



---

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



# Introduction to RDM and emerging good RDM practice in UK universities

---

Joy Davidson, *Digital Curation Centre*



Funded by:





# What is research data management and curation?

- Active management, adding value and maintaining access to research data over the course of its scholarly lifecycle.
- Part of good research practice



# UK Context

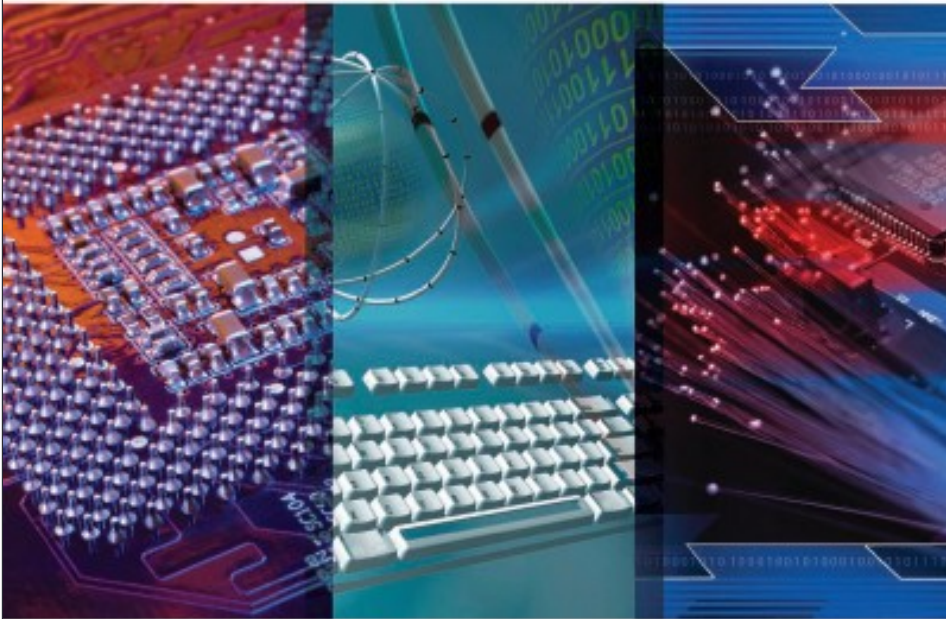


# What has been driving this agenda in the UK?

- Several factors making RDM a key issue for many UK HEIs
- These factors are international and are affecting research practice in many countries (US, Canada, EU)

# RCUK Review of e-Science 2009

BUILDING A UK FOUNDATION FOR THE TRANSFORMATIVE  
ENHANCEMENT OF RESEARCH AND INNOVATION



*“Data sets  
are  
becoming  
the new  
instruments  
of science”*

# Digital data as the new special collections?

Sayed Choudhury, Johns  
Hopkins



Roman de la Rose: Lutenist and singers in a garden  
British Library Harley MS 4425, f.14v  
Copyright © The British Library Board



## Give us back our crown jewels

Our taxes fund the collection of public data - yet we pay again to access it. Make the data freely available to stimulate innovation, argue Charles Arthur and Michael Cross

Charles Arthur and Michael Cross  
The Guardian, Thursday 9 March 2006  
[Article history](#)

Research  
data:  
institutional  
crown  
jewels?



# What do UK funders expect?



# Ultimately funders expect:

- timely release of data
  - once patents are filed or on (acceptance for) publication
- open data sharing
  - minimal or no restrictions if possible
- preservation of data
  - typically 5-10+ years if of long-term value



See the RCUK Common Principles on  
Data Policy:

[www.rcuk.ac.uk/research/Pages/DataPolicy.aspx](http://www.rcuk.ac.uk/research/Pages/DataPolicy.aspx)

● Full Coverage   ● Partial Coverage   ○ No Coverage

Research Funders	Policy Coverage		Policy Stipulations					Support Provided			
	Published outputs	Data	Time limits	Data plan	Access/sharing	Long-term curation	Monitoring	Guidance	Repository	Data centre	Costs
AHRC	●	●	●	●	●	●	○	●	○	●	●
BBSRC	●	●	●	●	●	●	●	●	●	●	●
CRUK	●	●	●	●	●	●	●	●	●	○	○
EPSRC	●	●	●	●	●	●	●	●	○	○	●
ESRC	●	●	●	●	●	●	●	●	●	●	●
MRC	●	●	●	●	●	●	○	●	●	○	●
NERC	●	●	●	●	●	●	●	●	●	●	●
STFC	●	●	●	●	●	●	●	●	●	●	●
Wellcome Trust	●	●	●	●	●	●	●	●	●	●	●

<http://www.dcc.ac.uk/resources/policy-and-legal/overview-funders-data-policies>

## In this section

[Briefing Papers](#)

[How-to Guides](#)

[Developing RDM Services](#)

[Curation Lifecycle Model](#)

[Curation Reference Manual](#)

### Policy and legal

[Overview of funders' data policies](#)

#### Funders' data policies

[AHRC](#)

[BBSRC](#)

[Cancer Research UK](#)

[EPSRC](#)

[ESRC](#)

[► MRC](#)

[NERC](#)

[STFC](#)

[Wellcome Trust](#)

[Institutional data policies](#)

[Policy tools and guidance](#)

## MRC - Medical Research Council

The MRC [position statement](#) supports open and unrestricted access to published research.

It also has a [policy on research data sharing](#). This policy was revised in September 2011.

### POLICY STIPULATIONS

#### Time Limits

The MRC expects publications to be made openly accessible as soon as possible and in any event within six months of the journal publisher's official date of final publication.

Data should be made available in a timely and responsible manner. A limited defined period of exclusive use of data for primary research is reasonable.

Data must be securely maintained for a minimum of ten years after completion of the research

#### Data Plan

Applicants are expected to submit a [data management plan](#) at the grant proposal stage.

## Useful links

[MRC statement in support of open access to published research](#)

[MRC policy on research data-sharing](#)

[MRC guidance on DMPs](#)

[MRC good research practice guide](#)

# Working with UK Universities



# Institutional Engagement programme

- Pilot phase ran from Spring 2011 to Spring 2013
- Funded by HEFCE under Universities Modernisation Fund
- Provided up to 60 days of resource to 21 institutions
- Aimed to:
  - Increase capacity and capability in Research Data Management
  - Drive efficiencies in the HE sector by sharing models and lessons



# The engagement model

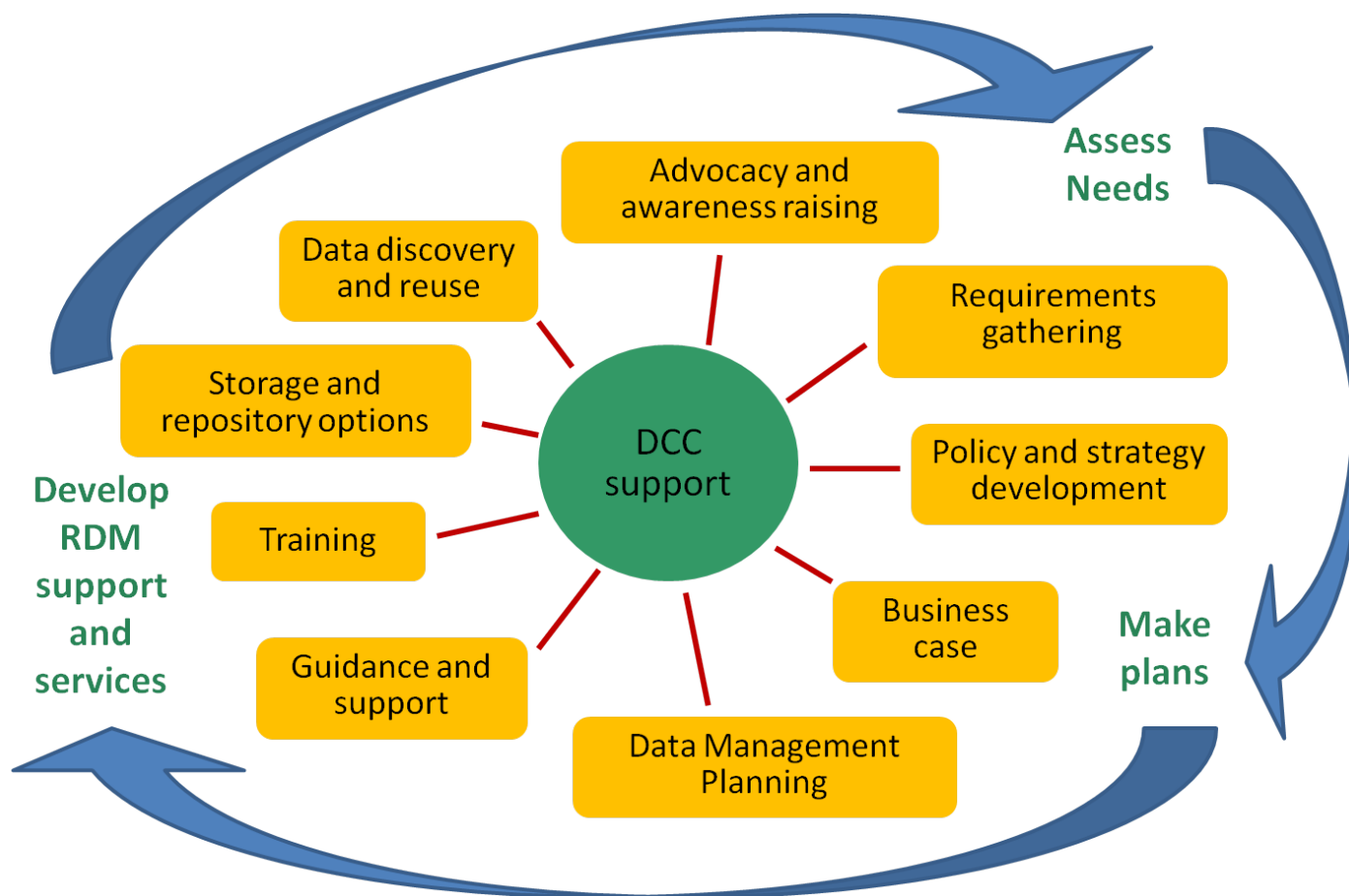
- Efficient information sharing across HE sector
- Redesign of DCC support



# Who we worked with

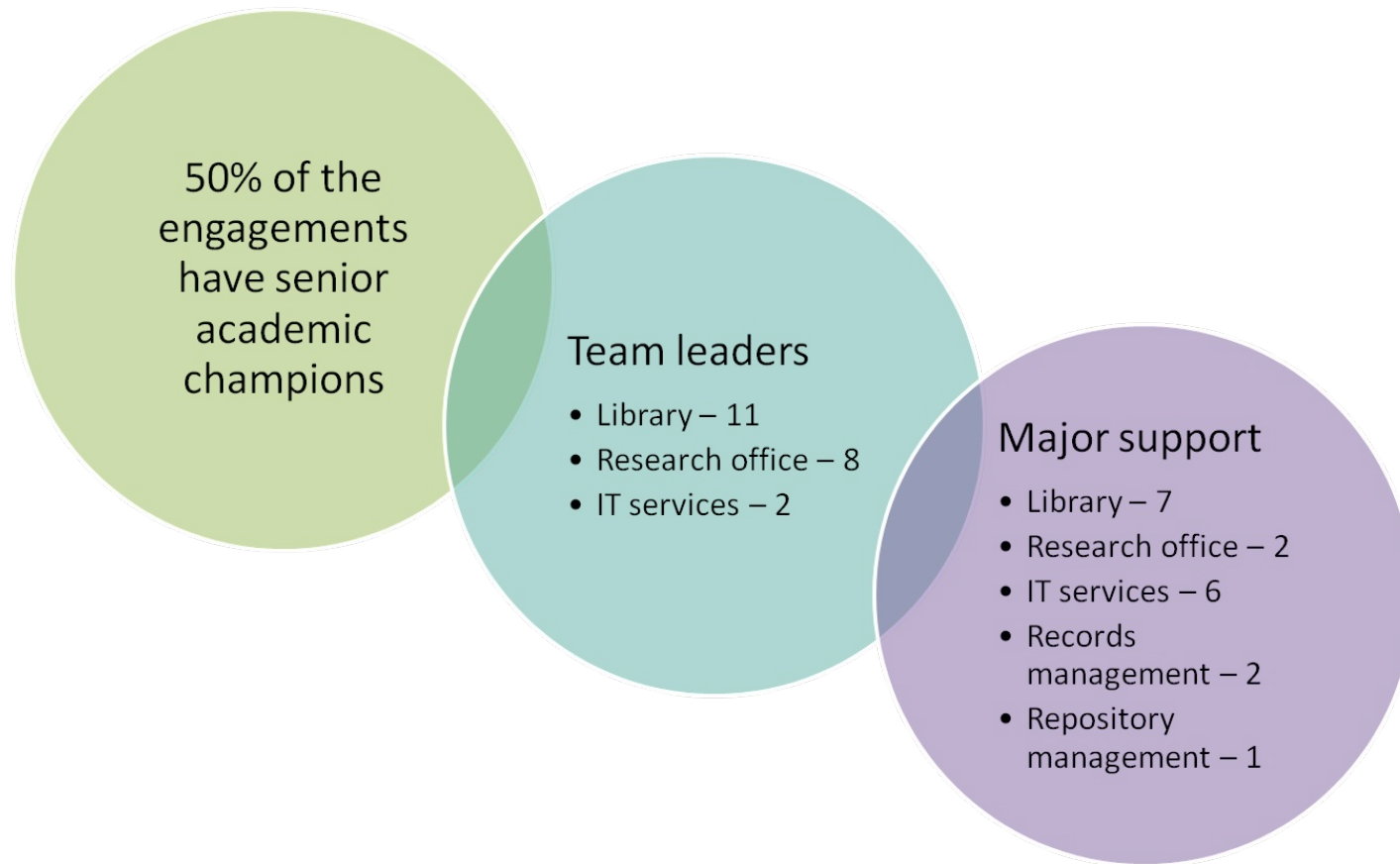


# DCC tailored support



[www.dcc.ac.uk/tailored-support](http://www.dcc.ac.uk/tailored-support)

# Roles of main participants



## Examples of support: **CARDIO at Queen Mary**

- Advised on the mini quiz and full CARDIO
- Ran a workshop to reach consensus on results
- Wrote up findings to feed into strategy
- ‘heatmap’ approach being used to guide roadmap development

DCC support fits into the curation strand of the IT Transformations

Project at QMUL

# Fostering skills development



# Guidance webpages

University of Glasgow

Subjects A-Z | Staff A-Z | Academic units A-Z

Enter your keywords here SEARCH

## Data management site support for researchers


- Creating your data
- Organising your data
- Accessing your data
- Looking after your data
- Training
- Who can help?
- Suggest a change

Home > Services A-Z > Data management support for researchers

## Data management support for researchers

Research data takes many forms, ranging from measurements, numbers and images to documents and publications. These webpages will help you plan, create, organise and look after your data.

### Creating your data




Do you know what formats to use or how to manage IPR and ethics? Who can help you explain this in your data plan?

### Organising your data



Can you name your files, your folders and control way that makes data easy to use?

### Accessing your data



Can you and your collaborators access all of the files you need? How does the storage you choose affect options for remote access?

### Looking after your data



How can you store and back up your data securely? Can the way you select and care for your data make a difference to the long-term?

[Bidding for funding](#) | [Data planning](#) | [Choosing file formats](#) | [Intellectual property rights](#) | [Data protection](#) | [Freedom of Information](#) | [Naming files](#) | [Structuring data](#) | [Documentation](#) | [Storage](#) | [Remote access](#) | [Back-up](#) | [Selection](#) | [Preservation](#) | [Repositories](#) | [Data sharing](#) | [Research data management at the University of Glasgow](#)



Students | Staff | A-Z

search site... go

## Research data

- Research data
- Introduction
- Planning for data
- Keeping data safe
- Sharing & reusing data
- Archiving data
- Support, advice & training
- Policy
- Contact us (email)

## Research data

Research data is any material collected, observed, or created for the purposes of analysis to generate original research results, irrespective of the format of data. It may be digital, paper based or in other forms.

### Overview of research data management, sharing and archival

Use the links below to find help for particular tasks, or browse the menu on the left by category.

## What to do when...

### ...planning a project

- Write a data management plan
- Find out what your funder expects
- Use tools and templates to write a data management plan
- Request research storage, *if only* (BUCS)

### ...working on a project

- Learn how to keep your data safe
- Organise and document your files
- Write or update your data management plan

### ...finishing a project

- Find out what your funder expects
- Select data for archival
- Choose a suitable archive/repository
- Digitise non-digital data

### ...using someone else's data

- Learn about sharing and reusing data
- Find reusable data

### ...somebody wants your data

- Learn how to license research data (Digital Curation Centre)
- Read the [Freedom of Information FAQ](#)

### ...you want to know more

- Contact us ([research-data@bath.ac.uk](mailto:research-data@bath.ac.uk))
- Attend a data management workshop (for research staff; for postgraduates)
- Other sources of help

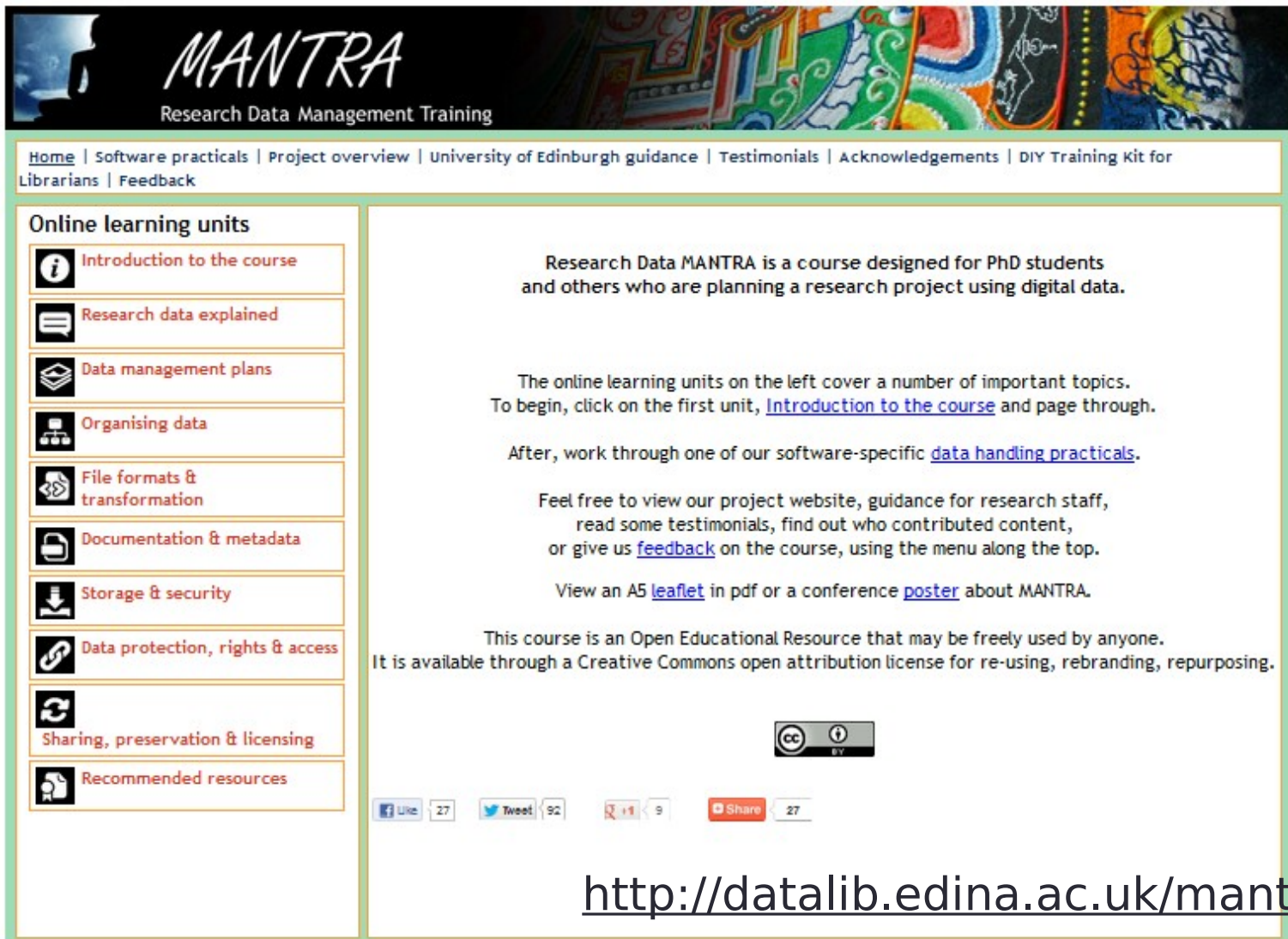
Video: How not to manage data  
[Data Sharing and Management Snafu in 3 short acts](#)

[www.gla.ac.uk/datamanagement](http://www.gla.ac.uk/datamanagement)



[www.bath.ac.uk/research/data](http://www.bath.ac.uk/research/data)

# Online training for PhD students













The screenshot shows the homepage of the MANTRA (Research Data Management Training) website. The header features the MANTRA logo and a navigation menu. A sidebar on the left lists online learning units. The main content area provides an overview of the course and its resources.

**MANTRA**  
Research Data Management Training

[Home](#) | [Software practicals](#) | [Project overview](#) | [University of Edinburgh guidance](#) | [Testimonials](#) | [Acknowledgements](#) | [DIY Training kit for Librarians](#) | [Feedback](#)

### Online learning units

-  [Introduction to the course](#)
-  [Research data explained](#)
-  [Data management plans](#)
-  [Organising data](#)
-  [File formats & transformation](#)
-  [Documentation & metadata](#)
-  [Storage & security](#)
-  [Data protection, rights & access](#)
-  [Sharing, preservation & licensing](#)
-  [Recommended resources](#)

Research Data MANTRA is a course designed for PhD students and others who are planning a research project using digital data.


The online learning units on the left cover a number of important topics. To begin, click on the first unit, [Introduction to the course](#) and page through.





After, work through one of our software-specific [data handling practicals](#).

Feel free to view our project website, guidance for research staff, read some testimonials, find out who contributed content, or give us [feedback](#) on the course, using the menu along the top.

View an A5 [leaflet](#) in pdf or a conference [poster](#) about MANTRA.

This course is an Open Educational Resource that may be freely used by anyone. It is available through a Creative Commons open attribution license for re-using, rebranding, repurposing.



 27  92  9  27

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

# Jorum

**Deposit resources**

Initial Questions → Upload → Describe → Describe → Describe → Describe → CC License

**Describe Item**

**Intended audience:**

- Doctoral researcher
- Librarian/Information Professional
- Postgraduate researcher
- Professional/senior level
- Research staff
- Support staff

Select one or more options from the drop down menu to describe the intended target audience of your resource.

**Skills:**

- Project planning
- Information searching, review and discovery
- Assessment and analysis of information sources
- Creation and collection
- Description
- Attribution and citation

Select one or more options from the drop down menu to indicate the skills that this resource aims to teach.

**Ability level:** Intermediary (understands)

Choose one entry from the drop down list to indicate the level on the audience's prior knowledge or experience.

**Delivery:**

- Classroom-based courses (lecture or workshop)
- Individual tuition
- Online courses
- Training materials

Select one or more options from the drop down menu to indicate how the resource can be used or delivered.

**Range:**

- Faculty
- Institution
- Inter-institutional
- International (open to non-UK HEIs)
- National/regional
- Not applicable

Select one or more options from the drop down menu to indicate which range of users the resource is relevant to particular groupings.

< Previous Save & Exit Next >



Click to edit Master text styles  
Second level

Third level

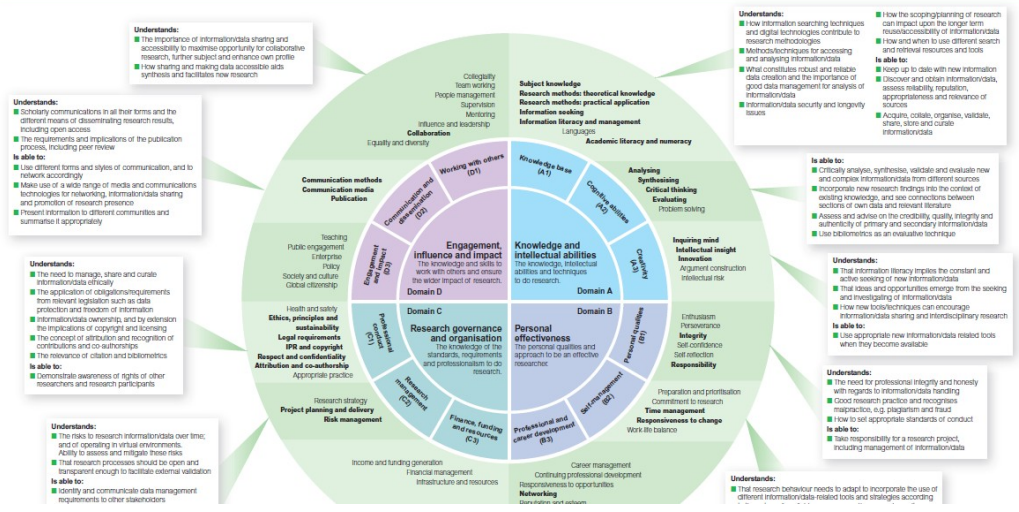
Fourth level

Fifth level

# FOSTER

<http://www.fosteropenscience.eu/>

## Information literacy lens on the Vitae Researcher Development Framework using the SCONUL Seven Pillars of Information Literacy



# Vitae RDF



# Examples of support: Capacity building at UEL

- Co-designed a series of modules for training RDM support staff
- Developed online learning resource, presentations and exercises
- Delivered training with UEL colleagues
- Helped to build awareness and confidence in dealing with RDM

SupportDM: <http://www.uel.ac.uk/trad/outputs/resources/>

Review of the course:

<http://datamanagementuel.wordpress.com/2013/06/19/are-you-rdm-ready-from-zeroes-to-heroes/>



# Supporting data management planning



# What is a DMP?

A short plan that outlines:

- what data will be created and how
- how it will be managed (storage, back-up, access...)
- plans for data sharing and preservation

# Some other funders that require DMPs or equivalent



Department  
for International  
Development

BILL & MELINDA  
GATES foundation



European  
Commission



GenomeCanada



National Science Foundation  
WHERE DISCOVERIES BEGIN



ACADEMY OF FINLAND



Vetenskapsrådet



National Institutes of Health  
*Turning Discovery Into Health*



National Institute for  
Health Research

DFG Deutsche  
Forschungsgemeinschaft



For research on  
diseases of poverty

UNICEF • UNDP • World Bank • WHO

FOSTER

# Five common themes

- Description of data to be collected / created  
(i.e. content, type, format, volume...)
  - Standards / methodologies for data collection & management
  - Ethics and Intellectual Property  
(highlight any restrictions on data sharing e.g. embargoes, confidentiality)
  - Plans for data sharing and access  
(i.e. how, when, to whom)

# DMPonline



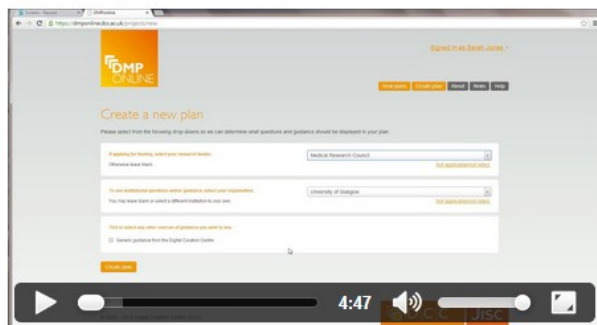
<https://dmponline.dcc.ac.uk>

[Home](#) [About](#) [News](#) [Help](#)

Welcome.

DMPonline has been developed by the **Digital Curation Centre** to help you write data management plans.

Screencast on how to use DMPonline



## Sign in

[Forgot your password?](#)

Remember me

Sign in

[Or, sign in with your institutional credentials](#) (UK users only)

## Sign up

New to DMPonline? Sign up today.

[Contact us](#) | [Terms of use](#) | [DMPonline previous version](#)

© 2004 - 2014 Digital Curation Centre (DCC)



# Main features of DMPonline

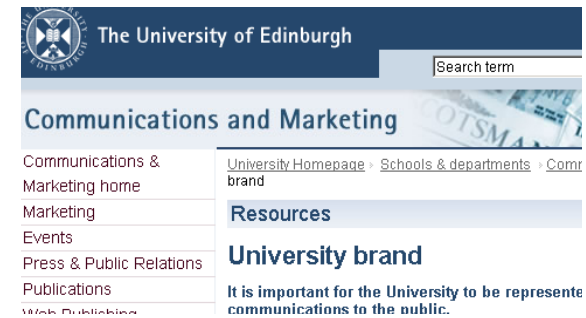
- Templates for different requirements (funder or institution)
- Tailored guidance (funder, institutional, discipline-specific etc)
- Ability to provide examples and boilerplate text
- Supports multiple phases (e.g. pre- / during / post-project)
- Granular read / write / share permissions
- Customised exports to a variety of formats
- API for systems interoperability

# Institutions can customise DMPonline

Profile local support via custom  
guidance and boilerplate text



Add your  
logo, URL etc



Select / write  
desired questions

[www.dcc.ac.uk/news/customising-dmponline](http://www.dcc.ac.uk/news/customising-dmponline)



# Facilitating research data discovery



# Archiving - external data centres

## Research funders' data centres...



**British Atmospheric Data Centre**  
NATIONAL CENTRE FOR ATMOSPHERIC SCIENCE  
NATURAL ENVIRONMENT RESEARCH COUNCIL



## Structured databases



**ChemSpider**  
The free chemical database



## Registries of international data centres

**Databib**

**re3data.org**  
REGISTRY OF RESEARCH DATA REPOSITORIES

## Disciplinary & community initiatives



**GoGeo**



# Institutional data repositories



<http://datashare.is.ed.ac.uk>



[www.dspace.cam.ac.uk](http://www.dspace.cam.ac.uk)

Not intended to replace national, subject or other established data repositories

Acknowledge hybrid environment



Research Data at Essex and DataPool at Southampton



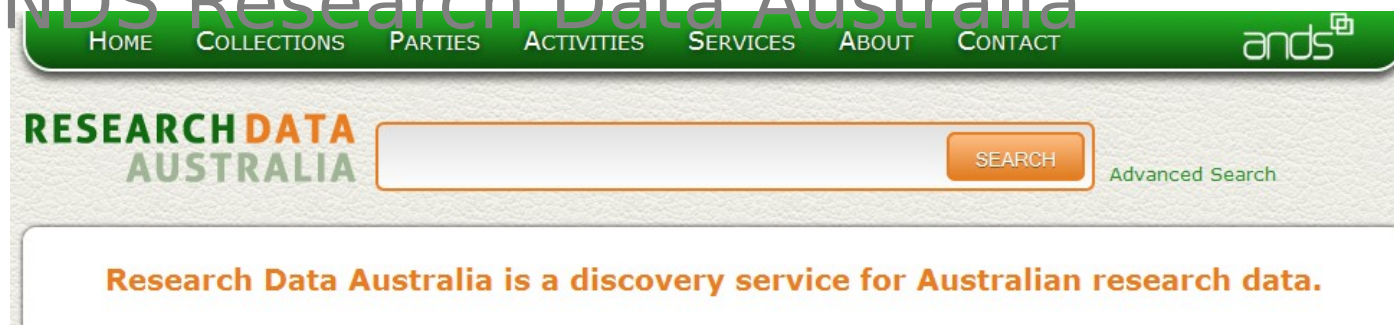
<https://databank.ora.ox.ac.uk>

# Data catalogues

- Oxford is developing its DataFinder tool  
<http://blogs.it.ox.ac.uk/damaro>
- Research Data @ Essex has developed a profile based on DataCite, Inspire and DDI standards  
[www.data-archive.ac.uk/media/395364/rde\\_march2013\\_repositoryoutputs.pdf](http://www.data-archive.ac.uk/media/395364/rde_march2013_repositoryoutputs.pdf)
- C4D is developing a research data extension to the CERIF standard - <http://cerif4datasets.wordpress.com>
- CKAN is being explored by various projects  
<http://ckan.org>

# Research Data Registry pilot

- aim is to improve data discovery and provide a national level service
- aggregate metadata relating to data collections or datasets held in UK HEIs and subject data centres
- have mapped metadata schema to develop crosswalks
- trialling ANDS Research Data Australia



The screenshot shows the top portion of the ANDS Research Data Australia website. At the top is a green navigation bar with white text links: HOME, COLLECTIONS, PARTIES, ACTIVITIES, SERVICES, ABOUT, and CONTACT. To the right of these links is the ANDS logo, which consists of the letters 'ands' in a lowercase, sans-serif font with a small square icon to its upper right. Below the navigation bar is a white search area. On the left of this area is the text 'RESEARCH DATA AUSTRALIA' in a bold, sans-serif font, with 'RESEARCH DATA' in green and 'AUSTRALIA' in orange. To the right of this text is a white search input field with an orange border. To the right of the input field is an orange button with the word 'SEARCH' in white. Further to the right is a link for 'Advanced Search' in green text. Below the search area is a white box with a thin orange border containing the text: 'Research Data Australia is a discovery service for Australian research data.' in orange.



# Jisc RDRDS Pilot Partners

## Data centres:

- ▶ UK Data Archive
- ▶ NERC Data Catalogue Service
  - ▶ BADC
  - ▶ BODC
  - ▶ EIDC
  - ▶ NEODC
  - ▶ NGDC
  - ▶ PDC
  - ▶ UKSSDC
  - ▶ ADS

## Universities:

- ▶ Edinburgh
- ▶ Glasgow
- ▶ Hull
- ▶ Lincoln
- ▶ Leeds
- ▶ Oxford
- ▶ Oxford Brookes
- ▶ St Andrews
- ▶ Southampton

# RDRDS Metadata Work

## DDI Codebook 2.5

- ▶ UK Data Archive

## DataCite 3

- ▶ Edinburgh (TBC)
- ▶ Oxford (TBC)
- ▶ Hull (TBC)

## OAI-PMH Dublin Core

- ▶ Oxford Brookes (TBC)

## UK Gemini 2.2

- ▶ NERC Data Catalogue Service

## EPrints 3

- ▶ Glasgow
- ▶ Leeds
- ▶ Lincoln (TBC)
- ▶ Southampton

# Metadata catalogues: defining levels



# RDRDS ongoing investigations

- ▶ Would another platform suit us better?
- ▶ Would another internal metadata scheme suit us better than RIF-CS?
- ▶ What use cases should the registry target?
- ▶ How can we add value to the registry's records?
- ▶ Could the registry add value to other systems?

# Assessing RDM costs and benefits



● Full Coverage

○ Partial Coverage

Click to edit Master text styles

## Second level

Research Funders	Policy Coverage		Policy Third level					Support Provided			
	Published outputs	Data	Time limits	Data plan	Access/sharing	Long term curation	Monitoring	Guidance	Repository	Data centre	Costs
AHRC	●	●	●	●	●	○	○	●	○	○	○
BBSRC	●	●	●	●	●	●	●	●	●	○	●
CRUK	●	●	●	●	●	●	●	○	●	○	○
EPSRC	●	●	●	○	●	●	●	○	○	○	●
ESRC	●	●	●	●	●	●	●	●	●	●	○
MRC	●	●	●	●	●	●	○	○	●	○	○
NERC	●	●	●	●	●	●	●	●	●	●	○
STFC	●	●	●	●	●	●	●	○	●	○	○
Wellcome Trust	●	●	●	●	●	●	●	●	●	○	●

<http://www.dcc.ac.uk/resources/policy-and-legal/overview-funders-data-policies>

# Collaboration to Clarify the Costs of Curation (4C)

Welcome

http://ccex.org

## Welcome to CCEX

Login / Register

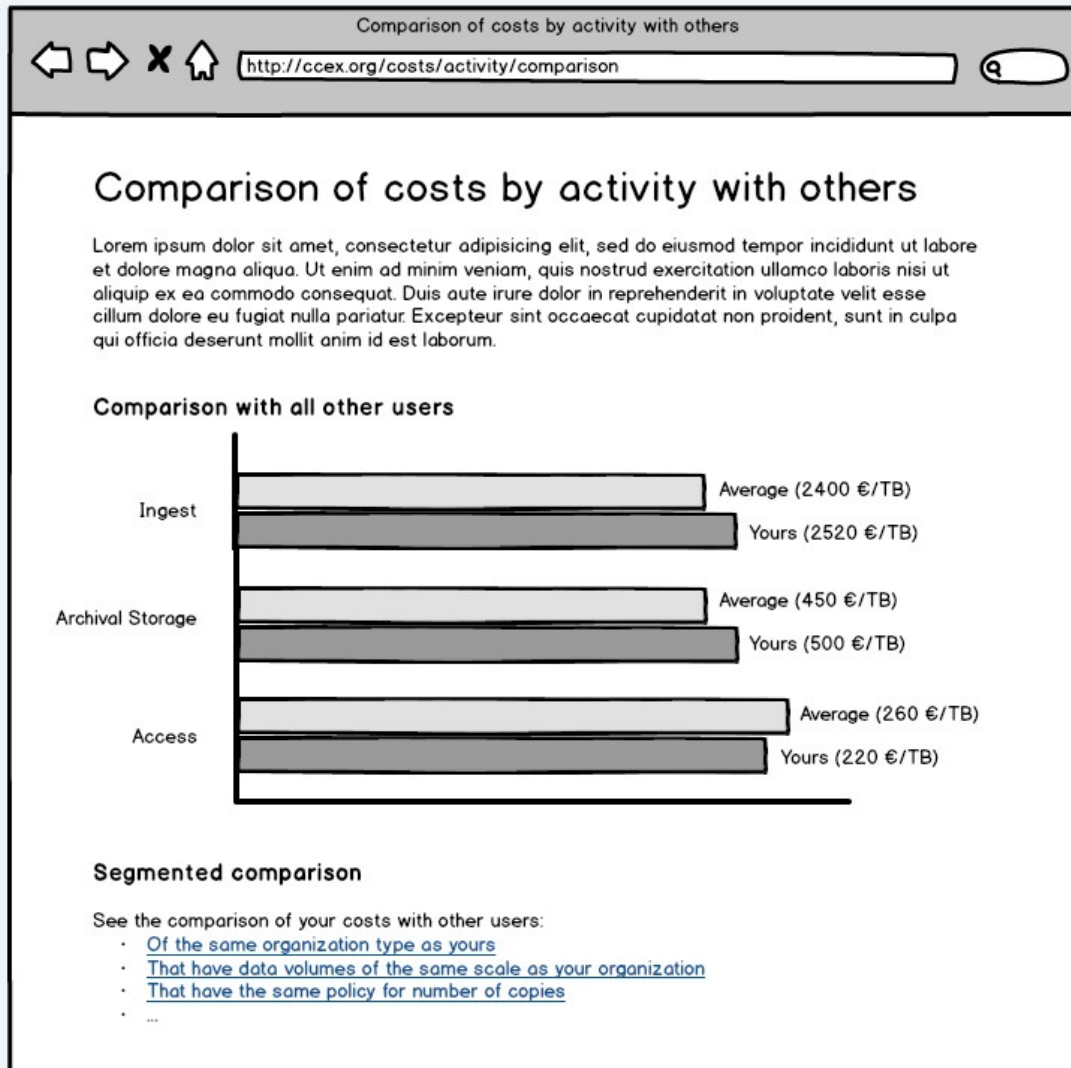
Lorem ipsum dolor sit amet, consectetur adipiscing elit. Curabitur tempus condimentum est ut convallis. Suspendisse venenatis justo id lorem viverra pretium. Fusce vitae turpis et eros eleifend semper sed nec felis. Morbi interdum risus nulla. Nullam varius tortor condimentum feugiat tempor. Vivamus sed justo tincidunt, feugiat purus vel, fringilla urna. Nulla tortor turpis, interdum ac ligula sed, congue viverra odio. Proin id feugiat neque, a scelerisque elit. Vestibulum vehicula ipsum dui, quis bibendum dolor facilisis non. Vivamus vehicula commodo ligula ac, fringilla Praesent ac pretium tortor. Sed pulvinar condimentum dui. Ut massa libero, feugiat porta felis sit amet, dictum ullamcorper augue. Nunc faucibus nisl magna, vitae adipiscing arcu faucibus sed. Donec aliquet massa velit, ut dictum dui ornare vel. Proin et tortor consequat, sodales leo eget, imperdiet elit.

### What the CCEX can do for you

- Define your organization profile
- Compare your costs by activity
- Compare your costs by capital and labour
- Browse through cost models descriptions and comparisons
- Know what cost models you should use
- Know about other's risks and benefits
- Connect with others with the same motivations
- Stay informed about all news and events
- Know about suppliers and their services



# Collaboration to Clarify the Costs of Curation (4C)



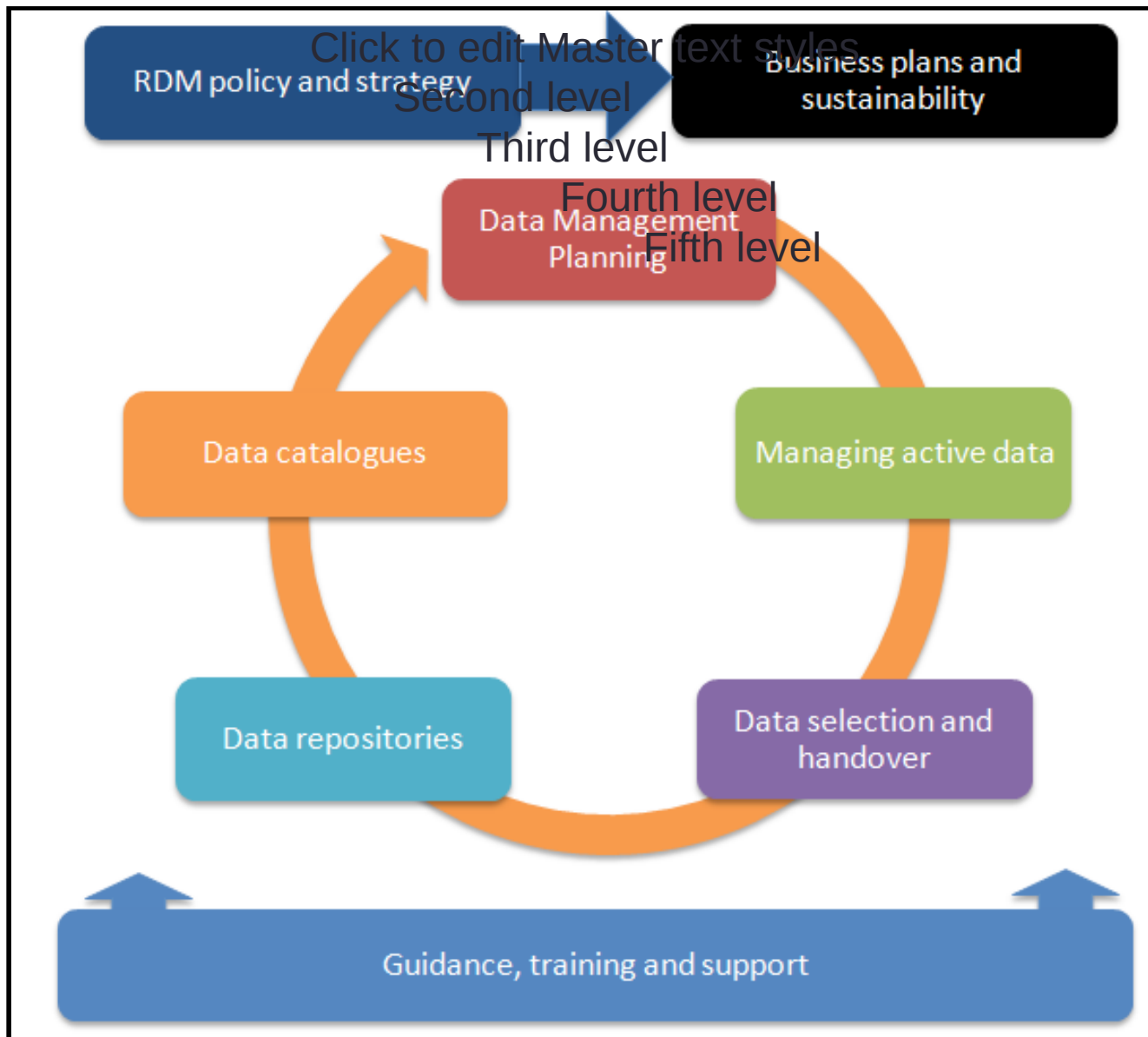
<http://4cproject.eu/>



# Lessons we can share



# Components of RDM services



# UK research data policies



“Statement of commitment”  
□ Infrastructure □ policy



Baseline of RCUK Code  
+ procedures & support



“10  
commandments”  
mutual promises  
aspirational



legal tone / language  
a section in uni DM  
policy  
useful guide as  
appendix



Based on Edin.  
with a few  
additions



# Sample roadmaps: University of Bath

- Based on Monash University RDM strategy
- Identifies the current position and proposes activity
- Defines roles and responsibilities and timeframes

Objectives	Actions	Milestones	Roles & Responsibilities
<b>1.1 Develop the data management skills and knowledge of Bath researchers</b>	<ul style="list-style-type: none"> <li>• Initiate and coordinate a program of research data management (RDM) skills development for Bath researchers</li> <li>• Work with Faculties, DTCs and the Bath Graduate Schools to explore embedding research data management skills development in coursework curricula and professional development offerings</li> </ul>	<ul style="list-style-type: none"> <li>• Pilot RDM skills development to new Bath researchers from September 2013</li> <li>• High quality training delivered (participant feedback, impact notes)</li> </ul>	<ul style="list-style-type: none"> <li>• <b>Associate Deans of Research</b></li> <li>• <b>Doctoral Training Centres and Graduate Schools</b></li> <li>• <b>Library/UKOLN/BUCS</b></li> <li>• Learning &amp; Teaching Enhancement Office &amp; Researcher Development Unit</li> </ul>
<b>1.2 Develop the skills and knowledge of key professional services staff</b> <ul style="list-style-type: none"> <li>• <b>Library</b></li> <li>• <b>BUCS</b></li> <li>• <b>RDSO</b></li> </ul>	<ul style="list-style-type: none"> <li>• Initiate and coordinate a program of research data management skills development for key professional services staff at Bath</li> <li>• Seek opportunities for professional services staff to network and participate in communities of practice – internally, nationally and internationally</li> <li>• Promote research data management career paths within professional support services where appropriate</li> </ul>	<ul style="list-style-type: none"> <li>• Pilot RDM skills development to key professional services staff from September 2013</li> <li>• High quality training delivered (participant feedback, impact notes)</li> <li>• Support Bath staff participation in DCC meetings (RDMF, IDCC etc.)</li> </ul>	<ul style="list-style-type: none"> <li>• <b>Library/BUCS</b></li> <li>• <b>RDSO</b></li> <li>• Staff Development Unit</li> <li>• Data Scientist*</li> </ul>

<http://www.bath.ac.uk/rds0/University-of-Bath-Roadmap-for-EPS-RC.pdf>

# Architecture models

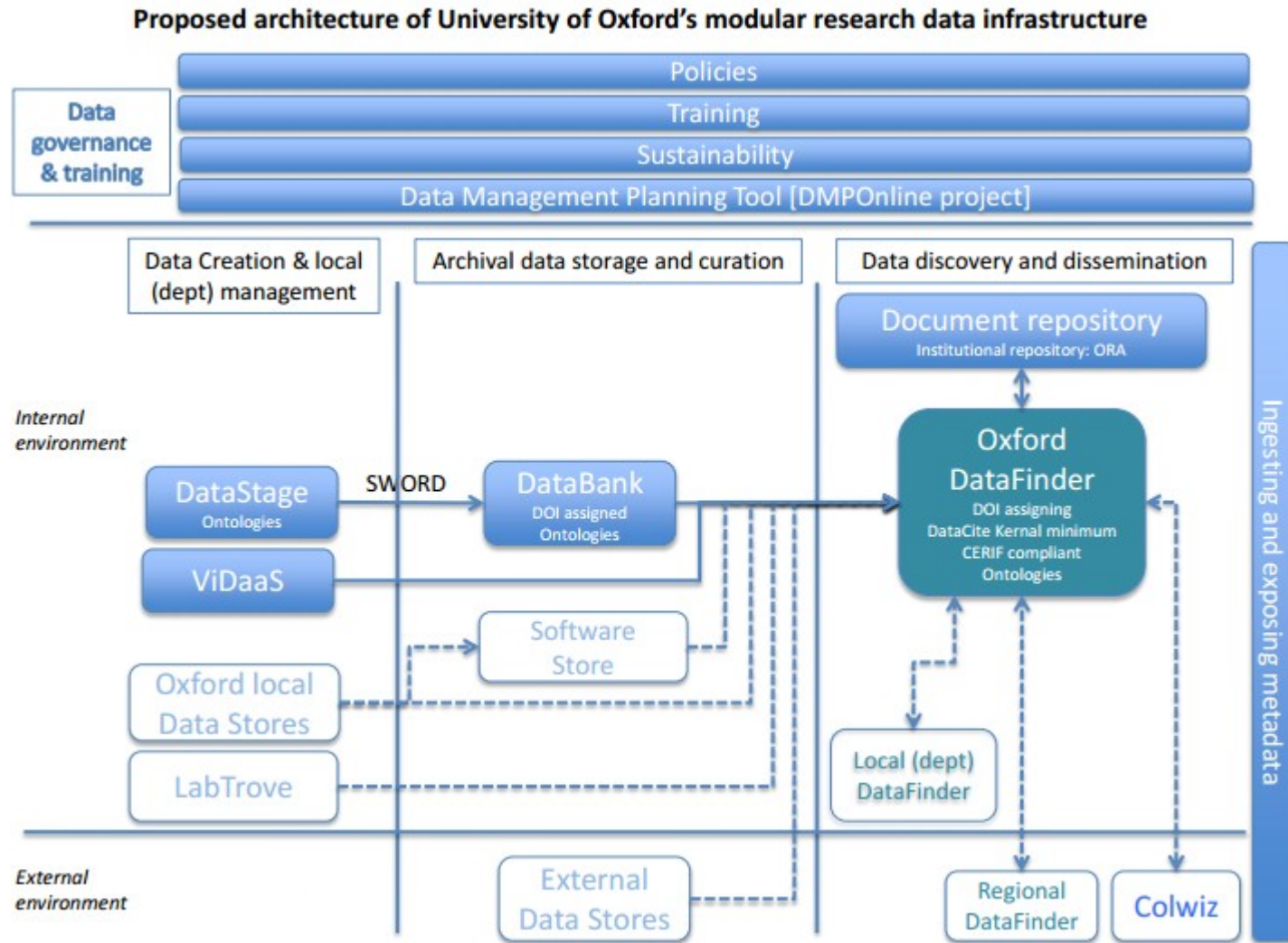


Diagram courtesy of Sally Rumsey, University of Oxford

# Thanks - any questions?

DCC guidance, tools and case studies:  
[www.dcc.ac.uk/resources](http://www.dcc.ac.uk/resources)

Follow us on twitter:  
@digitalcuration and #ukdcc



D|C|C

because good research needs good data

