

Research Data Management (RDM) Roadmap

August 2012 – July 2016

Information Services RDM Policy Implementation Committee

University of Edinburgh

September 2015: Version 2.0

Document Status

This is a living document written by the Information Services (IS) Research Data Management (RDM) Action Group and approved by the RDM Steering Committee.

Introduction

The University of Edinburgh is a world-leading centre of academic excellence with a mission for the *creation, dissemination and curation of knowledge*. Information Services, a support group within the University, contributes to the University's mission by striving to provide a *Knowledge Management and Information Service appropriate for supporting and enabling learners, researchers and teachers in a world-class University*.

The University's Research Data Management Policy, passed by the Senate in May 2011, is made up of ten aspirational statements affirming both the researchers' and the University's responsibilities.¹ A research data storage paper was submitted to the University's IT Committee in 2010, with six key recommendations for the University's data infrastructure².

In order to implement the policy and the data storage recommendations, an RDM Policy Implementation Committee was convened in 2012 by the then Vice Principal Knowledge Management and Chief Information Officer, Jeff Haywood. Chaired by John Scally, the then Director of Library and University Collections, its membership had representation across IS and it was charged with delivering services that will met those policy objectives. The Vice Principal also established a Steering Committee led by Professor Peter Clarke from the School of Physics, with membership representing each of the

¹ <http://www.ed.ac.uk/is/research-data-policy>

² <https://www.wiki.ed.ac.uk/download/attachments/146528191/100715-RDSWG-Report.pdf?version=1>

three colleges, IS and Edinburgh Research and Innovation (ERI). Their role is to provide oversight to the activity of the action group, ensuring the services meet the needs of University researchers.

The Steering group is continuing to provide valuable insight into the needs of researchers across the University and its input has been vital in the development of Roadmap 2.0. The implementation of Roadmap 2.0 is the responsibility of the RDM Action Group led by Stuart Lewis, Head of Research and Learning Services. This group contains representatives from across IS each of whom is tasked with delivering separate elements of Roadmap V2.0 whilst ensuring that the entire programme remains integrated.

Development

The Executive Summary of the Information Services Plan, 2012-13 stated³, “Research data management & storage – policies, training, curation, preservation, baseline 0.5Tb/user,” was a major IS-led project for the year. Roadmap 1.2 set out a high level plan for its delivery, noting objectives, outcomes, deliverables and target dates for the 3 year period August 2012-May 2015, across four strategic areas:

- a. **data management planning,**
- b. **active data infrastructure,**
- c. **data stewardship,** and
- d. **data management support.**

The roadmap follows up the business case submitted to the University IT Committee on 3 June, 2012 by Jeff Haywood⁴. Whereas that document estimated a cost of £1M one-off, and £250K recurrent, this Roadmap does not include itemised costs, which are to be agreed as part of the planning process.

The purpose of Roadmap 2.0 is to provide clear information to all University staff on the progress that has been made in delivering the RDM programme for the University and on the areas where work remains to be done or where new tools and services are being evaluated or developed for the benefit of researchers across the University. Researchers and other staff will be able to follow progress through Blog Posts on the Research Data Blog⁵, or by regular updates to the Research Data Management webpage.

³ <http://www.ed.ac.uk/schools-departments/information-services/about/strategy-planning/annual-plans-reports>

⁴ https://www.wiki.ed.ac.uk/download/attachments/146528191/Paper_A_business_case_RDS_RDM_Feb2012_penultimate-1.pdf?version=1

⁵ <http://datablog.is.ed.ac.uk/>

Timeframe

Roadmap 2.0 covers the period from August 2012 when the RDM project began until July 2016. Phases 0, 1, & 2 are now completed and Phases 3 & 4 have been added. In addition a number of services and activities have been marked as Ongoing as there is no end date for the delivery of these.

Phase 0: August 2012 – August 2013: largely a planning phase, with some pilot activity and early deliverables.

Phase 1: September 2013 – May 2014: Initial rollout of primary services.

Phase 2: June 2014 – May 2015: Continued rollout; maturation of services.

Phase 3: September 2015 – December 2015

Phase 4: January 2016 – July 2016

Ongoing: These are activities such as awareness raising and training which are already underway and will continue as a rolling programme available to both new and existing staff.

Authors

The roadmap was developed by members of the IS RDM Action Group: Cuna Ekmekcioglu (L&UC), Sarah Jones (DCC), Stuart Lewis (L&UC), Stuart Macdonald (EDL), Kerry Miller (L&UC), Robin Rice (EDL), Orlando Richards (ITI), Tony Weir (ITI), John Scally (L&UC), Neil Bruce (USD), Marshall Dozier (LAS), Annie Donnelly(LAS), Mike Holmes (ITC).

Data Management Planning				
Support and services for planning activities that are typically performed before research data is collected or created.				
Addresses RDM policy clauses 3, 4.				
Objective	Outcomes	Actions	Deliverables	Target date
1. Tailored DMP assistance for PIs submitting research proposals	Better alignment between IS, ERI and schools	Analyse recent DMPs in research proposals	Set of successful examples of submitted DMPs gathered from ERI and schools	Phase 3
	PIs have access to appropriate, accurate information about IS RDM services for their grant proposals	Trial fast turnaround advisory service with nominated PIs	DMP 'response team' established within IS if deemed necessary	Phase 4
2. Customise DMP Online for optimal UoE use	Clear online service provision for those PIs needing a data management plan	Evaluate DMPonline and analyse user requirements	University branded DMP Online tool	Phase 1 (Completed)
		Refine tool or fully adopt the DCC version, adding UoE specific questions and guidance	Boilerplate text about IS services for use in DMPs	Phase 3 (Underway)
		Advertise to schools that customised DMP templates can be created for their researchers as required	Research support staff and research committees are aware of the potential to customise it for their school.	Phase 3 (Underway)
		Create customised DMPonline templates for schools on request	Respond to demand from schools as it arises	Ongoing
		Metrics and reporting of DMP creation and use	Monthly reports on number of new DMP's created	Ongoing

3. Software Management Planning (SMP)	Investigate need for a Software Management Planning system or process	Work with Software Sustainability Institute (SSI) to identify if an SMP system or process is required	Report on potential of a SMP system identifying next steps	Phase 3
		If required, examine options for an SMP system or process that will fit with existing university RDM systems in collaboration with SSI	Report on the available options with a recommendation of which option(s) to pursue further	Phase 4
4. Embedding RDM questions and guidance within local ethics approval processes	Local ethics processes include RDM issues and provide appropriate guidance	Work with college research administrators to understand existing ethics processes and establish how RDM can be included in these processes	An understanding of the existing ethics processes in each college	Phase 3
		Embed RDM within local ethics processes	Colleges have a template ethics process which includes relevant aspects of RDM	Phase 4

DRAFT

Active Data Infrastructure

The facilities to store active data (data that is actively being used in current research activities), to provide access to that storage through a number of channels, and tools to assist in working with the data.

Addresses RDM policy clauses 5, 8.

Objective	Outcomes	Actions	Deliverables	Target date
5. To provide a globally accessible cross platform file store with sufficient capacity to satisfy majority of researcher use cases	A large common data storage infrastructure using standards compliant technologies with initial access mechanisms available to present this storage, and with automated off-site data backup	Confirm high level requirements for resilience and recoverability of the storage infrastructure (e.g. single or dual site)	Confirmation of high level requirements and initial implementation of DataStore.	Phase 0 (Completed)
		Pilot trial implementation to confirm suitability		Phase 0 (Completed)
		Purchase sufficient infrastructure of appropriate performance and capacity		Phase 0 (Completed)
		Agree allocation and administration processes; Cost service		Phase 1 (Completed)
	Phased implementation of the DataStore across all schools.	Construct file store with initial access mechanisms to support native access from common desktop platforms.	Roll out DataStore on School by School basis	Phase 1 (Completed)
		Review of existing DataStore allocation and deallocation policies	Review current storage quota and allocation policy	Phase 3 (Underway)
			Review deallocation policy	Phase 3 (Underway)
			Implement changes to allocation policy	Phase 4
		Implement deallocation policy	Phase 4	

6. Provide additional data access mechanisms to better support mobile devices and external collaboration	DataStore meets more researcher requirements	Gather requirements on researcher use cases and pilot appropriate technologies.	Additional data access mechanisms to DataStore	Phase 1 (Completed)
		To implement those services which are demonstrated to be needed and for which appropriate technologies are available - DataSync		Phase 2 (Completed)
7. To provide mechanisms to address backup and synchronisation of mobile devices	Ensure recoverability of mobile data.	Confirm technical requirements. If existing mechanisms cannot be used to meet this requirement then cost and develop appropriate service.	Understanding of requirements and, if appropriate, services to ensure mobile data recoverability.	Phase 2 (Completed)
		Rollout DataSync service (promotion, documentation, support)	DataSync made available to researchers	Phase 3
8. Integrate DataStore with the University's Research Computing Infrastructure	Allow easy transfer of data between Datastore and Eddie for processing	Establish mechanisms for quick and easy transfer of data between DataStore and Eddie	Automatic transfer mechanism for users and/or documentation allowing users to do manual transfers	Phase 3
9. Investigate options for an institutional Electronic Lab Notebook (ELN) service	Evaluate potential of ELN's for use across the university	Investigate the requirements for ELNs to see if there is a single system which can meet the needs of projects in many different disciplines	Recommendation of whether or not to adopt a single ELN	Phase 3
		Implement a central ELN system if required	Develop implementation plan	Phase 4
10. Investigate the need for additional source code control systems	Provide additional systems, in addition to Subversion, if required	Investigate the need for, and options to address, source code control systems.	Report on whether additional source code systems are required?	Phase 3
		If appropriate, introduce additional service(s)	Decide if Gitlab or another solution is best and implement selected system	Phase 4
11. Investigate and pilot secure data services	Provide researchers with access to a secure data service which is integrated as appropriate with other systems	Run pilots (Trusted Research Environments (TREs) via virtual machines)	Report on the suitability of the piloted system	Phase 4
		Provide secure data capture and storage service	Select and implement chosen system	Phase 4+

Data Stewardship				
Tools and services to aid in the description, deposit, and ongoing management of completed research data outputs.				
Addresses RDM policy clauses 6, 7, 9, 10.				
Objective	Outcomes	Actions	Deliverables	Target date
12. To develop the data repository for enhanced deposit and discovery of data collections generated by University researchers	Number of new data collections added to the repository	Pilot use of Edinburgh DataShare by 2-3 research groups to identify user requirements	Case studies and use cases based on piloted research groups.	Phase 0 (Completed)
	Metrics show increased use of data collections in repository	Depositor Workflow and usability		Phase 1 (Completed)
		Enhancements made to data repository	Load balancing and software upgrades	Phase 1 (Completed)
			SWORD remote deposit	Phase 1 (Completed)
			End User Interface improvements	Phase 3 (Underway)
			Multimedia Display	Phase 4
13. To provide a registry of research data assets in support of the University RDM policy	Researchers will have a system for recording the location and description of their data assets	Hold focus group(s), interviews and discussions to determine nature of the Data Asset Register (DAR)	Deliverable requirements are determined (Report)	Phase 1 (Completed)
		Analyse results of scoping exercise and agree system specification	System specification	Phase 2 (Completed)
	The University will have a record of its data assets linked to research project information	Rollout out across University	Work with research groups within each school to begin process of populating PURE with dataset details	Phase 3 (Underway)

14. To ensure efficient interoperation between all RDM systems as well as PURE	'Joined up' set of data services serving each stage of the data lifecycle	<p>Investigate opportunities for simplify and merge RDM systems to reduce duplication of effort and simplify processes for researchers</p> <p>Design to incentivise good behaviour, e.g. metadata input, open sharing, designation of data custodians</p> <p>Develop interoperability between systems to allow users to transfer metadata records and datasets between systems as appropriate</p>	<p>Schematic flowchart showing where flows can occur</p> <p>Implement technologies to put optimal flows in place</p> <p>Mechanically joined up set of RDM services</p>	<p>Phase 3 (Underway)</p> <p>Phase 3 (Underway)</p> <p>Phase 4</p>
15. To provide continuity of access for data assets with long-term value	<p>Researchers trust DataShare to keep data safe and accessible over the long term</p> <p>Preservation and access requirements of funded grants are fulfilled</p>	<p>Develop acquisition, appraisal, preservation, retention and disposal strategies and procedures for research data services</p> <p>Promote DataShare as a trusted research data repository for the whole University</p> <p>Develop preservation and access services that fulfil the requirements of research funders</p>	<p>Service definitions include information about risks and guarantees over time</p> <p>Investigate requirements for long-term digital preservation of research data</p> <p>Achieve a recognised trusted repository status (such as Data Seal of Approval)</p> <p>Investigate options for tracking data access dates</p> <p>Implement preservation planning procedures in DataShare</p>	<p>Phase 0 (Completed)</p> <p>Ongoing</p> <p>Phase 2 (Completed)</p> <p>Phase 3 (Underway)</p> <p>Phase 2 (Completed)</p> <p>Phase 4</p>

16. Provide a service to ensure integrity and long term retention of golden copy research data	Develop a DataVault service that will provide long-term, secure, storage of data from DataStore	Requirements gathering to confirm the nature of this service.	Confirmed requirements for data vaulting.	Phase 2 (Completed)
		Develop interim service	Develop service and conduct initial user testing	Phase 3 (Underway)
		Construct and launch interim service university wide.	Interim service available to researchers	Phase 3
		Develop fully featured archival service	Full service built and tested	Phase 4
17. Software preservation investigation	Understand the requirements and scope for software preservation at the UoE	Form expert advisory group of local research software developers	Creation of advisory group	Phase 3
		Investigate and document the requirements and scope for software preservation at the UoE	Scoping report for future software preservation actions	Phase 4
18. Sharing large data investigation	Enable publication and transfer of large datasets	Investigate mechanisms and policies for the publication and transfer of large datasets	Report on available options for enabling large dataset publication and transfer	Phase 3
		Implement mechanisms for the publication and transfer of large datasets	Implement a suitable mechanism for enabling large dataset publication and transfer	Phase 4
19. Take part in Jisc investigations into creating a national Research Data Discovery Service (RDDS)	Local research data descriptions are included in the national RDDS	Populate PURE and DataShare with metadata and datasets respectively	More metadata and datasets in local systems	Ongoing
		Engage with RDDS project	Local data appears within the RDDS	Ongoing

Data Management Support				
General consultancy and support services Addresses RDM policy clauses 1, 2, 4.				
Objective	Outcomes	Actions	Deliverables	Target date
20. Raise awareness of University and funder policies and advocate for good practice in RDM for all research projects	More academic and support staff aware of University and funder policies	Contacting School Research Admins	Awareness raising sessions for different audiences	Phase 1 (Completed)
		Deliver at least 1 awareness raising session per school		Phase 2 (Completed)
		Present session at research committees for each school	Meet with other support groups and committees to gain support for DMP	Phase 1 (Completed)
		Present for other relevant committees, such as - Library Academic Computing, etc.		Phase 2 (Completed)
Ongoing programme of awareness raising presentation for PGR's, and Research Staff	On demand delivery of awareness raising presentations across all schools and research groups	Ongoing		
21. Create a new website for RDM with links to RDM services and RDM how-to guides	Published on IS website; linked from School/College websites across the University	New website for RDM	Updated and streamlined RDM pages linked from the School/College websites	Phase 1 (Completed)
		Update Content in collaboration with Library Academic Support Librarians		Phase 1 (Completed)
	Review and refresh the RDM website to reflect feedback from researchers	Conduct user testing of the existing website. Redesign / redevelop website based on outcomes from testing.	New page on costing RDM in grant proposals	Phase 3
	Form sub-group(s) to consider costing expectations and sources, one group for Infrastructure and one for support services.	Phase 4		
Add RDM costing information for PIs		Phase 4		

22. Maintain, develop and promote online training modules	Remaining MANTRA units written, tested and finalised	Update Software	Update and redevelop MANTRA	Phase 1 (Completed)
		Refresh and redesigned homepage		Phase 1 (Completed)
	Increased usage from within UoE as measured by Google Analytics and feedback from schools	Reach out to more schools to embed MANTRA in their PhD training programmes (with IAD)	Improve take-up of MANTRA within existing participating schools	Ongoing
		Monitor take-up and gather feedback; revise and develop modules accordingly; consider discipline-specific approaches		Ongoing
		Deliver training sessions at schools on request	Face to face sessions with PhD training programmes delivered	Ongoing
			Create and seed JISCmail list	Phase 3 (Underway)
	MANTRA User group to inform future development of MANTRA	Develop a MANTRA User group for trainers from other national and international institutions to exchange ideas	Organise face to face meetings if possible at conferences and other external meetings	Ongoing
	Ensure continued national and international uptake of MANTRA	Deliver RDM MOOC based on MANTRA in collaboration with University of South Carolina Chapel Hill	Completed modules for delivery to learners Publish MOOC Produce regular statistics	Phase 3 Phase 4 Ongoing

23. Create tailored, on-demand training for research groups and professionals	Greater awareness of the importance of RDM across the University	Reach out to a range of groups offering short training workshops	Workshops and training sessions scheduled and delivered to schools and others	Ongoing
	Training trialled which could be delivered to others	Pilot self-directed training with Library Academic Support Librarians	Four trained librarians in first instance and an evaluation with recommendations for further rollout	Phase 0 (Completed)
		Run subsequent course of 4 librarians		Phase 1 (Completed)
		Run course for new librarians or provide refreshers as required	On demand delivery of course	Ongoing
	Develop courses for online delivery	Use Xerte tool to convert courses for online delivery	Courses made available online for self-directed learning	Phase 4
24. Compliance monitoring	Monitor compliance with university and funder RDM policies and requirements	Develop processes for working with ERI and other depts. to provide data on whether PI's and other researchers are complying with relevant policies and requirements	Agree methods of checking compliance	Phase 3
			Periodic reports of awareness and compliance	Phase 4
25. Trial in-depth data management consultancy service	Evidence for demand of consultancy service by PIs and schools/research units	Trial consultancy service with individual researchers and schools/research units.	Meet RDM-related requests for in-depth support	Phase 3 (Underway)
		Determine which IS sections are responsible for delivery and provide a co-ordinated service	Evidence gathering for potential RDM in-depth contracted services for grant-funded projects	Phase 4
		Put in place required services for grants where RDM support is costed into the grant	Meet RDM-related requests for in-depth support	Phase 4
		Contact each new grant holder to offer a short meeting about how we can support their work	Contact new grant holders, using info provide by ERI to offer targeted support for their project	Phase 4

26. Provide day to day RDM support	Researchers have access to timely advice on all aspects of RDM	Provide support service online, by email, by phone, or in person to researchers when required	Researchers provide positive feedback on the accessibility and helpfulness of RDM support activities.	Ongoing
		Promote services to potential users through training, and awareness raising.	Increasing numbers of researchers are aware of the support available and make use of it as required.	Ongoing
	Provide RDM follow-up support during and after research / grant	Contact PI's at grant award and grant completion to offer support with all aspects of RDM	PI's and other researchers are made aware of the support available to them at the grant application stage, during their grant, and at the end.	Phase 3

DRAFT