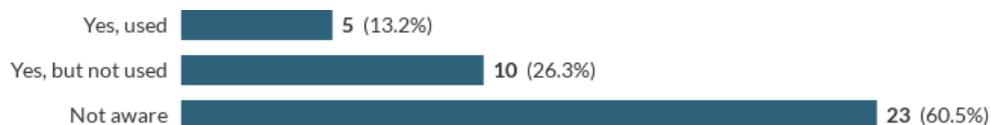
Q6. Do you create structured metadata describing your research data?

Yes	4 (10.5%)
No	34 (89.5%)

Q7. Do you work with sensitive data (e.g. commercial data, personal data)?

Yes	13 (34.2%)
Not at the moment but may in future	10 (26.3%)
No	15 (39.5%)

Q8. Do you share your final data output?

Yes	18 (47.4%)
No	20 (52.6%)

Q8a. If you answered 'Yes', how do you share your data?

Formally, by depositing into a repository/archive	9 (50%)
Informally, e.g., by request, via email, etc.	9 (50%)

Q9. Do you think an RDM consultancy service (e.g. help with writing Data Management Plans) for researchers would be useful?

Very useful	8 (21.1%)
Useful	8 (21.1%)
Don't know	20 (52.6%)
Not useful at all	2 (5.3%)
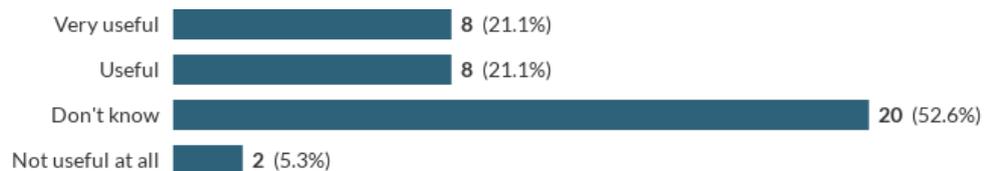
Q13. Would you be willing to take part in a further evaluation event such as an interview or a focus group?

Yes	10 (26.3%)
No	28 (73.7%)

Respondents by School or Research Group (College of Science & Engineering)

School	No. respondents
Informatics	8
Engineering	6
Mathematics	5
Chemistry	4
GeoSciences	3
Physics & Astronomy	3
Biological Sciences	1

Respondents by School or Research Group (other Colleges)

School	No. respondents
Edinburgh Parallel Computing Centre	4
Edinburgh College of Art	2
Social & Political Science	1

Note: *1 respondent withheld information about School or Research Group*

SUMMARY FINDINGS

- Whilst approximately one-third of respondents knew about the EPSRC expectations on research data, over half did not know the detail and 5 respondents (13.2%) knew nothing about it.

  o This demonstrates that further awareness raising and engagement exercises with grant holders are required on the part of University of Edinburgh, EPSRC and RCUK. A funder by funder approach to this may be worth trying.

- Only one-third of respondents knew about the University of Edinburgh's RDM Policy whereas one-fifth knew nothing about the institutional RDM policy.

  o This would suggest that further awareness raising and engagement exercises with grant holders are required on the part of University of Edinburgh in particular institutional and researcher responsibilities.

- There was a lack of awareness of University of Edinburgh RDM Service offerings with 52.6% of respondents unaware of DataStore up to 84.2% of respondents being unaware of MANTRA.

    o Further awareness raising, engagement, and training on the complete suite of RDM Services is required from the University of Edinburgh RDM programme for grant holders and research administrators.

- Approximately two-fifths of respondents were of the opinion that an RDM Consultancy Service (e.g. help with writing Data Management Plans) would be a useful addition to the service offerings.

- Over two-thirds of respondents do not link their data from their publications, whilst only one in ten created structured metadata to describe their research data.

- Over three-fifths of respondents work with or may work with sensitive data in the future. Of those 18 respondents who share their final data output, half do so informally (by request, via email) whereas the other half do so via deposit in an archive or repository.

CONCLUSION

The results of our survey suggest that further awareness raising and engagement exercises are required both on the part of University of Edinburgh, but also EPSRC and RCUK in relation to the management of research data at the institutional level. The University of Edinburgh RDM programme will aim to increase training and awareness raising commitment in relation to the EPSRC expectations over the coming months. It will also aim to feedback any additional findings from follow-up interviews and from tailored awareness raising with EPSRC grant holders and funded researchers to EPSRC in light of their Policy Framework on Research Data and 1 May compliancy date.

APPENDIX 1.

EPSRC expectations concerning the management and provision of access to EPSRC-funded research data

EPSRC has the following clear expectations of organisations in receipt of EPSRC research funding ([clarification and guidance on the interpretations of these expectations (PDF 135KB)](#)):

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](#)).

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 2. EPSRC Expectations Awareness Survey

EPSRC Expectations Awareness

Welcome

Welcome to the EPSRC expectations awareness survey. In April 2011, the Engineering & Physical Sciences Research Council (EPSRC) set out its research data management expectations for institutions in receipt of EPSRC grant funding, which included the development of an institutional 'Roadmap'.

In order to implement the University's Research Data Management (RDM) Policy, the University has invested significantly in research data management services, infrastructure (incl. storage and security) and support as detailed in the University's RDM Roadmap.

This survey aims to find out more about your awareness of those EPSRC expectations concerning the management and provision of access to EPSRC-funded research data as detailed in the EPSRC Policy Framework on Research Data (http://www.epsrc.ac.uk/about/standards/researchdata/). As such we would value your views and contributions relating to your own experiences and understanding of EPSRC's Policy Framework, the University's RDM Policy and RDM Service offerings.

The survey is completed anonymously, and takes around 5 minutes to complete.

Note that once you have clicked on the CONTINUE button at the bottom of each page you can not return to review or amend that page.

Data Protection

All data collected in this survey will be held anonymously and securely. Data will be reviewed within this institution only, and aggregated and published anonymously.

Cookies, personal data stored by the Web browser, are not used in this survey.

About your awareness of EPSRC expectations and RDM tools and services

1 Which statement best describes your awareness of EPSRC's expectations on research data? ✱ *Required*

- ○ I know nothing at all about it
- ○ I have heard about it but don't know any details
- ○ I know about it in some detail

2 Which statement below best describes your awareness of University of Edinburgh policy on research data management (http://www.ed.ac.uk/is/research-data-policy)? ✱ *Required*

- ○ I don't know if it has a policy
- ○ I know it doesn't have a formal policy
- ○ University of Edinburgh has a policy but I don't know what it is
- ○ University of Edinburgh has a policy and I know what it is

3 Are you aware of the following University of Edinburgh (RDM) tools and Services?

	✱ *Required*		
	Yes, used	Yes, but not used	Not aware
DataStore, the new 0.5TB file store	○	○	○
DMPonline, web-based data management planning tool (https://dmponline.dcc.ac.uk/)	○	○	○

DataShare, University's open access data repository (http://datashare.is.ed.ac.uk/)	○	○	○
MANTRA - RDM training online course (http://datalib.edina.ac.uk/mantra/)	○	○	○
RDM training workshops and awareness-raising sessions (http://edin.ac/1usi2JP)	○	○	○
Data Asset Register (DAR) using PURE as a catalogue to collect description metadata about research data	○	○	○

4 Do you link to your research data from your publications? ✱ *Required*

○ Yes   ○ No

5 Do you create or use non-digital data?

○ Yes   ○ No

5.a If you answered 'Yes' will you be able to make your non-digital data available for sharing in the event of a valid request for access to the data?

○ Yes   ○ No

6 Do you create structured metadata describing your research data? ✱ *Required*

○ Yes   ○ No

7 Do you work with senstive data (e.g. commercial data, personal data)? ✳ *Required*

- ○ Yes
- ○ Not at the moment but may in future
- ○ No

8 Do you share your final data output? ✳ *Required*

- ○ Yes
- ○ No

8.a If you answered 'Yes', how do you share your data?

- ○ Formally, by depositing into a repository/archive
- ○ Informally, e.g., by request, via email, etc.

9 Do you think an RDM consultancy service (e.g. help with writing Data Management Plans) for researchers would be useful? ✳ *Required*

- ○ Very useful
- ○ Useful
- ○ Don't know
- ○ Not useful at all

About you

10 Full name

11 Academic role ✱ *Required*

12 School or Research Group ✱ *Required*

13 Would you be willing to take part in a further evaluation event such as an interview or a focus group? ✱ *Required*

  ○ Yes   ○ No

13.a Your email:

Please enter a valid email address.

Thank you!

Thank you for completing this survey. If you have any questions after today, please contact Cuna Ekmekcioglu at cuna.ekmekcioglu@ed.ac.uk.

APPENDIX 3 - EPSRC grant holders interview questions

Overview

1. To start with, could you give me a general overview of the project, its aims and the data related to the project (including ancillary materials)?

Organisation

2. What format(s) are the data in? Can they be converted into open standards formats?
3. What software is used to generate your data? What software is required to use or understand the data?
4. Do you create any documentation or metadata?
5. Where are your files currently stored? How have you backed up your data?

Storage and sharing

6. Expectation V states that metadata for discovery of datasets, i. e. specifying what research data exists, why, when and how it was generated, and how to access it, should be published within 12 months of the data being generated. How realistic is this expectation with regard to the current practice in your research field?
7. Expectation VII states that data have to be preserved for a minimum of ten years after the last request of access. Do you have any concerns regarding this expectation?
8. Have you ever published research data and linked to them from an article you published in an academic journal? How did you do this? (which repository / journal, link via DataCite DOI?)
9. Looking at the 'data sharing matrix' please tell me about the reasons behind your sharing choices.
10. Please describe any conditions or constraints placed on the sharing of this data (e.g. agreements with commercial partners, concerns about intellectual property ownership, secondary use / copyright clearance issues, confidentiality clauses, need for anonymisation, etc.)
11. Do you have concerns about attribution when sharing your data?
12. If you were to share your data, would you want to be able to obtain usage?
13. If you were to share your data, would you like an embargo time period on access? (if different types of data require different embargoes please specify)
14. Are there any services that the University of Edinburgh does not provide that you would have to cost in your grant applications?
15. Thinking about provision at the University of Edinburgh, are there additional support services you would like to see?

About this interview

16. Do you have any thoughts on the relevance of the questions in this interview to your project? (For example, were any inappropriate, or are there other are as you'd like to discuss?)

Data sharing matrix: data types, levels of sharing and data volumes anticipated

Please list the data types from the research project you've chosen for discussion, and mark which level of sharing you think is most appropriate. During our meeting, I'll ask a few questions about why you've selected these options.

List each type of data here (planning documents, raw data, analysis, etc)	Wouldn't share with anyone	Would share only with my collabor-ators	Would share with others in my field	Would share with other academics outside my field	Would share with the general public

	What volume of data do expect to generate during the lifetime of the project?	What volume of data do expect to preserve after the end of the project?
10MB or less		
10MB – 100MB		
100MB – 500MB		
500MB – 1GB		
1GB – 10GB		
10GB – 100GB		
100GB – 500GB		
500GB – 1TB		
1TB – 10TB		
10TB – 100TB		
More than 100TB		