



Facilitate Open Science Training for European Research

WORKSHOP 1:
Open Research Data in Social Sciences and Humanities (ADP)

June 18th 2014, Social Science Data Archives, Faculty of Social Sciences,
University of Ljubljana





DATA LIBRARIAN EXPERIENCE: RDM PROGRAMME @ EDINBURGH

Stuart Macdonald
EDINA & Data Library
University of Edinburgh

stuart.macdonald@ed.ac.uk

RDM and Open Data , Slovenian Social Sciences Data Archive,
University of Ljubljana, Slovenia, 18 June, 2014



- Background
 - EDINA & Data Library
- Defining research data
- Research data management
 - benefits, drivers, funders
- RDM programme
 - policy and implementation
- RDM services and support
 - RDM planning
 - data stewardship
 - awareness raising and training
- JISC

BACKGROUND

- EDINA and University Data Library (EDL) together are a division within Information Services (IS) of the University of Edinburgh.
- EDINA is a Jisc-funded National Data Centre providing national online resources for education and research.
- The Data Library assists Edinburgh University users in the discovery, access, use and management of research datasets.

Data Library Services: <http://www.ed.ac.uk/is/data-library>

EDINA: <http://edina.ac.uk/>



EDINA

- Mission statement: *“We develop and deliver online services and digital infrastructure for UK research and education drawing upon knowledge and expertise gained through research, innovation and development.”*
- Networked access to a range of online resources for UK FE and HE
- Services free at the point of use for staff and students in learning, teaching and research through institutional subscription
- Focus on service but also undertake R&D (projects ➤ services)
 - delivers about 20 online services
 - 5 - 8 major projects (incl. services in development)
 - employs about 80 staff (Edinburgh & St Helens)

DATA LIBRARY

- finding...
- accessing ...
- using ...
- troubleshooting ...
- managing ...

Primarily supporting research in the social sciences but not exclusively so

Building relationships with researchers via postgraduate teaching activities, research support projects, IS Skills workshops, Research Data Management training and through traditional reference interviews.

DEFINING RESEARCH DATA

- Research data are collected, observed or created, for the purposes of analysis to produce and validate original research results.
- Research data can be generated for different purposes and through different processes in a multitude of digital formats
- Both analogue and digital materials are ‘data’.
- Digital data can be:
 - created in a digital form ('born digital')
 - converted to a digital form (digitised)

TYPES OF RESEARCH DATA

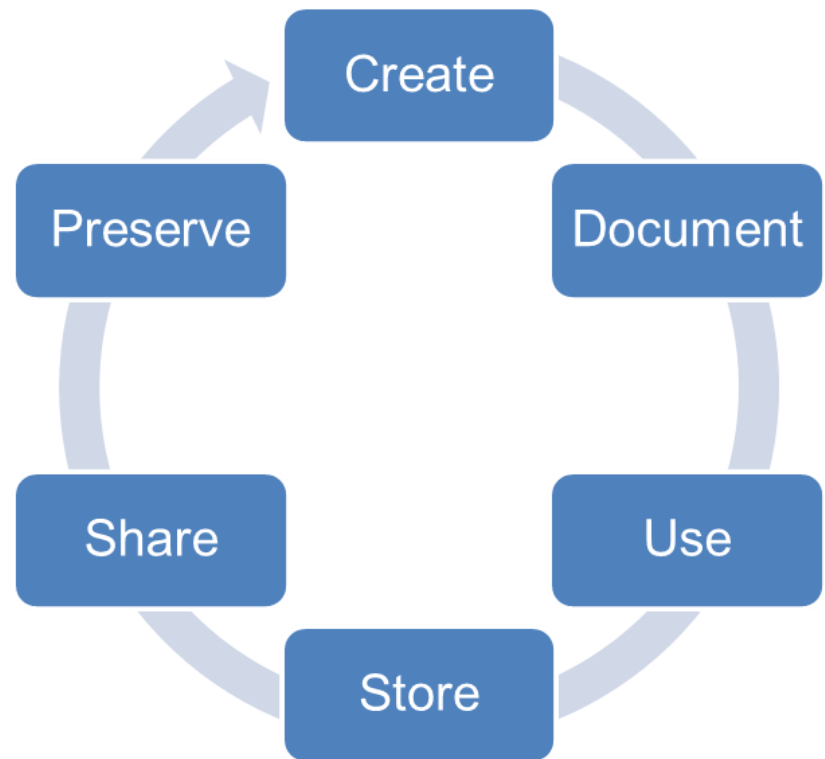
- Instrument measurements
- Experimental observations
- Still images, video and audio
- Text documents, spreadsheets, databases
- Quantitative data (e.g. household survey data)
- Survey results & interview transcripts
- Simulation data, models & software
- Slides, artefacts, specimens, samples
- Sketches, diaries, lab notebooks ...

RESEARCH DATA MANAGEMENT

- Research data management is caring for, facilitating access to, preserving and adding value to research data throughout their lifecycle.
- Data management is one of the essential areas of responsible conduct of research.
- It provides a framework that supports researchers and their data throughout the course of their research and beyond.

ACTIVITIES INVOLVED IN RDM

- Data Management Planning
- Creating data
- Documenting data
- Accessing / using data
- Storage and backup
- Sharing data
- Preserving data



BENEFITS

Managing your data means that you will:

- Meet funder / university / industry requirements.
- Ensure data are accurate, complete, authentic and reliable - as per good research practice.
- Ensure research integrity and replication.
- Enhance data security & minimise the risk of loss.
- Protect important IPR.
- Increase efficiency - save time & resources.
- Increase impact by sharing data (increase in citations 9 - 30% : [Piwowar & Vision 2013](#))

DRIVERS



Organisation for Economic Co-operation and Development describes data as a public good that should be made available

Research Councils UK in its 2009 code of good research conduct says data should be preserved and accessible for 10 years +



Research Funder data policies increasingly demanding of institutional commitment and provisions...

COMMON THEMES ACROSS FUNDING BODIES

- What data will be created? (format, types, volumes etc)
- What standards and methodologies will you use?
- How will ethics and Intellectual Property be managed? (highlight any restrictions on data sharing e.g. embargoes, confidentiality)
- What are the plans for data sharing and access?
- What is the strategy for long-term preservation?

RDM PROGRAMME @ EDINBURGH

- an institutional approach

Edinburgh Data Audit Framework (DAF) Implementation Project
(May - Dec 2008)

A JISC-funded pilot project produced 6 case studies from research units across the University in identifying research data assets and assessing their management, using DAF methodology developed by the Digital Curation Centre.

2 main outcomes:

- Develop university research data management policy
- Develop services & support for RDM (in partnership IS)

DAF Implementation Project: <http://ie-repository.jisc.ac.uk/283/>

UNIVERSITY OF EDINBURGH RDM POLICY

- University of Edinburgh is one of the first Universities in UK to adopt a policy for managing research data:

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

- The policy was approved by the University Court on 16 May 2011.
- It's acknowledged that this is an aspirational policy and that implementation will take some years.

Research Data Management Policy

This policy for managing research data was approved by the University Court on 16 May, 2011.

The University adopts the following policy on Research Data Management. It is acknowledged that this is an aspirational policy, and that implementation will take some years.

1. Research data will be managed to the highest standards throughout the research data lifecycle as part of the University's commitment to research excellence.
2. Responsibility for research data management through a sound research data management plan during any research project or programme lies primarily with Principal Investigators (PIs).
3. All new research proposals [from date of adoption] must include research data management plans or protocols that explicitly address data capture, management, integrity, confidentiality, retention, sharing and publication.
4. The University will provide training, support, advice and where appropriate guidelines and templates for the research data management and research data management plans.
5. The University will provide mechanisms and services for storage, backup, registration, deposit and retention of research data assets in support of current and future access, during and after completion of research projects.
6. Any data which is retained elsewhere, for example in an international data service or domain repository should be registered with the University.
7. Research data management plans must ensure that research data are available for access and re-use where appropriate and under appropriate safeguards.
8. The legitimate interests of the subjects of research data must be protected.
9. Research data of future historical interest, and all research data that represent records of the University, including data that substantiate research findings, will be offered and assessed for deposit and retention in an appropriate national or international data service or domain repository, or a University repository.
10. Exclusive rights to reuse or publish research data should not be handed over to commercial publishers or agents without retaining the rights to make the data openly available for re-use, unless this is a condition of funding.

POLICY IMPLEMENTATION

RDM Programme in 3 phases:

- **Phase 0:** August 2012 - August 2013: 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.

Full details of the programme is available at:

<http://edin.ac/1eE3sav>

COMMITTEES

An RDM Policy Implementation Committee was set up by the Vice Principal Knowledge Management charged with delivering services that will meet RDM policy objectives:

- Membership from across IS
- Iterate with researchers to ensure services meet the needs of researchers

The Vice Principal also established a Steering Committee led by Prof. Peter Clarke with members of Research Committee from the 3 colleges, IS, DCC and Edinburgh Research and Innovation (ERI).

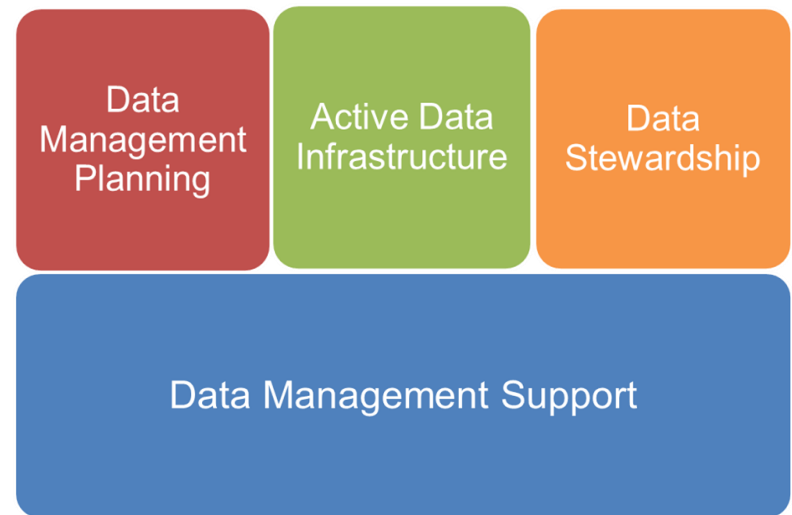
Their role is to:

- Provide oversight to the activity of the Implementation Committee
- Ensure services meet researcher requirements without harming research competitiveness

RDM SERVICES AND SUPPORT

- Services already in place:
 - Data management planning
 - Active working file space = DataStore
 - Data publication repository = DataShare
- Services in development:
 - Long term data archive = DataVault
 - Data Asset Register (DAR)
- RDM support: Awareness raising, training & consultancy

RDM Roadmap



<http://edin.ac/1u3sKqy>

RESEARCH DATA MANAGEMENT PLANNING

Support and services for planning activities that are performed at the conceptual stage before research data are collected or created

- Tailored DMP assistance for researchers submitting research proposals
- Customised instance of *DMPonline* toolkit for University of Edinburgh use

WHAT IS A DATA MANAGEMENT PLAN (DMP)?

DMPs are written at the start of a project to define:

- What data will be collected or created?
- How the data will be documented and described?
- Where the data will be stored?
- Who will be responsible for data security and backup?
- Which data will be shared and/or preserved?
- How the data will be shared and with whom?

DMP SUPPORT

- Academic Support Librarians have received RDM training, including training on writing Data Management Plans.
- CHSS Research Administrators staff received training in May to provide support at the grant application stage. This will be extended to the Research Administrators in CSE & MVM.
- ERI staff will be receiving RDM training.
- Tailored DMP courses for research staff and PGRs are being delivered.
- MANTRA also has a module on DMP for self-paced learning.
- General DMP queries can be sent to the IS Helpline who will direct them as appropriate.

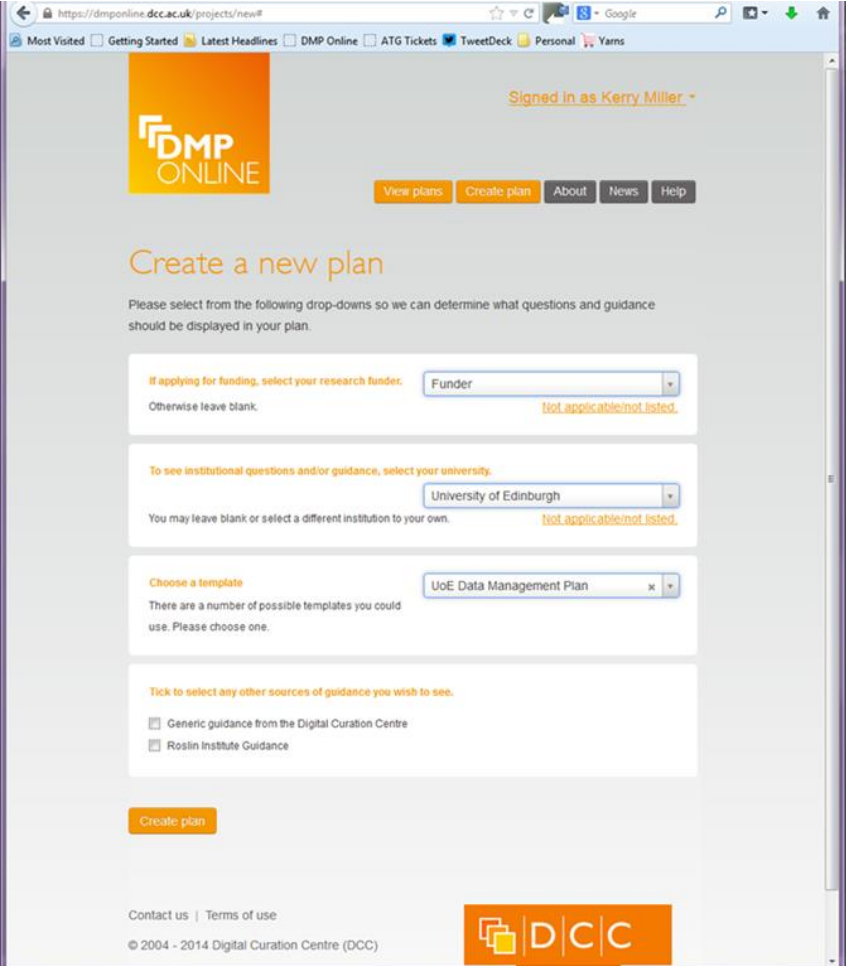
DMP ONLINE TOOLKIT

Free and open web-based tool to help researchers write plans:

<https://dmponline.dcc.ac.uk/>

It features:

- Templates based on different requirements
- Tailored guidance (disciplinary, funder etc.)
- Customised exports to a variety of formats
- Ability to share DMPs with others



The screenshot shows the DMP ONLINE website interface. At the top, there is a navigation bar with the DMP ONLINE logo and links for 'View plans', 'Create plan', 'About', 'News', and 'Help'. The user is signed in as 'Kerry Miller'. The main heading is 'Create a new plan', followed by instructions to select from drop-downs. The form includes three main sections: 1. 'Funder' selection with a dropdown menu and a 'not applicable/not listed' link. 2. 'University' selection with a dropdown menu and a 'not applicable/not listed' link. 3. 'Choose a template' section with a dropdown menu showing 'UoE Data Management Plan'. Below these are checkboxes for 'Generic guidance from the Digital Curation Centre' and 'Roslin Institute Guidance'. A 'Create plan' button is at the bottom of the form. The footer contains contact information and the Digital Curation Centre (DCC) logo.

TEMPLATES AND GUIDANCE

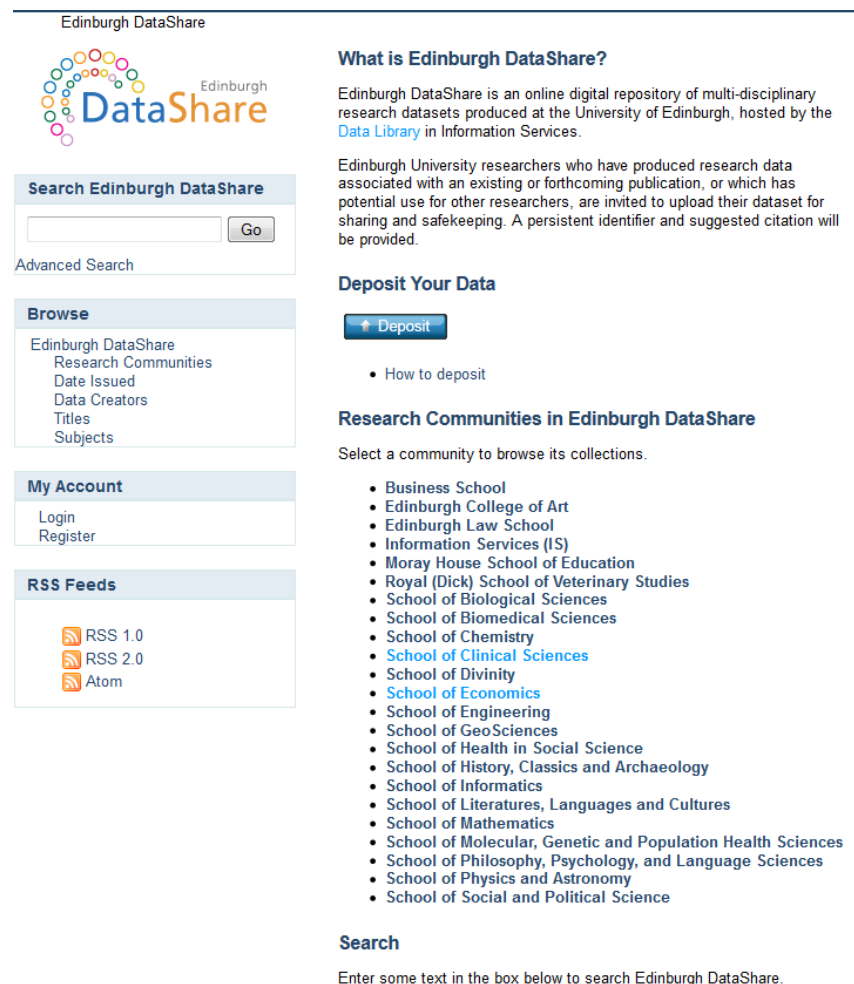
- Edinburgh University Templates and Guidance are still in draft.
- Edinburgh University Guidance is provided for those applying to: AHRC, BBSRC, CRUK, ESRC, MRC, NSF, NERC, STFC, & Wellcome Trust.
- Edinburgh University Templates are available for Researchers and PGRs not applying to any of the above.
- Customised Guidance is given for those working at the Roslin Institute.

DATASTORE

- Facility to store data that are actively used in current research activities
- Provision: 1.6PB storage initially
- 0.5 TB (500GB) per researchers, PGR upwards
- Up to 0.25TB of each allocation can be used to create “shared” group storage
- Cost of extra storage: £200 per TB per year= 1TB primary storage, 10 days online file history, 60 days backup, DR copy
- Early adopters in three Colleges: Geosciences; Philosophy, Psychology, & Language Sciences; Population Health Sciences; Biological Sciences; Clinical Neurosciences; and Social & Political Sciences
- CHSS: all personal spaces rolled out, waiting for information required for construction of group spaces. CHSS IT will roll it out to individual staff and ensure they have the right information to use the space.
- All other schools /research centres in CSE & MVM will follow.

DATASHARE

- Edinburgh DataShare is the University data repository for publishing your research data openly:
<http://datashare.is.ed.ac.uk>
- It will help you disseminate your research, get credit for your data collection efforts, and preserve your data for the long-term.
- It backs up the University Research Data Management policy.
- It can help you comply with funder requirements to preserve and share your data.



The screenshot shows the Edinburgh DataShare website. At the top, it says "Edinburgh DataShare" and features the logo, which consists of a cluster of colorful circles around the text "DataShare". Below the logo is a search bar with a "Go" button. There are four main navigation sections: "Advanced Search", "Browse", "My Account", and "RSS Feeds". The "Browse" section lists categories like "Edinburgh DataShare", "Research Communities", "Date Issued", "Data Creators", "Titles", and "Subjects". The "My Account" section has "Login" and "Register" links. The "RSS Feeds" section lists "RSS 1.0", "RSS 2.0", and "Atom". On the right side, there is a "What is Edinburgh DataShare?" section, a "Deposit Your Data" section with a "Deposit" button and a "How to deposit" link, and a "Research Communities in Edinburgh DataShare" section with a list of 20 different schools and departments.

Edinburgh DataShare

Edinburgh DataShare

Search Edinburgh DataShare

Go

Advanced Search

Browse

- Edinburgh DataShare
- Research Communities
- Date Issued
- Data Creators
- Titles
- Subjects

My Account

- Login
- Register

RSS Feeds

- RSS 1.0
- RSS 2.0
- Atom

What is Edinburgh DataShare?

Edinburgh DataShare is an online digital repository of multi-disciplinary research datasets produced at the University of Edinburgh, hosted by the Data Library in Information Services.

Edinburgh University researchers who have produced research data associated with an existing or forthcoming publication, or which has potential use for other researchers, are invited to upload their dataset for sharing and safekeeping. A persistent identifier and suggested citation will be provided.

Deposit Your Data

Deposit

- How to deposit

Research Communities in Edinburgh DataShare

Select a community to browse its collections.

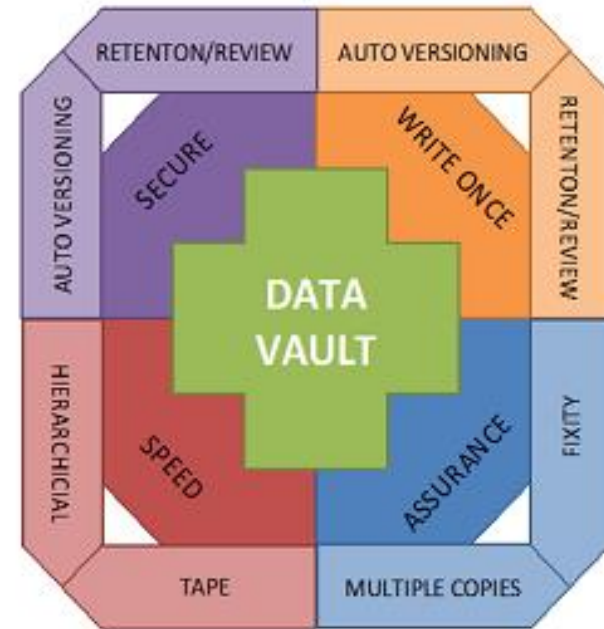
- Business School
- Edinburgh College of Art
- Edinburgh Law School
- Information Services (IS)
- Moray House School of Education
- Royal (Dick) School of Veterinary Studies
- School of Biological Sciences
- School of Biomedical Sciences
- School of Chemistry
- School of Clinical Sciences
- School of Divinity
- School of Economics
- School of Engineering
- School of GeoSciences
- School of Health in Social Science
- School of History, Classics and Archaeology
- School of Informatics
- School of Literatures, Languages and Cultures
- School of Mathematics
- School of Molecular, Genetic and Population Health Sciences
- School of Philosophy, Psychology, and Language Sciences
- School of Physics and Astronomy
- School of Social and Political Science

Search

Enter some text in the box below to search Edinburgh DataShare.

DATA VAULT

- Safe, private, store of data that is only accessible by the data creator or their representative
- Secure storage:
 - File security
 - Storage security
 - Additional security: encryption
- Long term assurance
- Automatic versioning



<http://datablog.is.ed.ac.uk/2013/12/20/thinking-about-a-data-vault>

DATA ASSET REGISTER

- a catalogue of data assets produced by researchers working for the University of Edinburgh,
- will be a key component of the University of Edinburgh Research Data Management (RDM) systems
- will give researchers a single place to record the existence of data assets they have produced so that they can be discovered, accessed, and reused as appropriate.

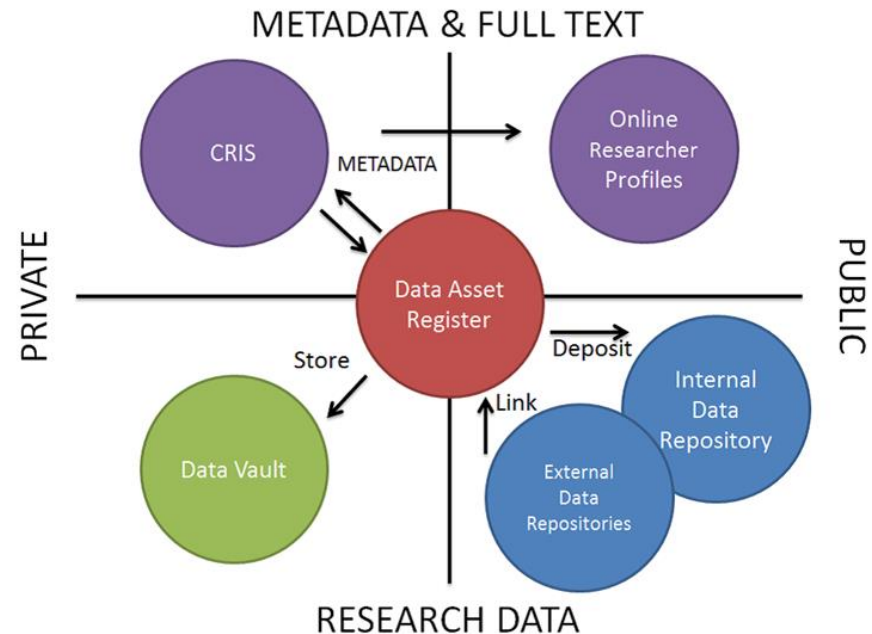
<http://datablog.is.ed.ac.uk/2013/12/12/thinking-about-research-data-asset-registers>

INTEROPERATION

Systems do not live in isolation, and become more powerful and more likely to be used if they are integrated with each other.

However, the last thing that we want is to introduce further systems that need to be fed with duplicate information.

This means interoperation for some or all of the components



RDM SUPPORT

Making the most of local support!

- RDM team will work with the Research Administrators in each School.
- Academic Support Librarians (who represent each of the 22 Schools).
- IT staff in each School.
- ERI staff. They will be receiving RDM training.
- Each School's Ethics Committee
- Bespoke RDM email address or queries can be sent to the Helpline who will direct them as appropriate.

COMMUNICATIONS PLANS

There are a number of different groups within the university and outside with whom we need to communicate our RDM programme.

This will be done through a variety of communication activities.

Target Audiences

1. University of Edinburgh staff need to understand the principles of RDM and how it is practiced and supported within the University:
 - Research active staff
 - IS and School/college support staff
 - Other university committees and groups (research policy group, library committee, IT committee, knowledge strategy committee)
2. External collaborators and stakeholders such as funding bodies, Russell Group, national and international RDM community e.g. RDA, DANS, ANDS, COAR, DPC, DCC

KEY MESSAGES:

Co-ordinated, Consistent, Coherent

There are three key messages which will need to be tailored and made timely and relevant to our target audiences.

The core of each message must be maintained to ensure that everyone gains the same level of understanding.

1. The University is committed to and has invested in RDM
 - services, training, support
2. What is meant by Research Data Management?
 - definitions, data lifecycle, responsibilities
3. The University is supporting researchers
 - encourage good research practice, effect culture change

AWARENESS RAISING

- Introductory sessions on RDM services and support for research active and research admin staff in Schools / Institutes / Research Centres
- Contact Cuna Ekmekcioglu at cuna.ekmekcioglu@ed.ac.uk for a session for your School/Research Centre
- RDM website: <http://www.ed.ac.uk/is/data-management>
- RDM blog: <http://datablog.is.ed.ac.uk>
- RDM wiki: <https://www.wiki.ed.ac.uk/display/RDM/Research+Data+Management+Wiki>

The screenshot shows the University of Edinburgh's Research Data Management website. At the top, there is the university's logo and name, a navigation menu with 'Schools & departments', and a search bar. Below the header is a banner for 'Information Services' with a photo of people at computers. A red navigation bar contains 'Research Data Management' and a 'Contact' link. The main content area features a breadcrumb trail: 'University home > Schools & departments > Information Services > Research Support > Research Data Management > Research Data Management Home'. The central focus is the 'Research Data Management Programme' section, which includes an icon-based graphic, a title, and a brief description. To the right, there are three vertical panels: 'Why is data management important for your research?' with a sub-link 'Why is data management important?', 'Storage & backup solutions' with a sub-link 'Storage & backup solutions', and 'Sharing & publishing data' with a sub-link 'Sharing & publishing data'. At the bottom, there are three more panels: 'DMP ONLINE Data management planning' with a sub-link 'Data management planning', 'RDM training' with a sub-link 'RDM training', and 'Help & support' with a sub-link 'Help & support'. A 'Quick links' section at the bottom right lists 'Research Data MANTRA', 'Research Data Blog', 'PURE', and 'ERI: Research Support and Development'.

<http://www.ed.ac.uk/is/data-management>

TRAINING: MANTRA

- MANTRA is an internationally recognized self-paced online training course developed here for PGR's and early career researchers in data management issues.
- Anyone doing a research project will benefit from at least some part of the training - discrete units
- Data handling exercises with open datasets in 4 analytical packages: R, SPSS, NVivo, ArcGIS



The screenshot shows the MANTRA website interface. At the top, there is a banner with the text "MANTRA Research Data Management Training" and a colorful abstract graphic. Below the banner, a description states: "MANTRA is an online course designed for researchers or others planning to manage digital data as part of the research process." To the right of this text are four icons representing different user roles: "Research Student", "Career Researcher", "Senior Academic", and "Information Professional". Below the description is a navigation menu with the following items: "HOME", "ABOUT", "ACKNOWLEDGEMENTS", "DIY TRAINING KIT FOR LIBRARIANS", and "FEEDBACK". Underneath the navigation menu is a section titled "LEARNING UNITS" which contains a grid of nine buttons, each with an icon and a dropdown arrow. The buttons are: "Research data explained", "File formats & transformation", "Data protection, rights & access", "Data management plans", "Documentation & metadata", "Sharing, preservation & licensing", "Organising data", "Storage & security", and "Software practicals". The "Software practicals" button is highlighted in orange.

<http://datalib.edina.ac.uk/mantra>

TRAINING: TAILORED COURSES

- A range of training programmes on research data management (RDM) in the form of workshops, power sessions, seminars and drop in sessions to help researchers with research data management issues
 - <http://www.ed.ac.uk/schools-departments/information-services/research-support/data-management/rdm-training>
 - Creating a data management plan for your grant application
 - Research Data Management Programme at the University of Edinburgh
 - Good practice in Research Data Management
 - Handling data using SPSS
 - Handling data with ArcGIS
- <http://edin.ac/1kRMPv3>

PROGRESS SO FAR

- ✓ Data Share - Live Now
- ✓ *DMPonline* - Live Now
- ✓ Website - Live Now

- Data Management Planning Support - Aug 2014
- Data Store - Roll-out completed by Dec 2014
- Training - Ongoing
- Awareness Raising - Ongoing

- Data Asset Register - Dec 2014
- Data Vault - Dec 2014

JISC MRD PROGRAMME 2009-2013

- Jisc is ‘a registered charity [.. that ..] champions the use of digital technologies in UK education and research.’ It is funded by all the UK tertiary education funding councils.
- Jisc funded two Managing Research Data (MRD) strands:
 - Strand 1 (Oct. 2009 - Sept. 2011)
 - Piloting RDM infrastructures within institutions
 - Improving practice in RDM planning
 - Developing tools to help institutions plan their RDM practice
 - Demonstrating the benefits of citing, publishing and research data
 - Developing RDM Training materials - up-skilling researchers and support staff

- Strand 2 (Oct. 2011 - Jul. 2013) - Working closely with the DCC this strand aimed to improve institutional RDM capability.
 - 17 institutional projects to help universities pilot or develop RDM infrastructures to provide quality support for research
 - 8 projects for helping research groups or departments fulfill the requirements of research funders by implementing DMPs
 - 2 projects to customize DCC's DMP Online tool for institutional use
- These activities were complemented by work:
 - to develop disciplinary focused RDM training materials for other stakeholders, including discipline liaison librarians and research liaison officers.
 - Managing Research Data Strand 1: <http://www.jisc.ac.uk/whatwedo/programmes/mrd.aspx>
 - Managing Research Data Strand 2: <http://tinyurl.com/6w6g6qx>

THANK YOU!

Acknowledgements:

Dr. Cuna Ekmekcioglu (Vice Principal's Office)

Sarah Jones (Digital Curation Centre)

Stuart Lewis (Research & Learning Services)

Kerry Miller (Research & Learning Services)

Robin Rice (EDINA & Data Library)

Dr. John Scally (Library and Collections)

Tony Weir (IT Infrastructure)

