REPORT

The University of Nottingham

UNITED KINGDOM · CHINA · MALAYSIA

Supporting Research Data Management at the University of Nottingham Discussion paper

Author(s):	Caroline Williams, Laurian Williamson & Thomas Parsons
Audience:	Project stakeholders & JISC MRD
Published:	07/05/2013

Contents

Libraries and Research and Learning Resources
Supporting Research Data Management at the University of Nottingham   2012/2013
Page 1 of 14

1. Purpose of the paper

The purpose of this paper is enable discussion and agreement on how best to support Research Data Management (RDM) across RGS (Research Graduate Services), IT Services, and LRLR (Libraries and Research and Learning Resources). It is split into two sections: background and questions for consideration. It starts by articulating assumptions, describing where we are now, summarises the requirements of the funding bodies in relation to RDM, and presents data repositories in a wider context. Then moves on to the questions for us to discuss, they are:

1. What is the demand for support and how best can we predict demand?
2. What are the underlying principles upon which we should build a service model?
3. What are the aspects of RDM for which our Professional Service Departments are responsible?
4. How should we work across RGS, IT Services and LRLR to provide this support/service model?

2. Background

2.1. Assumptions

- That there is role to play for all three Professional Service Departments - RGS, IT Services, and LRLR - in supporting RDM.
- That all three Professional Service Departments need to agree a way of working together in support of RDM.
- That this agreed way of working will be presented by the Directors of all three Professional Service Departments to the PVC for Research for approval.
- That no additional funds will be made available for this activity from the University, and therefore support has to be delivered within existing staffing, unless an extremely compelling case can be made.

2.2. Where are we now?

- RGS and IT Services have delivered the RDM policy and its approval by Management Board (MB).
- IT Services have delivered enhanced research storage, support for technical aspects of data management plans for research bids, data security and classification policy, a training and communication plan in support of the RDM policy, and the e-workbook initiative.

Libraries and Research and Learning Resources
Supporting Research Data Management at the University of Nottingham   2012/2013
Page 2 of 14

- IT Services are developing a RDM technical infrastructure. Appendix 1 provides the conceptualisation of the managed cloud and research data archive service infrastructure produced by Sero Consulting, LRLR and with IT Services.
- The former IS (through securing external project funding from JISC) established a cross departmental project team and the ADMIRe project, the project team coordinated the development of the RDM website, gathered requirements from researchers, ran and co-ordinated a number of awareness raising events for Professional Service staff, delivered training and support – online MANTRA in Moodle and the RDM website, and proposed RDM Service Models to the February ADMIRe project steering group. The latter was rejected by the Steering Group. (Appendix 2 extracts the timelines for Level 1-3 service models which remain helpful overviews of the sociotechnical infrastructure and options for support provision).
- RGS and LRLR with input from IT Services have planned and organised a RDM senior stakeholder event for early June.
- The ADMIRe project ends at the end of June 2013 and the JISC deliverables will be complete.

2.3. RDM institutional obligations

One of the biggest drivers for good RDM practice is the clear "data" expectation of the EPSRC concerning those in receipt of EPSRC funding shall be compliant with the Policy Framework on Research Data expectations by May 2015. They are listed in Appendix 4.2.

In addition, the RCUK Common Principles on Data Policy states that research funders and institutions have requirements to preserve research data of 'long-term value'. Typically they expect data to be preserved for 10 years or more.

This is not restricted to RCUK, other funding bodies most notably the Wellcome Trust has a policy on data management and sharing[1] which expresses its commitment to ensuring that the outputs of the research it funds, including research data, are managed and used in ways that maximise public benefit.

And, now the University RDM policy[2] places expectations on both the researchers and the institution, and also the Research Ethics policy emphasises promotion and sharing of data.

---

[1] http://www.wellcome.ac.uk/About-us/Policy/Policy-and-position-statements/WTX035043.htm
[2] http://www.nottingham.ac.uk/research/research-data-management/creating-data/policies.aspx

Libraries and Research and Learning Resources
Supporting Research Data Management at the University of Nottingham   2012/2013
Page 3 of 14

2.4. National/subject data centres/repositories

Researchers can and do deposit data in subject repositories, data centres and structured databases. The extent to which national and subject data repositories can support intuitional compliance to funders' policies is difficult to gauge. Some useful information is available on the data deposit page on the RDM website[3], there is information on the type of data repositories and also where to find a subject repository (if it exists) for different disciplines. There is also a table outlining the key national data centres that can be used for depositing data. Some issues to consider:

- Funders may require researchers to deposit their research data in an appropriate public archive. This is important in order to facilitate the validation of results and further work by other researchers.

- Requirements of the data repositories/centres in terms of data quality and format.

- Not all funding bodies have a dedicated subject data centre service to support their researchers. For research that falls outside the remit of the existing subject data centres, the institution where the funded researcher is based is expected to maintain the data research outputs for the long-term.

The ADMIRe UoN RDM survey[4] asked researchers about subject and institutional repositories. When respondents were asked whether or not they would deposit their research data in a public subject/disciplinary repository, only 13% answered that they would. Only 3% did so because they were required to do so, and another 10% said they chose to of their own volition. 41% said they would not do so, but no reasons for this were collected.

The Databib[5] lists for some examples from across the world, as does the Datacite website[6].

In addition, there has been a recent announcement that Dryad[7] is now providing membership/subscription 'data package' options to institutions. This is quite important news, as it would entail off-site storage in a respected and perhaps sustainable data repository - see institutional pricing plans[8].

---

[3] http://www.nottingham.ac.uk/research/research-data-management/data-sharing-and-archiving/depositing-and-archiving.aspx
[4] http://eprints.nottingham.ac.uk/1893/
[5] http://databib.org
[6] http://www.datacite.org/repolist
[7] Dryad is a nonprofit repository for data underlying the international scientific and medical literature. It is funded by the National Science Foundation, US.
[8] http://datadryad.org/pages/pricing

Libraries and Research and Learning Resources
Supporting Research Data Management at the University of Nottingham   2012/2013
Page 4 of 14

3. Questions

3.1. What is the demand for RDM support?

It is difficult to predict demand for RDM support. The information we may wish to take into account is:

- A very quick and simple survey of staff from IT Services and LRLR which indicates that: last year there were 10 DMPs created with the support of IT Services, and 25 queries about research data storage (see Appendix 4).

- If we assume 600 research grant applications each year, then the level of support required could increase significantly following a successful awareness raising campaign.

- The ADMIRe UoN RDM survey[9] of researchers found that the areas which most people requested help (or how they would like help) are:

    o Research Data Management website for guidance and support

    o Greater data storage capacity

    o Data management support when writing a research proposal

    o Help to make better use of your final data sets (e.g. create website to showcase data)

    o Support regarding sensitive data.

However, with low awareness of funding body requirements, there is a need to provide more support than may be highlighted by researchers in this list.

3.2. What are the underlying principles for a service model?

Some suggestions for discussion:

- Provision of a single point of contact for all RDM queries e.g. DMP templates, data interviews, and data sharing statements – aligned with funders' requirements
- Provision of cross Professional Service advice, training and advocacy "what you need to do and when" for researchers
- "Policing" of funder/publisher mandates especially EPSRC and mandatory deposit to data services
- Quality assurance of DMPs and/or sharing best practice
- Hands-on curation (at the data planning stage and during research) or source of advice on data curation
- Metadata creation or quality assurance
- Data publication (including making data open)
- Supporting researcher make data deposits in national and international data centres

---

[9] http://eprints.nottingham.ac.uk/1893/

Libraries and Research and Learning Resources
Supporting Research Data Management at the University of Nottingham   2012/2013
Page 5 of 14

3.3. What are the aspects of RDM?

Below is an attempt to bring together the key areas for which the Professional Services are or could be responsible. In our meeting we may want to add to and refine this list (taking into account discussion of the principles above).

IT Services	RGS	LRLR
Continue to develop the RDM technical infrastructure including metadata store and data capture and transfer infrastructure	Policy and compliance	Metadata expertise
Recognised experts in data security and therefore source for advice and support for researchers in this area	Open Access gold fund management and OA coordination role	Recognised experts in Open Access and research communications
Recognised experts in DMP technical appendices and therefore source for advice and support for researchers in this area	Raising awareness, training and advocacy	Institutional repository – business owner (feeding requirement to Corporate Systems)
Raising awareness, training and advocacy	IPR	Data curation – potential source of advice and support for researchers, and data curation services
Active data management and storage		Raising awareness, training and advocacy
Data archiving and preservation		Data sharing and publication
		Copyright

3.4. How should we work across RGS, IT Services and LRLR to provide this support/service model?
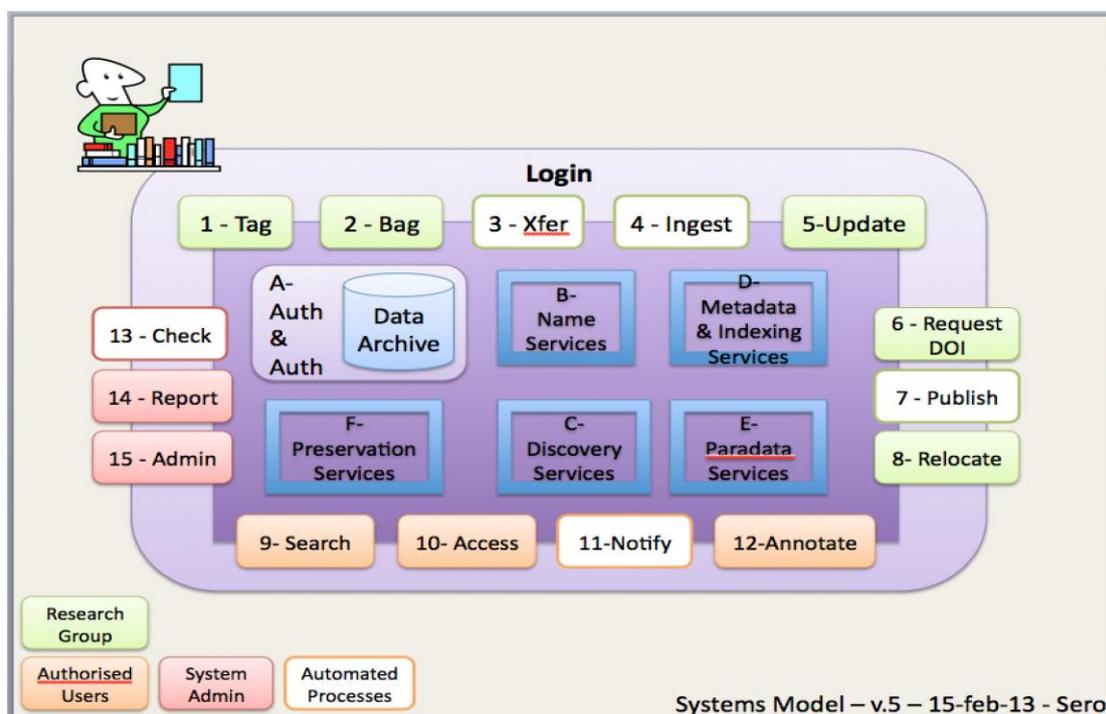
Some suggestions for discussion:
- Do we need a named coordinator/lead from one Professional Service, and if so, which Department should take on that role?

- In addition to the technical infrastructure work, are there any other stands of work that should be started or continued e.g. investigation of shared services or regional/national collaborative approaches, and monitoring demand, and if so, which Department should take on that work?

- Who is best placed to review and maintain the website?

- Who is best placed to be the first point of contact and manage referrals to others?

- Do we need to propose a project to develop an "early life" service?

Libraries and Research and Learning Resources
Supporting Research Data Management at the University of Nottingham   2012/2013
Page 6 of 14

- How do we best communicate and work together across our Services?
- How do we update and maintain training materials and initiatives?

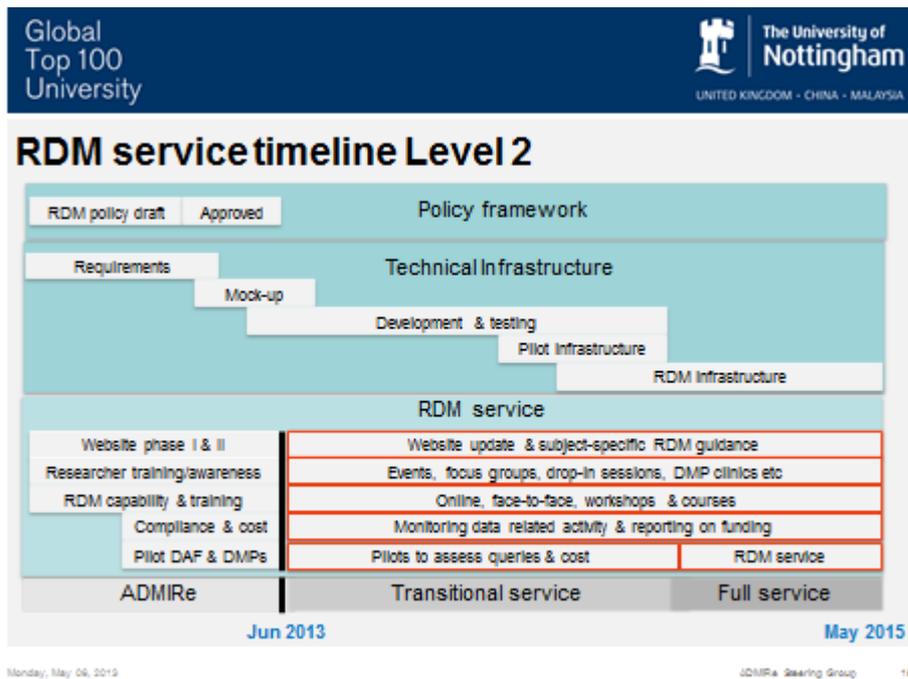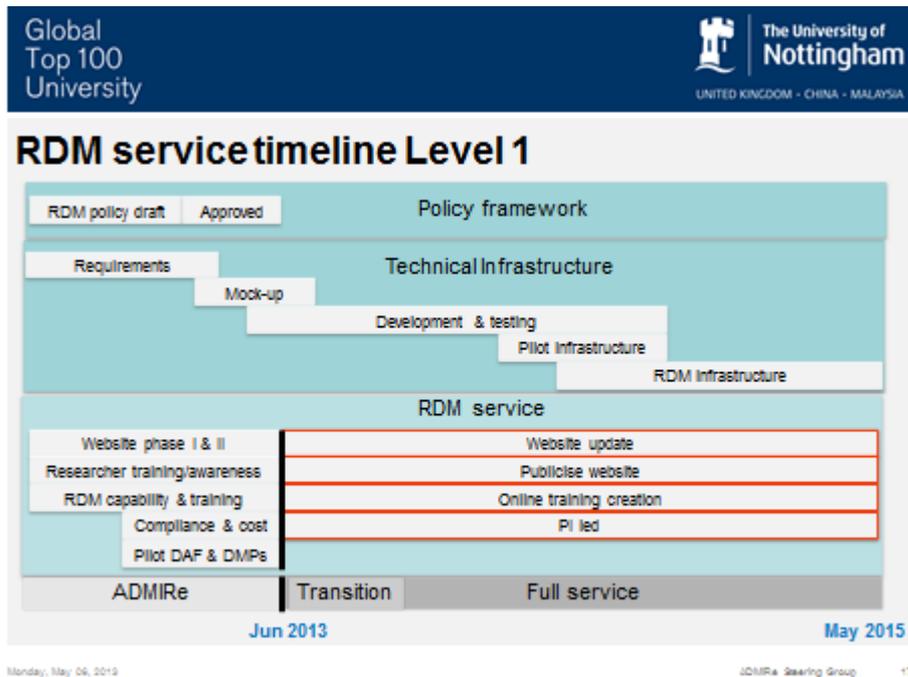Libraries and Research and Learning Resources
Supporting Research Data Management at the University of Nottingham   2012/2013
Page 7 of 14

4. Appendix

4.1. Managed cloud and research data archive service infrastructure (revised)

Function	Actors				Rq	Fr
	R	S	A			
1 – Tag				Enter metadata describing a bag of research data assets	M	M
2 – Bag				Zip the data files up in a bag	C	M
3 – Transfer	+			Transfer a bag to archival storage	C	M
4 – Ingest	+			Ingest a bag in to storage	C	M
5 – Update				Update (enhance, correct) metadata for a stored bag	O	L
6 – GetDOI				Get (public, private) DOIs for designated assets	C	L
7 – Publish	+	+		Publish assets appropriately on landing pages	C	L
8 – Relocate				Relocate assets and update locators	O	L
9 – Search				Search for assets by keyword or field	M	H
10 – Access				Access metadata and data according to permissions	M	M
11 – Notify	+	+	+	Notify actors automatically about data events	O	P
12 – Annotate				Create notes about a bag or its contents	O	L
13 - Check			+	Check (verify) that the contents of a bag are in order	M	P
14 – Report				Run reports on aspects of the system (DOI, bag, user)	O	L
15 - Administer				Administer permissions and system parameters	M	M

Libraries and Research and Learning Resources
Supporting Research Data Management at the University of Nottingham   2012/2013
Page 8 of 14

4.2. Proposed RDM service timelines Level 1-3

Libraries and Research and Learning Resources
Supporting Research Data Management at the University of Nottingham   2012/2013
Page 9 of 14

RDM service timeline Level 3

Libraries and Research and Learning Resources
Supporting Research Data Management at the University of Nottingham   2012/2013
Page 10 of 14
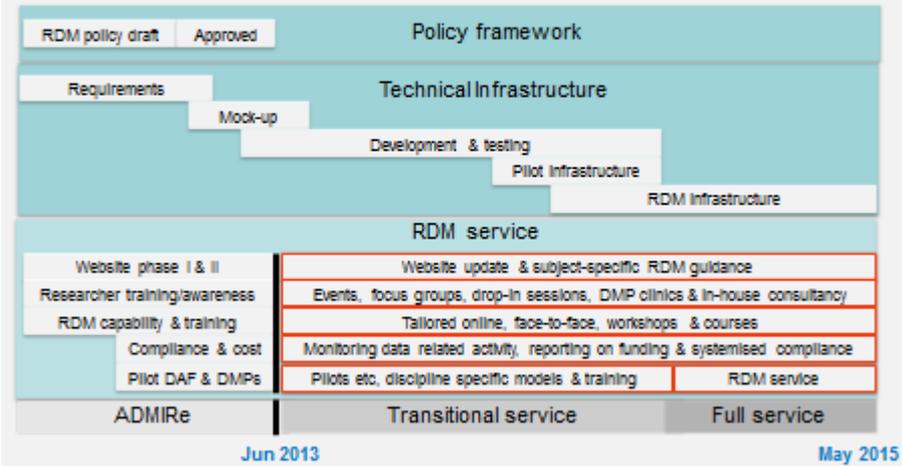
4.3. EPSRC Policy Framework on Research Data Expectations

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

Libraries and Research and Learning Resources
Supporting Research Data Management at the University of Nottingham   2012/2013
Page 11 of 14

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;

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.

Libraries and Research and Learning Resources
Supporting Research Data Management at the University of Nottingham   2012/2013
Page 12 of 14

4.4. Research Data Management queries from researchers survey

Survey overview

Number of respondents: 53

Launch date: 08 Feb 2013

Close date: 18 Feb 2013

Audience: IT Support and LRLR staff

Survey results: These were used by the ADMIRe project team to inform the work on RDM models and institutional capacity for RDM.

Thanks to Chris Middleton, Head of Academic Services and Caroline Gregory, Head of IT support for circulating the survey to their staff.

Question one: in which department are you working?

Question two: what is your job title? [confidential responses]

Question three: have you ever had to deal with a RDM query from a researcher?

Section 1: Department and role

1. In which department are you working?

Libraries and Research & Learning Resources (LRLR):	32.1%	17
IT Services:	67.9%	36

2. What is your job title?

View All Responses - There are too many responses to display on this page and so all the responses to this question are available on a separate page.

Section 2: Research Data Management queries from researchers

3. Have you ever had to deal with a RDM query from a researcher?

Planning a research idea:	n/a	11
Systematic review of the literature/data review:	n/a	5
Funding/Grant bid proposal:	n/a	12
Data management planning:	n/a	12
File organisation:	n/a	21
Local storage:	n/a	27
Data gathering:	n/a	13
Documentation and Metadata:	n/a	8
Data analysis:	n/a	8
Publishing research data:	n/a	6
Repository storage:	n/a	11
Data curation:	n/a	4
Data deposit:	n/a	1
Data discovery:	n/a	7
Accessing and re-using research data:	n/a	7
Never dealt with a RDM query:	n/a	20
Other (*please specify*):	n/a	3

View All Responses - There are too many responses to display on this page and so all the responses to this question are available on a separate page.

Responses for 'other':

- Data Collection website design
- Data recovery and data destruction (erasure)
- Haven't received a direct query but have promoted the topic in training sessions aimed at MPhil/PhD students.

Libraries and Research and Learning Resources
Supporting Research Data Management at the University of Nottingham   2012/2013
Page 13 of 14

Cross-tabulation of results:

Cross Tabulation

Results are cross tabulated by question "1. In which department are you working?"

3. Have you ever had to deal with a RDM query from a researcher?	Libraries and Research & Learning Resources (LRLR)	IT Services	Totals
Planning a research idea	3	8	11
Systematic review of the literature/data review	4	1	5
Funding/Grant bid proposal	2	10	12
Data management planning	1	11	12
File organisation	2	19	21
Local storage	2	25	27
Data gathering	4	9	13
Documentation and Metadata	2	6	8
Data analysis	0	8	8
Publishing research data	1	5	6
Repository storage	0	11	11
Data curation	2	2	4
Data deposit	0	1	1
Data discovery	3	4	7
Accessing and re-using research data	2	5	7
Never dealt with a RDM query	10	10	20
Other	1	2	3
Totals	39	137	176

Libraries and Research and Learning Resources
Supporting Research Data Management at the University of Nottingham   2012/2013
Page 14 of 14