Research Data Management (RDM) Roadmap August 2012 – January 2014

Information Services RDM Policy Implementation Committee

University of Edinburgh

November 2012: Version 1.0

Document Status

This is a living document of the IS Research Data Management (RDM) Policy Implementation Committee which has been approved by the RDM Steering Committee.

Introduction

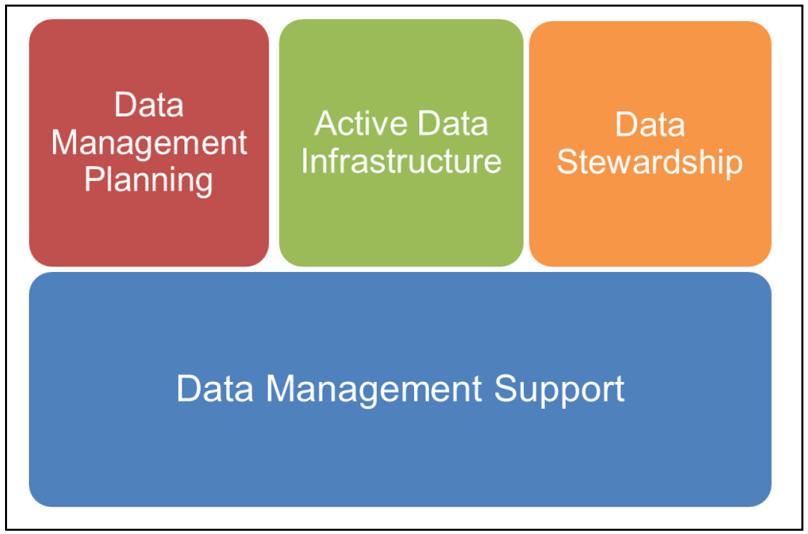
The University of Edinburgh is a world-leading centre of academic excellence with a mission for the *creation, dissemination and curation of knowledge*. Information Services (IS), a support group within the University, contributes to the University's mission by striving to provide a *Knowledge Management 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 infrastructure[2].

In order to implement the policy and the data storage recommendations, an RDM Policy Implementation Committee has been convened by the Vice Principal Knowledge Management and Chief Information Officer, Jeff Haywood. Chaired by John Scally, Director Library and Collections, its membership has representation across IS and it is charged with delivering services that will meet 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 three colleges, IS and Edinburgh Research and Innovation (ERI). Their role is to provide oversight to the activity of the Implementation Committee and its delegated action group, ensuring the services meet the needs of University researchers.

---

[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



Scope

The Executive Summary of the Information Services Plan, 2012-13 states[3], "Research data management & storage – policies, training, curation, preservation, baseline 0.5Tb/user," is a major IS-led project for the year. This roadmap sets out a high level plan for its delivery, noting objectives, outcomes, deliverables and target dates for the 18-month period July 2012-January 2014, across four strategic areas: **data management planning**, **active data infrastructure, data stewardship,** and **data management support**.

The roadmap follows up the business case submitted to the University IT Committee on 3 June, 2012 by Jeff Haywood[4]. 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.

Timeframe

The roadmap takes into account a two year planning horizon but is focused on what can be achieved in the current 18 month period. It is expected that after each phase, the roadmap will be reconstructed to account for further actions and deliverables towards objectives and outcomes in the next phases.

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

Phase 1: February – July 2013: Initial rollout of primary services.

Phase 2: August – January 2014: Continued rollout; maturation of services.

Authors

The roadmap was developed by members of the IS RDM Policy Implementation Committee's Action Group: Cuna Ekmekcioglu (VP's office), Sarah Jones (DCC), Stuart Lewis (L&C), Stuart Macdonald (EDL), Kerry Miller (DCC), Robin Rice (EDL), Orlando Richards (ITI), Tony Weir (ITI).

[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

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.

Responsible: DCC and RDM Action Group

Objective	Outcomes	Actions	Deliverables	Target date
1. Tailored DMP assistance for PIs submitting research proposals	Better alignment between IS, ERI and schools PIs have access to appropriate, accurate information about IS RDM services for their grant proposals	Analyse recent DMPs in research proposals Trial fast turnaround advisory service with nominated PIs	Set of successful examples of submitted DMPs gathered from ERI and schools DMP 'response team' established within IS if deemed required	Phase 0 Phase 2
2. Customise DMP Online for optimal UoE use	Clear online service provision for those PIs needing a data management plan	User requirements analysis Development of tool on own or DCC server.	University branded DMP Online tool Boilerplate text about IS services for use in DMPs	Phase 1 Phase 2

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.

Responsible: IT Infrastructure

Objective	Outcomes	Actions	Deliverables	Target date
3. 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) Pilot trial implementation to confirm suitability Purchase sufficient infrastructure of appropriate performance and capacity Agree allocation and administration processes; Cost service Construct file store with initial access mechanisms to support native access from common desktop platforms: Windows: CIFS, Expandrive (sshfs) Linux/Unix: NFS, sshfs, CIFS Mac: CIFS, Expandrive (sshfs), NFS	Confirmation of high level requirements and initial implementation of file store.	Phase 1

4. Provide additional data access mechanisms to better support mobile devices and external collaboration	File store meets more researcher requirements	Gather requirements on researcher use cases and pilot appropriate technologies. In particular to investigate need for: - Dropbox-like - WebDAV - Andrews File System To implement those services which are demonstrated to be needed and for which appropriate technologies are available	Additional data access mechanisms to file store	Phase 2
5. 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 requriement (eg Dropbox-like, existing data backup mechanisms) then cost and develop appropriate service.	Understanding of requirements and, if appropriate, services to ensure mobile data recoverability.	Phase 2
6. Provide a service to ensure integrity and long term retention of golden copy research data	Data Archive/Data Vault service	Requirements gathering to confirm the nature of this service. Develop, pilot, cost and construct service.	Confirmed requirements for data vaulting. Provision of appropriate service.	Phase 2

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.

Responsible: Data Library and Digital Library

Objective	Outcomes	Actions	Deliverables	Target date
7. 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 Metrics show increased use of data collections in repository	Pilot use of Edinburgh DataShare by 2-3 research groups to identify user requirements Develop repository according to user requirements using available software and protocols where possible.	Case studies and use cases based on piloted research groups Enhancements made to data repository	Phase 0 Phase 1
8. 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 The University will have a record of its data assets linked to research project information	Scoping exercise to determine capture and maintenance mechanisms, software, standards, metadata, usability Analyse results of scoping exercise and agree system specification	Scoping report Data asset registry implementation Registry is populated with a subset of current projects	Phase 0 Phase 1 Phase 2

Objective	Outcomes	Actions	Deliverables	Target date
9. 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	Investigate opportunities for potential automation of data and metadata flows to reduce duplication of effort Design to incentivise good behaviour, e.g. metadata input, open sharing, designation of data custodians	Schematic flowchart showing where flows can occur Implement technologies to put optimal flows in place	Phase 1 Phase 2
10. To provide continuity of access for data assets with long-term value	Researchers trust University RDM services to keep data safe according to service policies Preservation roles and responsibilities are known within IS and acted upon Preservation and access requirements of funded grants are fulfilled	Develop acquisition, appraisal, preservation, retention and disposal strategies and procedures for research data services Develop preservation and access services that fulfil the requirements of research funders	Service definitions include information about risks and guarantees over time Achieve a recognised trusted repository status (such as Data Seal of Approval) Investigate options for tracking data access dates Investigate requirements for long-term digital preservation of research data	Phase 1 Phase 2 Phase 1 Phase 2

Data Management Support

General consultancy and support services

Addresses RDM policy clauses 1, 2, 4.

Responsible: RDM Action group

Objective	Outcomes	Actions	Deliverables	Target date
11. Raise awareness of University and funder policies and advocate for the use of data management plans for all research projects	More academic and support staff aware of University and funder policies	Awareness raising sessions for different audiences Meet with other support groups and committees to gain buy-in for data management planning	A number of scheduled events and meetings; tailored presentations	Phase 0, 1, 2
12. Create and revise IS data management guidance	Published on IS website; linked from appropriate pages across University website	Revise web pages for researchers Form sub-group to consider costing expectations and sources. Add RDM costing information for PIs	Updated and streamlined guidance pages New page on costing RDM in grant proposals	Phase 0 Phase 2
13. Maintain, develop and promote online training modules	Remaining MANTRA units written, tested and finalised 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) Monitor take-up and gather feedback; revise and develop modules accordingly; consider discipline-specific approaches	Improve take-up of MANTRA within existing participating schools Face to face sessions with PhD training programmes delivered	Phase 0, 1 Phase 0, 1, 2

Objective	Outcomes	Actions	Deliverables	Target date
14. Create tailored, on-demand training for research groups and professionals	Greater awareness of the importance of RDM across the University Training trialled which could be delivered to others	Reach out to a range of groups offering short training workshops Pilot self-directed training with liaison librarians	Workshops and training sessions scheduled and delivered to schools and others Four trained librarians in first instance and an evaluation with recommendations for further rollout	Phase 0, 1, 2 Phase 1
15. 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. Determine which IS sections are responsible for delivery and provide a co-ordinated service	Meet RDM-related requests for in-depth support Evidence gathering for potential RDM in-depth contracted services for grant-funded projects	Phase 1 Phase 2