FOSTER

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*

CC BY

Funded by: Jisc

The DCC Mission

"Helping to build capacity, capability and skills in data management and curation across the UK's higher education research community"

Phase 3 Business Plan

FOSTER

www.dcc.ac.uk

JISC

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

FOSTER

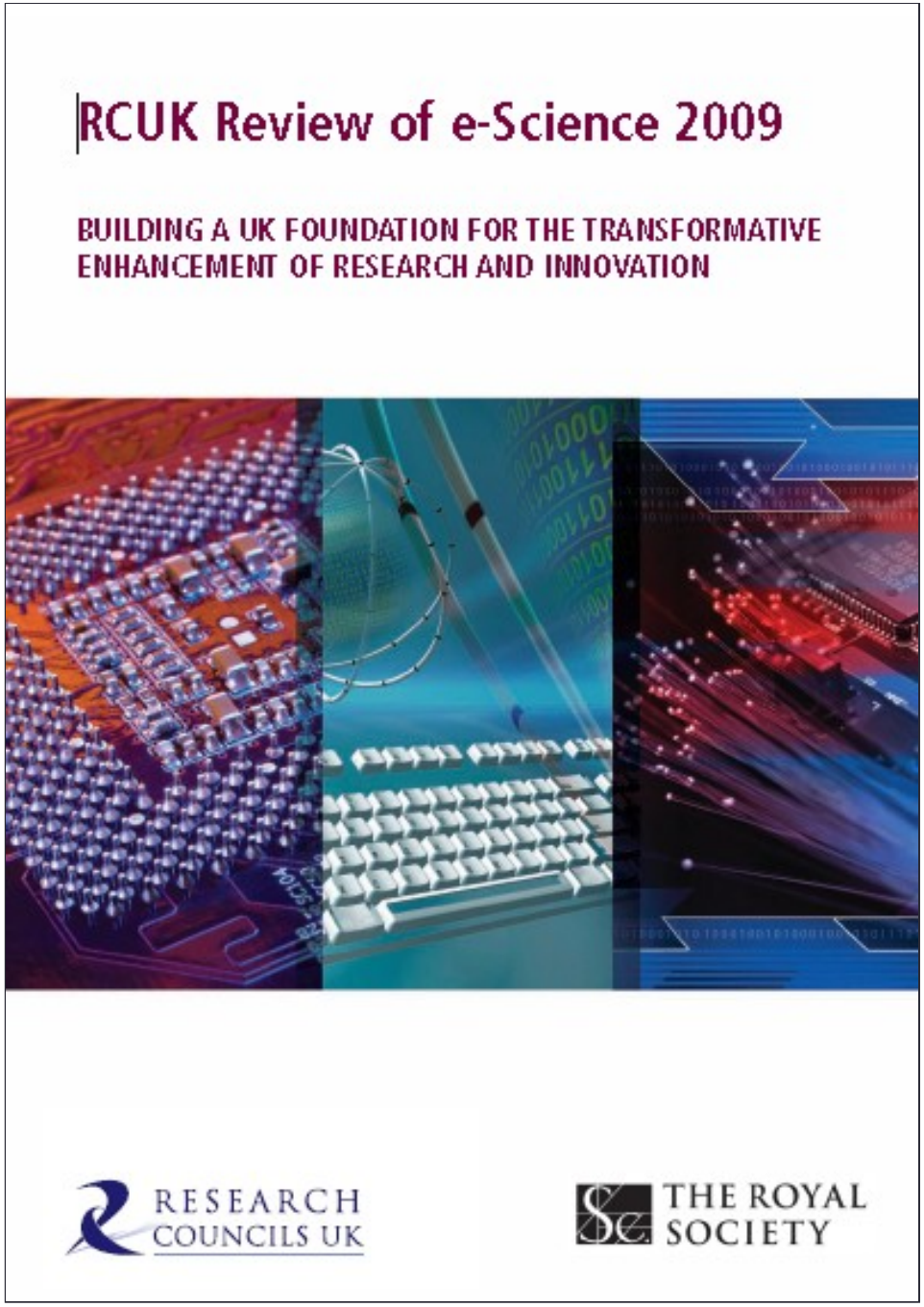
UK Context

FOSTER

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)

FOSTER



"Data sets are becoming the new instruments of science"

Dan Atkins, University of Michigan

Digital data as the new special collections?

Sayeed 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

guardian.co.uk

News | Sport | Comment | Culture | Business | Money | Life & style

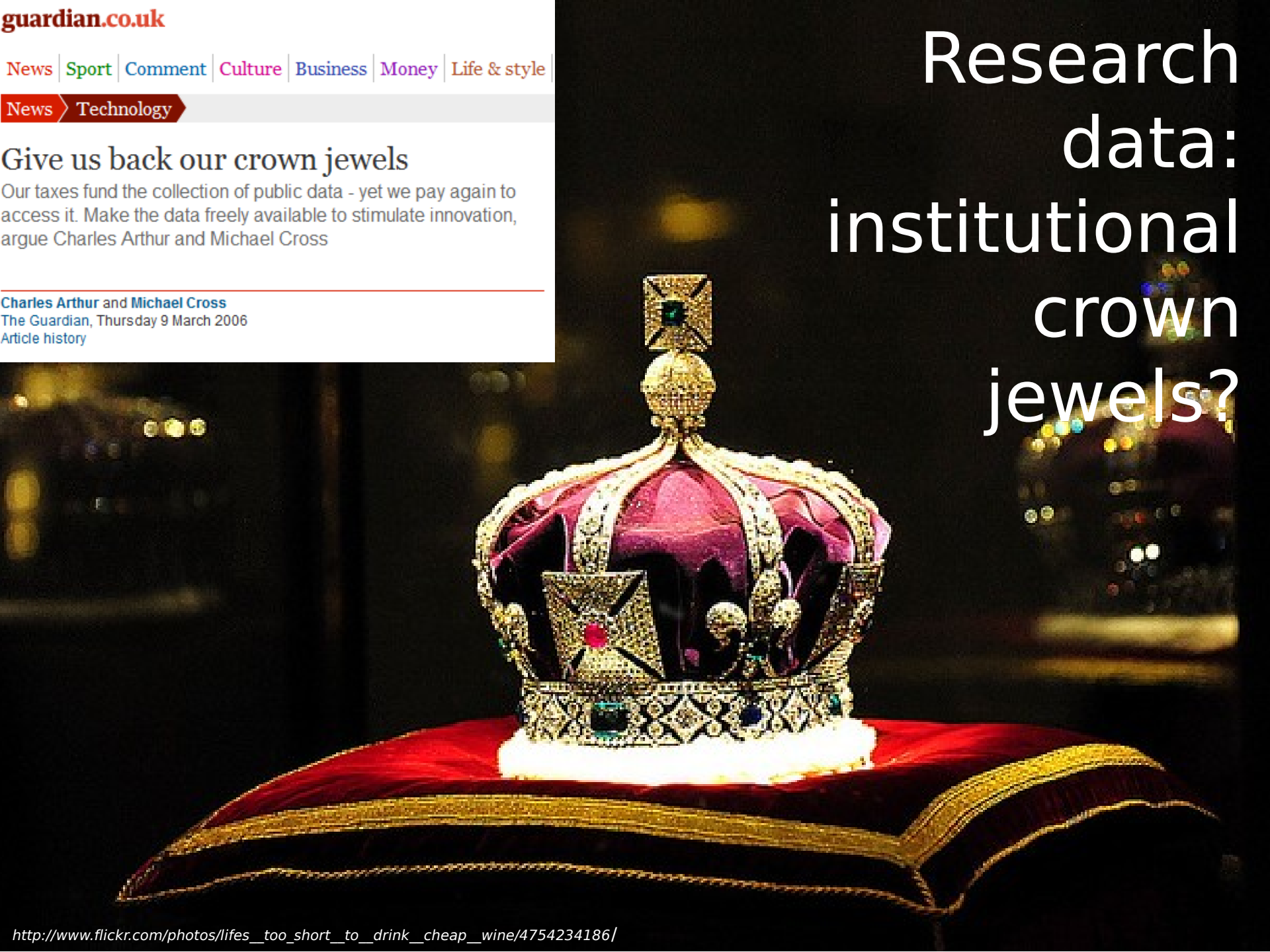
News › Technology

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?



http://www.flickr.com/photos/lifes__too_short__to__drink__cheap__wine/4754234186/

What do UK funders expect?

FOSTER

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

FOSTER

● Full Coverage ◐ Partial Coverage ○ No Coverage

	Policy Coverage		Policy Stipulations					Support Provided			
Research Funders	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

FOSTER

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

FOSTER

Working with UK Universities

FOSTER

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

FOSTER

The engagement model

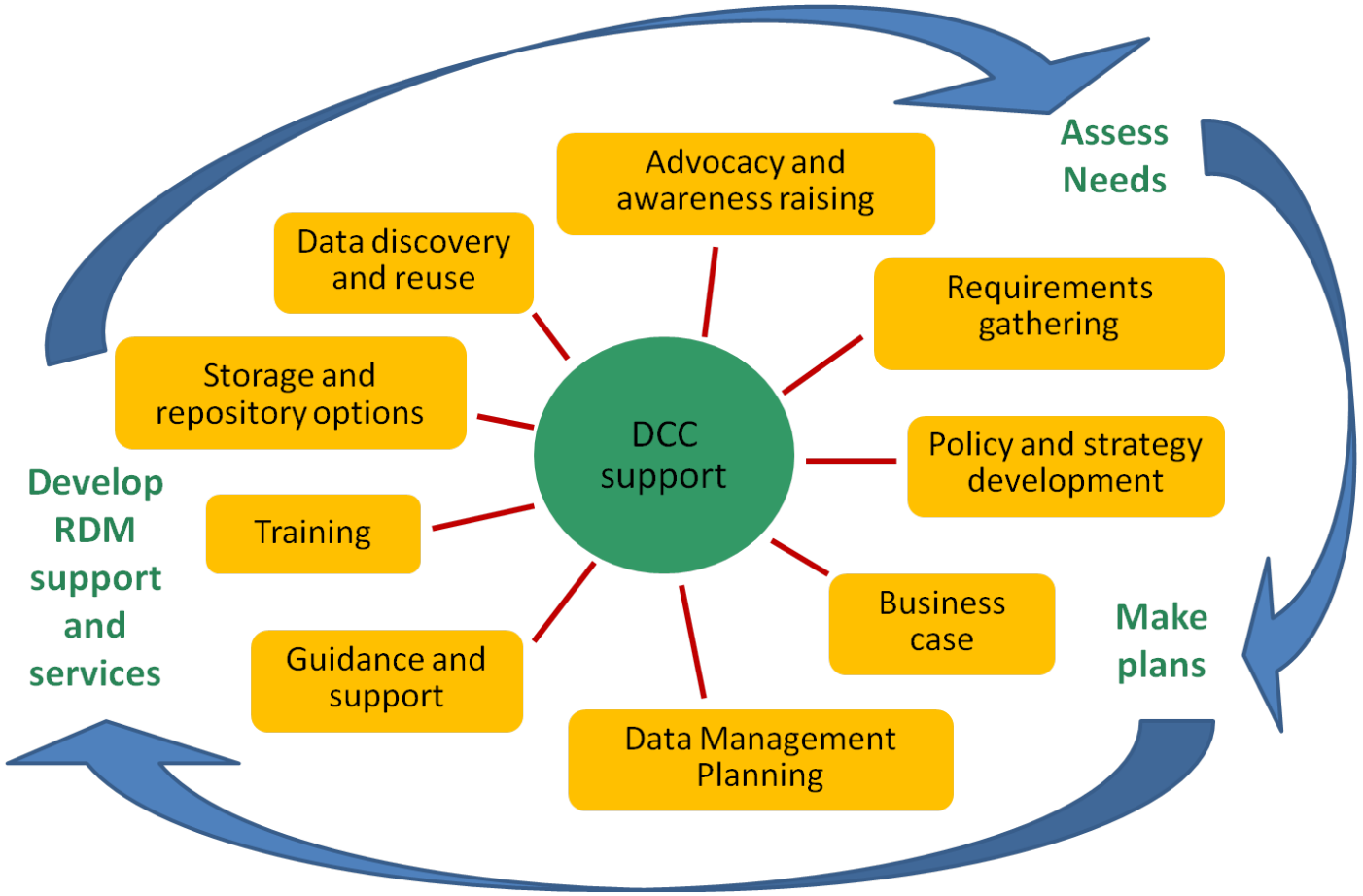
- Efficient information sharing across HE sector
- Redesign of DCC support

FOSTER

Who we worked with

FOSTER

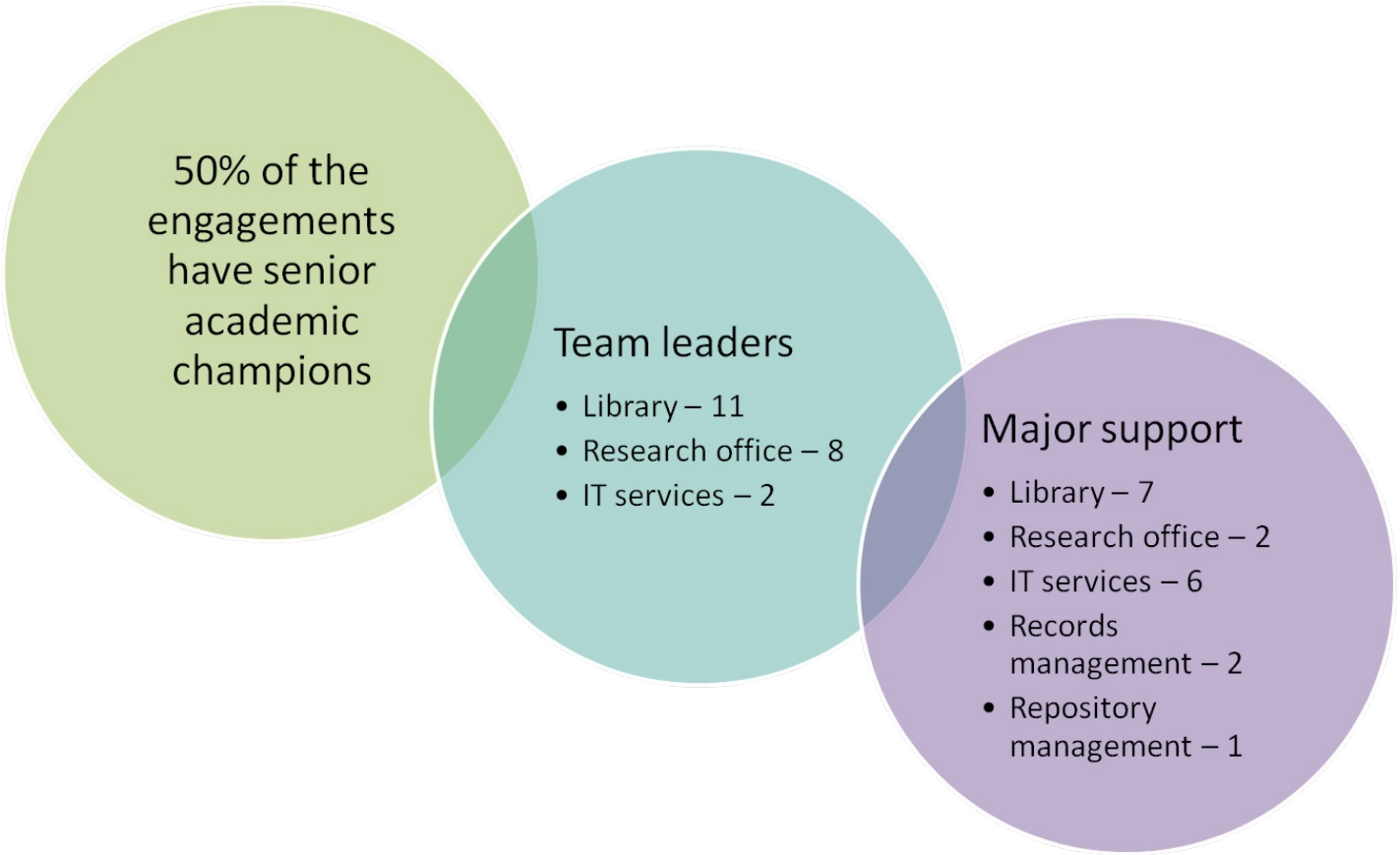
DCC tailored support



www.dcc.ac.uk/tailored-support

FOSTER

Roles of main participants



FOSTER

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

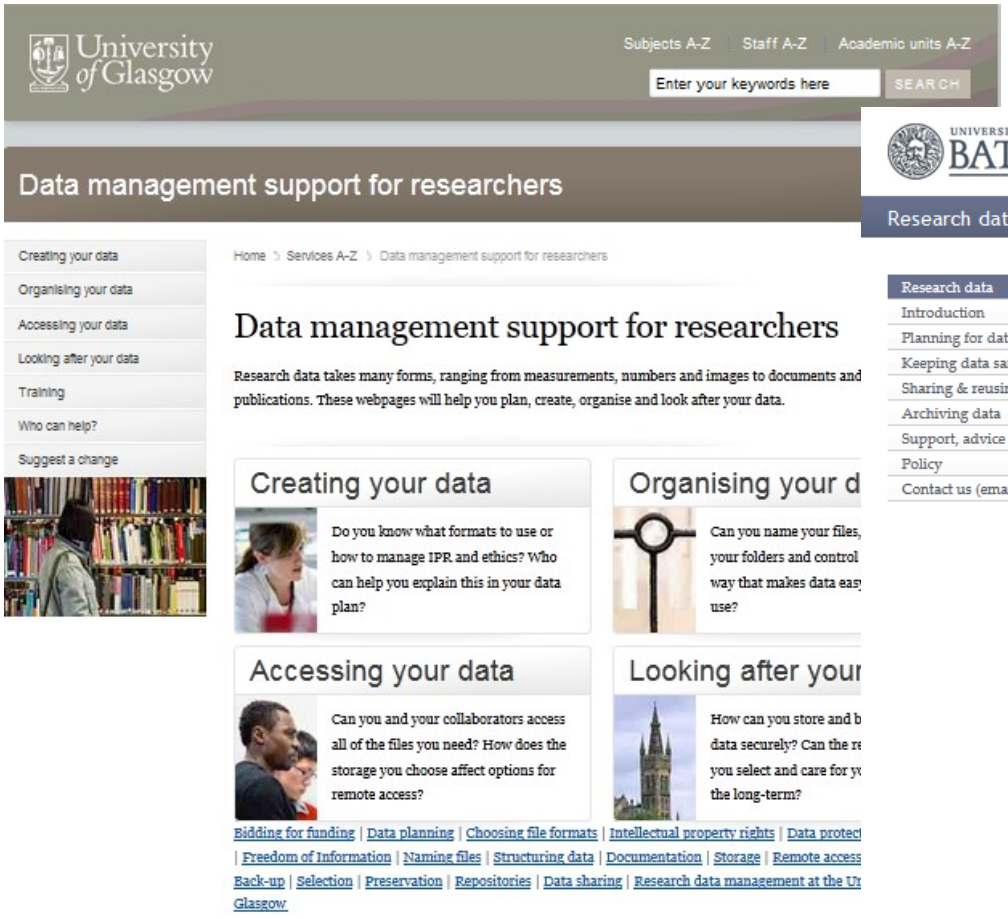
Project at QMUL

FOSTER

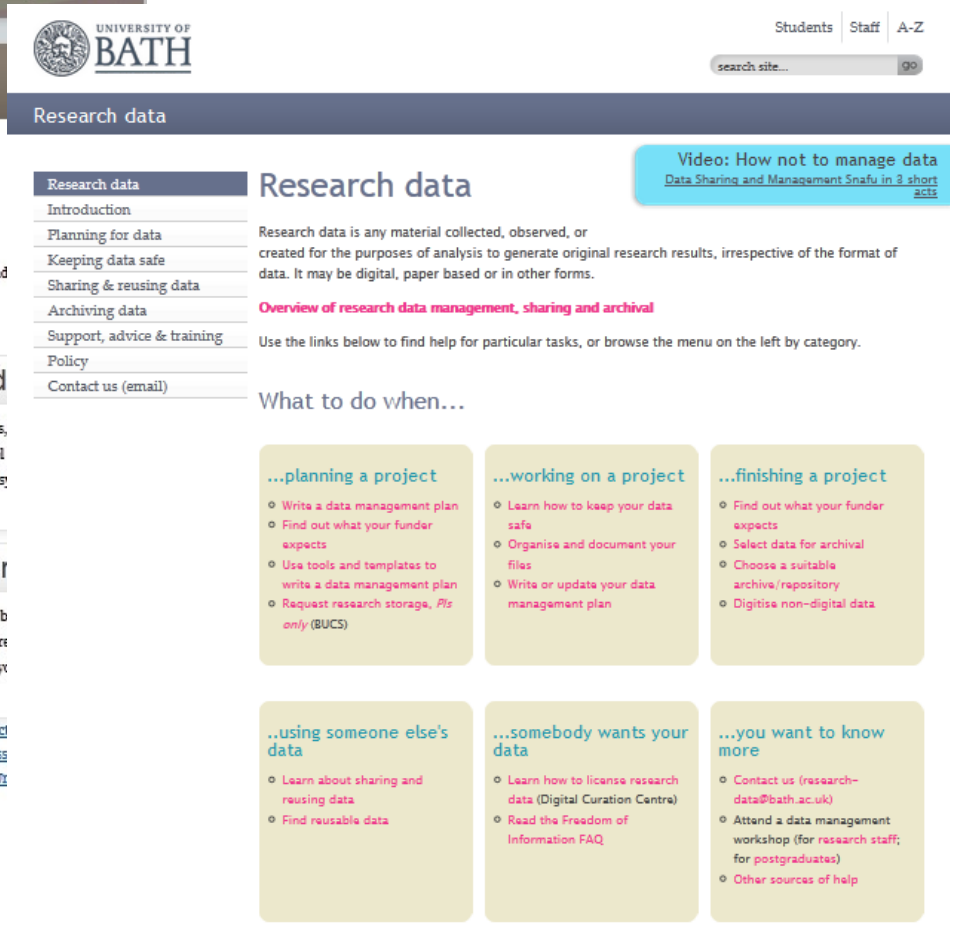
Fostering skills development

FOSTER

Guidance webpages



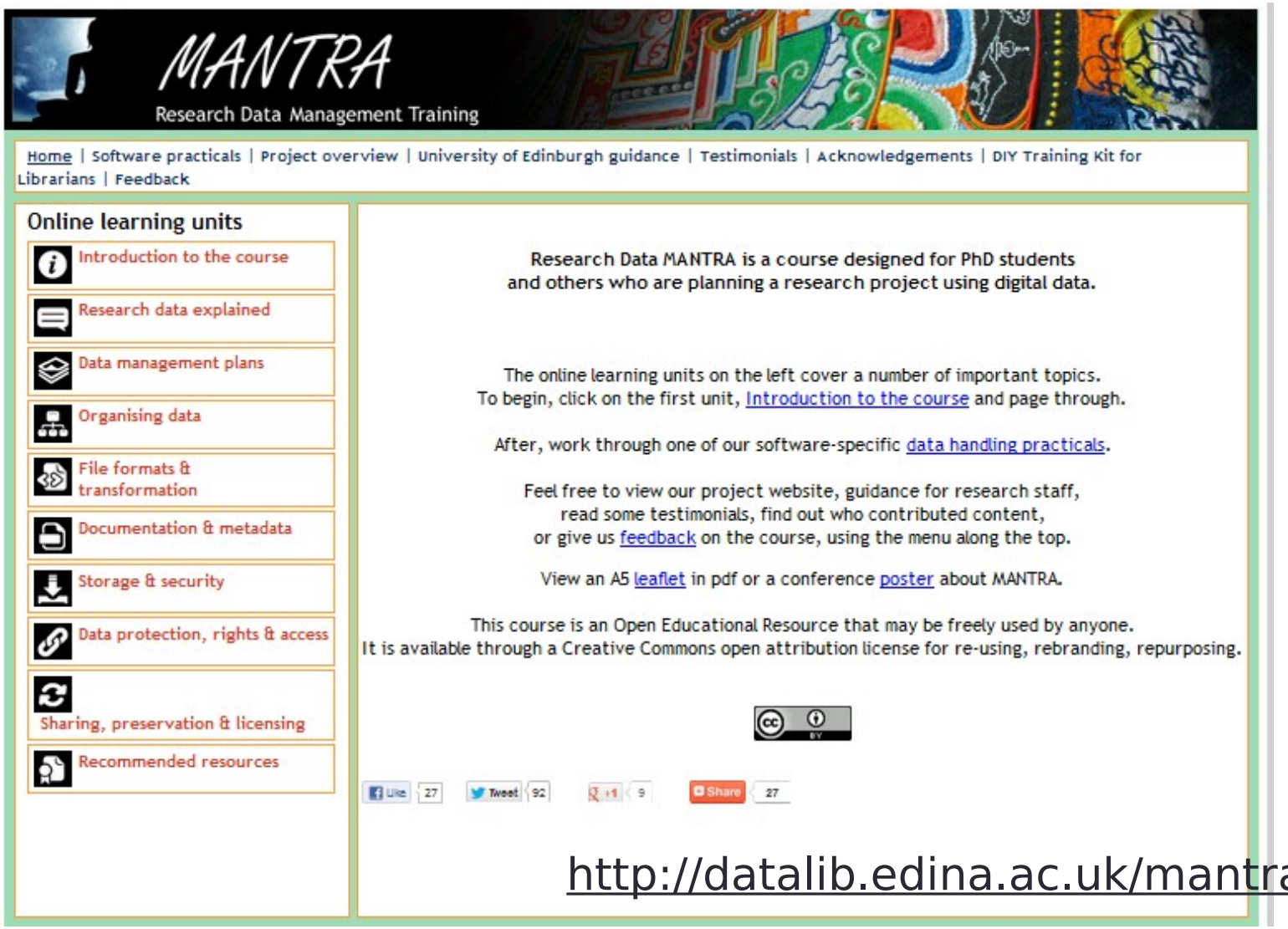
www.gla.ac.uk/datamanagement



www.bath.ac.uk/research/data

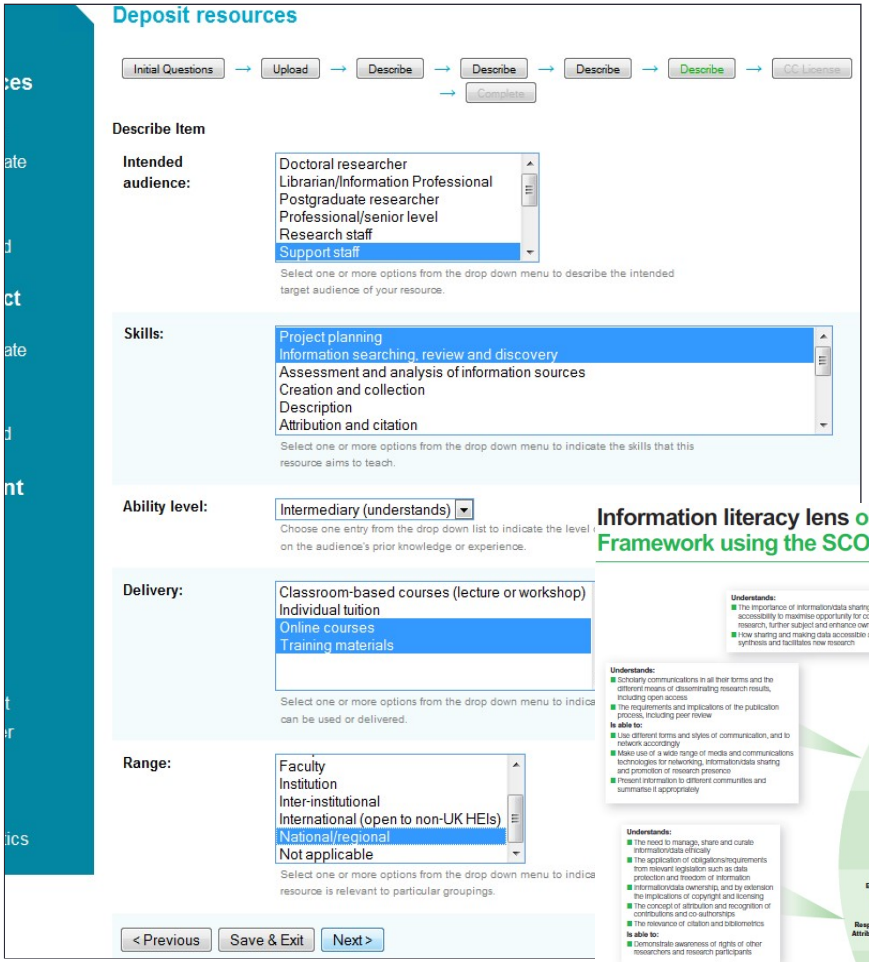
FOSTER

Online training for PhD students



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

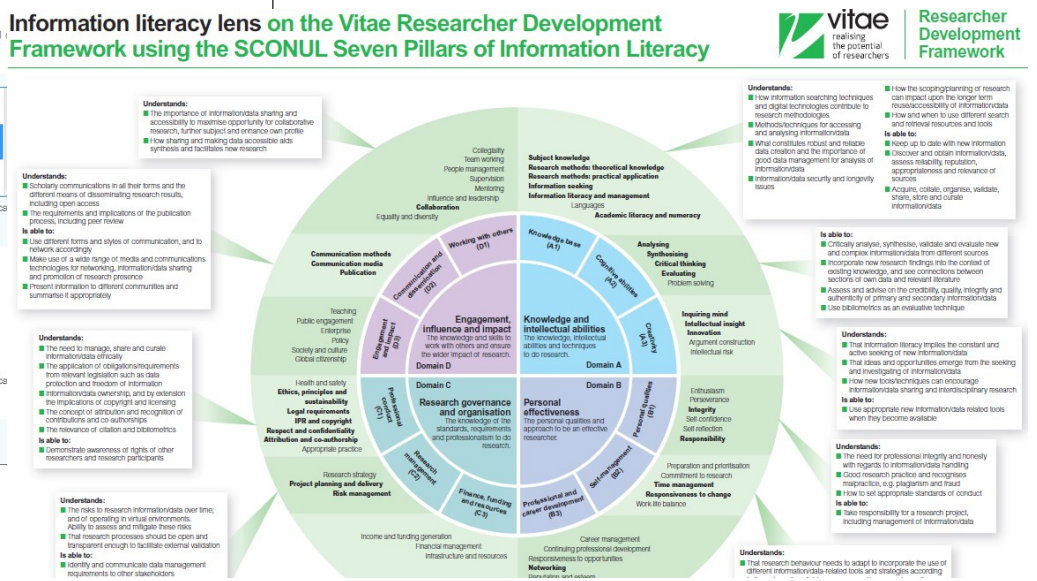
Jorum



Click to edit Master text styles
Second level
Third level
Fourth level
Fifth level

FOSTER

http://www.fosteropenscience.eu/



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/

FOSTER

Supporting data management planning

FOSTER

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

FOSTER

Some other funders that require DMPs or equivalent

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)

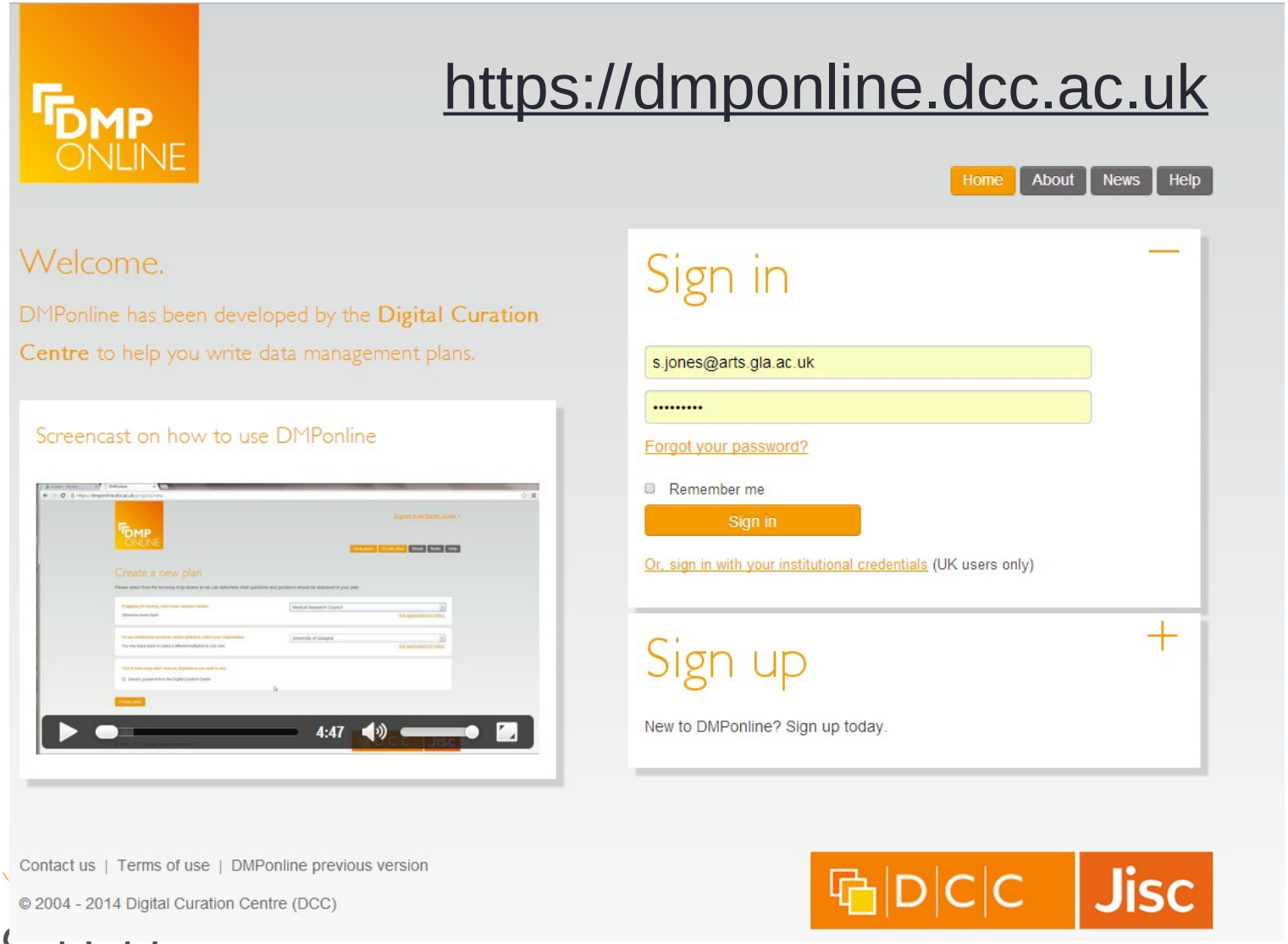
  - Plans for data sharing and access

  (i.e. how, when, to whom)

  - Strategy for long-term preservation

FOSTER

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
4:47
Sign in
s.jones@arts.gla.ac.uk
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
- Shibboleth authentication

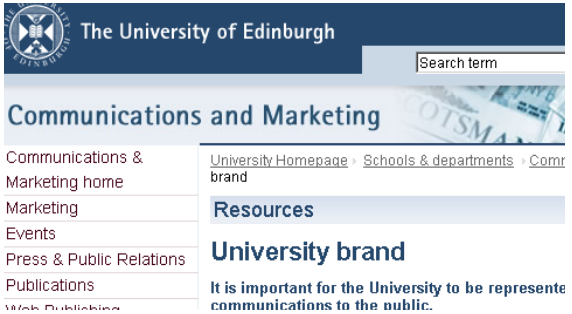
FOSTER

Institutions can customise DMPonline

Profile local support via custom guidance and boilerplate text

Select / write desired questions

Add your logo, URL etc



FOSTER

www.dcc.ac.uk/news/customising-dmponline

Facilitating research data discovery

FOSTER

Archiving – external data centres

Research funders' data centres…

ads ARCHAEOLOGY DATA SERVICE

esds Economic and Social Data Service

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

Structured databases

ChemSpider
The free chemical database

RCSB PDB
PROTEIN DATA BANK

Disciplinary& community initiatives

GoGeo

figshare

zenodo

Registries of international data centres

Databib

re3data.org
REGISTRY OF RESEARCH DATA REPOSITORIES

Institutional data repositories

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

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

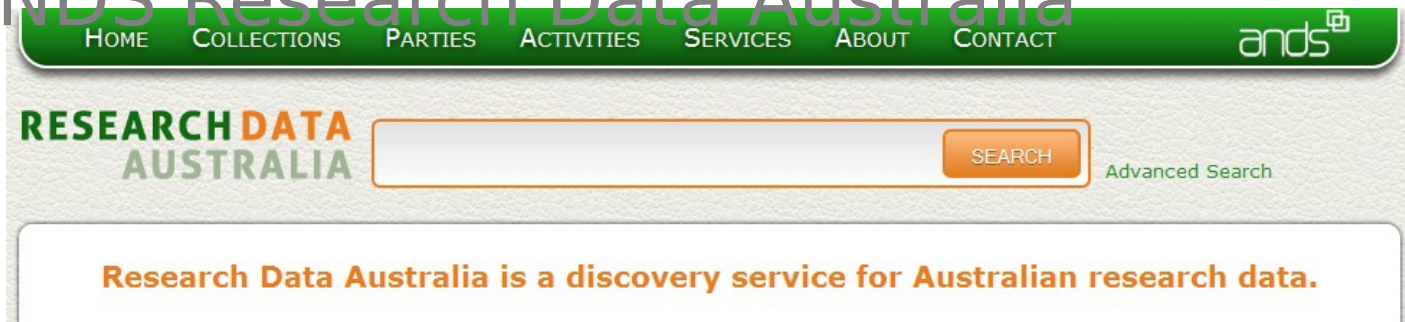
https://databank.ora.ox.ac.

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
- 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



FOSTER

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

FOSTER

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

FOSTER

Metadata catalogues: defining levels

FOSTER

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?

FOSTER

Assessing RDM costs and benefits

FOSTER

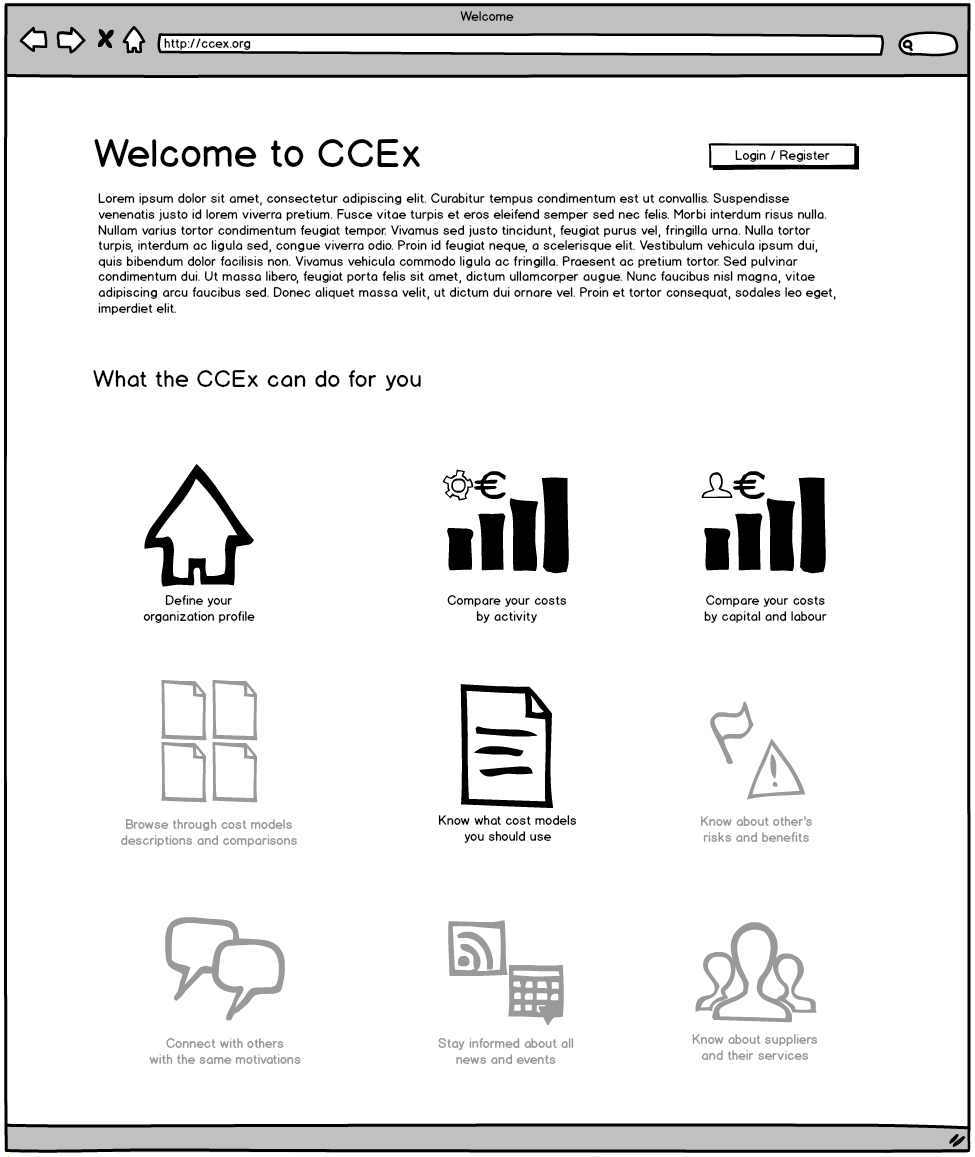
Full Coverage Partial Coverage No Coverage

Click to edit Master text styles
Second level
Third level
Fourth level
Fifth level

	Policy Coverage		Policy Stipulations					Support Provided			
Research Funders	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)



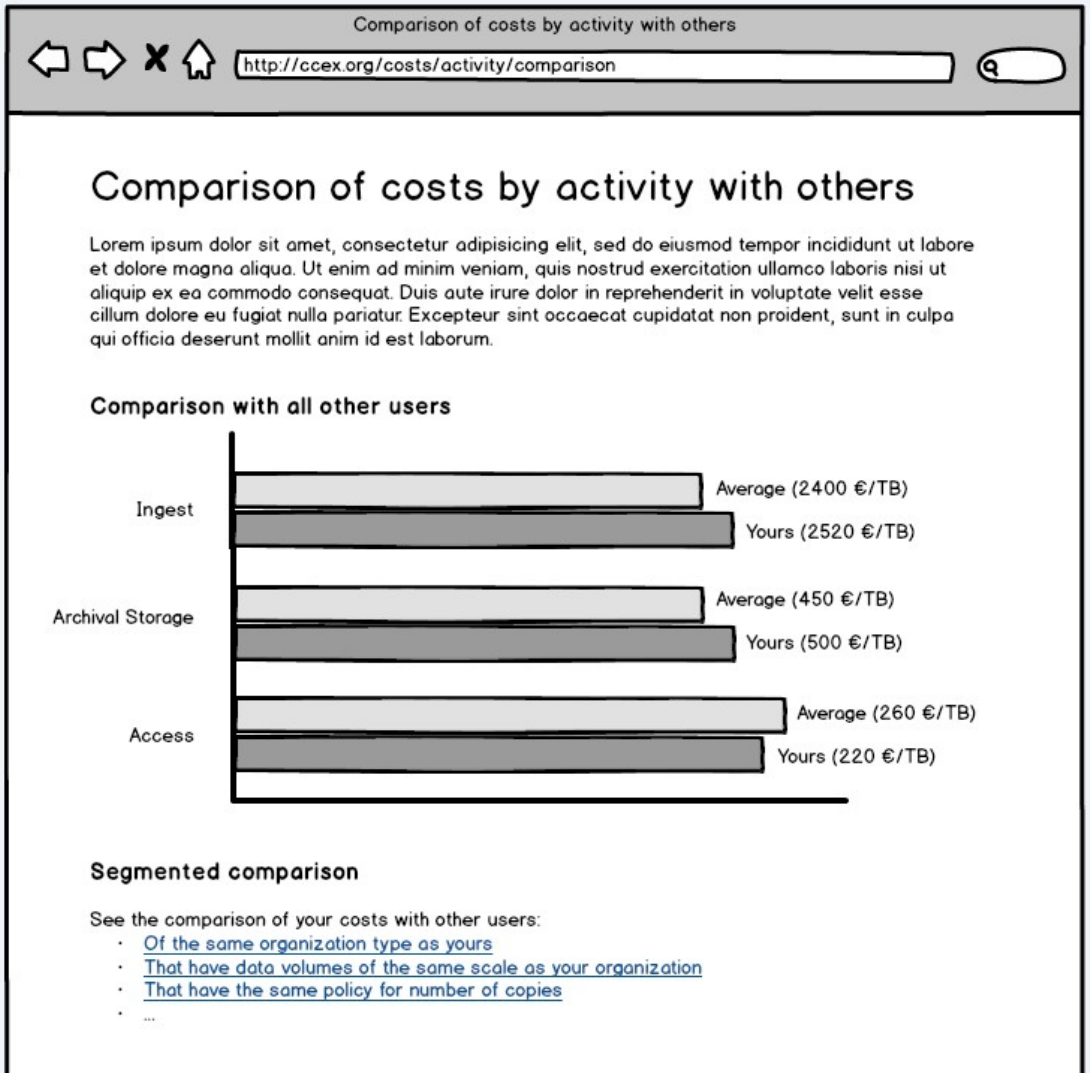
FOSTER

http://4cproject.eu/

4C

SEVENTH FRAMEWORK PROGRAMME

Collaboration to Clarify the Costs of Curation (4C)

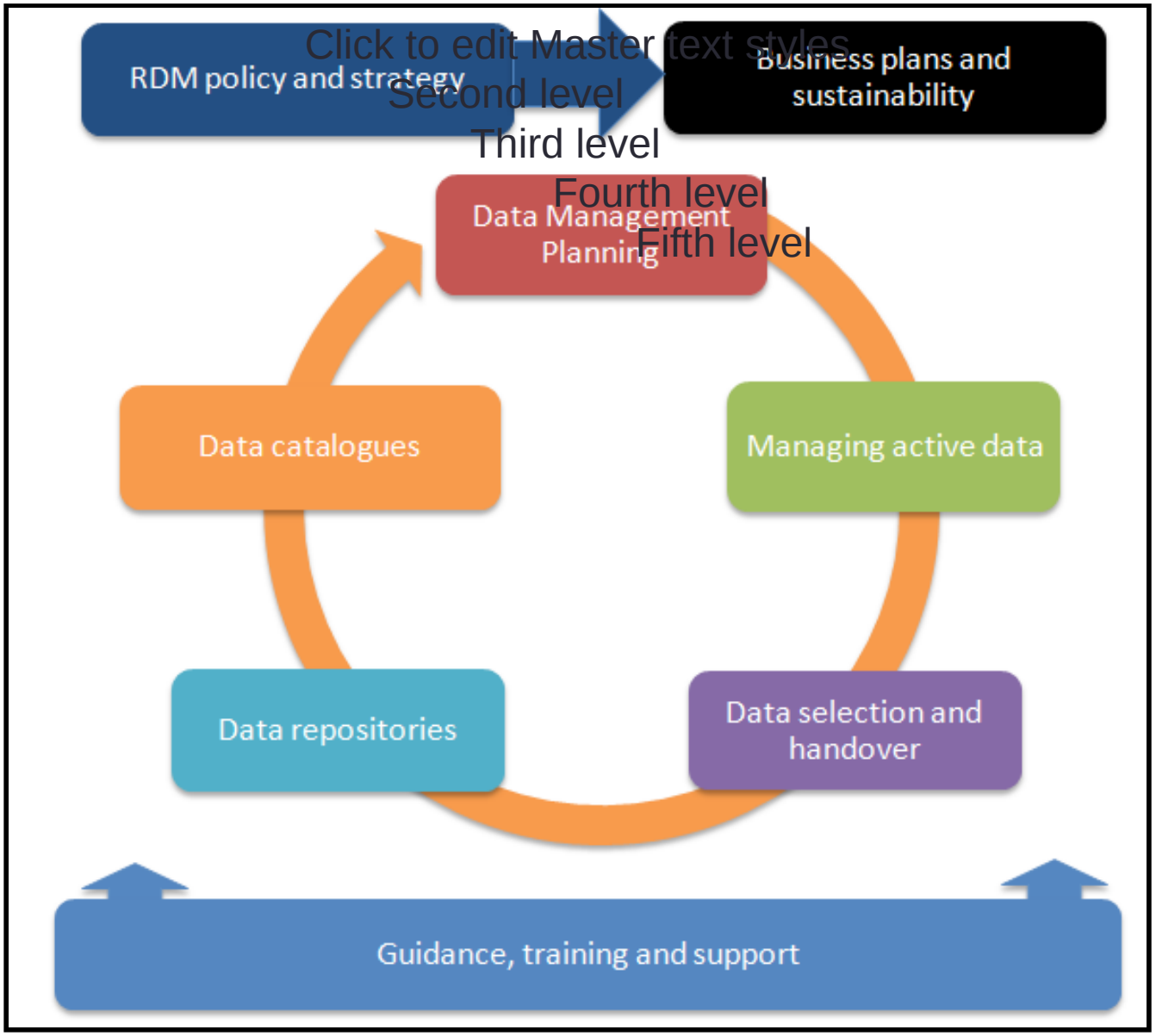


http://4cproject.eu/

Lessons we can share

FOSTER

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
1.1 Develop the data management skills and knowledge of Bath researchers	• Initiate and coordinate a program of research data management (RDM) skills development for Bath researchers • Work with Faculties, DTCs and the Bath Graduate Schools to explore embedding research data management skills development in coursework curricula and professional development offerings	• Pilot RDM skills development to new Bath researchers from September 2013 • High quality training delivered (participant feedback, impact notes)	• **Associate Deans of Research** • **Doctoral Training Centres and Graduate Schools** • **Library/UKOLN/BUCS** • Learning & Teaching Enhancement Office & Researcher Development Unit
1.2 Develop the skills and knowledge of key professional services staff • **Library** • **BUCS** • **RDSO**	• Initiate and coordinate a program of research data management skills development for key professional services staff at Bath • Seek opportunities for professional services staff to network and participate in communities of practice – internally, nationally and internationally • Promote research data management career paths within professional support services where appropriate	• Pilot RDM skills development to key professional services staff from September 2013 • High quality training delivered (participant feedback, impact notes) • Support Bath staff participation in DCC meetings (RDMF, IDCC etc.)	• **Library/BUCS** • **RDSO** • Staff Development Unit • Data Scientist*

http://www.bath.ac.uk/rdso/University-of-Bath-Roadmap-for-EPSRC.pdf

Architecture models

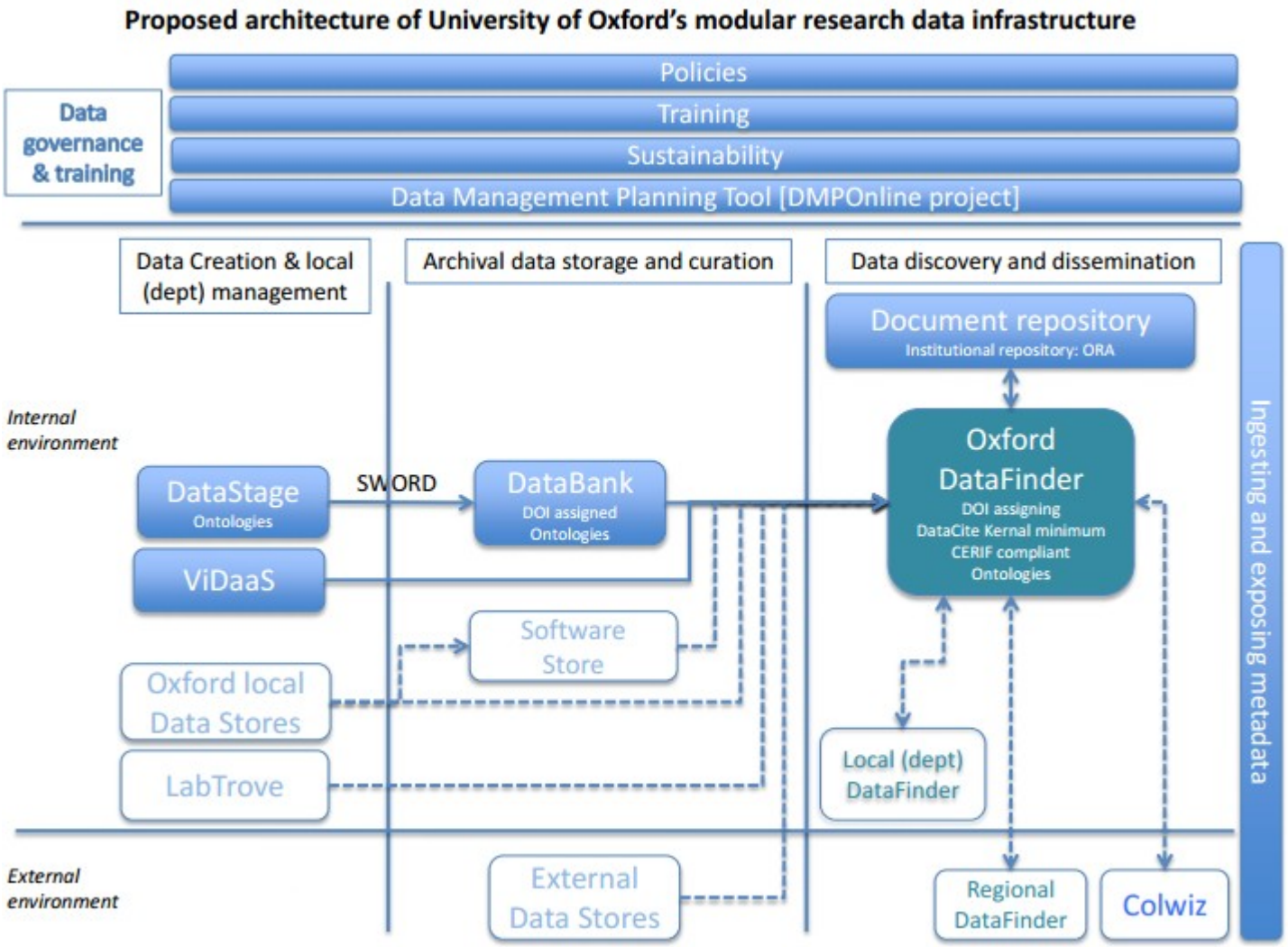


Diagram courtesy of Sally Rumsey, University of Oxford

Thanks – any questions?

DCC guidance, tools and case studies:
www.dcc.ac.uk/resources

Follow us on twitter:
@digitalcuration and #ukdcc

FOSTER

because good research needs good data