

Research data: what is being done at national level

Kevin Ashley
Digital Curation Centre
www.dcc.ac.uk
[@kevingashley](https://twitter.com/kevingashley)
Kevin.ashley@ed.ac.uk



Reusable with attribution: CC-BY



because good research needs good data

Topics

- Policy
- International actions
- UK as example of national & institutional action
- Comparison with other countries

G8UK - Endorses OA Open Data Charter Policy Paper 18 June 2013



Policy paper

G8 Open Data Charter and Technical Annex

Published 18 June 2013

Contents

1. Principle 1: Open Data by Default
2. Principle 2: Quality and Quantity
3. Principle 3: Usable by All
4. Principle 4: Releasing Data for Improved Governance
5. Principle 5: Releasing Data for Innovation
6. Technical annex



Kevin Ashley - Austrian e-infra workshop

RCUK policy - The 1-minute version

- Research data are a public good – make openly available in timely & responsible way
- Have policies & plans. Data with long-term value should be preserved & usable
- Metadata for discovery & reuse. Link publications & data
- Sometimes law, ethics get in the way. We understand.
- Limited embargos OK. Recognition is important – always cite data sources
- OK to use public money to do this. Do it efficiently.

EPSRC policy points

- Awareness of regulatory environment
- Data access statement
- Policies and processes
- Data storage
- Structured metadata descriptions
- DOIs for data
- Securely preserved for a minimum of 10 years **from last use**

Compliance
expected by 2015

Elsewhere...

- Similar policies in Canada, USA, Germany, Netherlands, EC, Norway
- Also charitable funders
- Differ in detail and in compliance mechanisms

Policies at lower level

- University policies becoming common
- Make clearer how researchers are supported locally
- In UK – Concordat between universities & funders defines points of agreement, responsibility

Lessons on policy

- Funders more effective when university is a clear target of the policy (e.g. EPSRC)
 - Almost all just place responsibility on principal investigator
- Sanctions not necessary – yet
- Must be able to show that compliance can be tracked
- See Wellcome OA as an example


because good research needs good data

[Accessibility](#)

Do you have 5 minutes to let us know what you think of this website? [Take part in our](#)

[Home](#)
[Digital Curation](#)
[About Us](#)
[News](#)
[Events](#)
[Resources](#)
[Training](#)
[Projects](#)
[Comm](#)

[Home](#) > [Resources for Digital Curators](#) > Policy and Legal

Policy and Legal

In this section

- Curation Reference Manual
- Curation Lifecycle Model
- Policy and Legal**
- Overview of Funders' Data Policies
- Funders' Data Policies
- Institutional Data Policies
- Policy Tools and Guidance
- Freedom of Information
- FAQs
- MRC Data Plan FAQs
- Open Source FAQs
- Data Management Plans
- Case Studies
- Tools and Applications
- Briefing Papers
- How-to Guides
- Standards
- Publications
- External Resources

Policy resources

[Overview of Funders' Data Policies](#)
A table and short summaries comparing research funders' policies

[Funders' Data Policies](#)
Detailed overview of each funder's policy, stating requirement for data plans, expectations on data sharing and available support.

[Institutional Data Policies](#)
A table listing example of UK universities research data policies.
Add your examples!

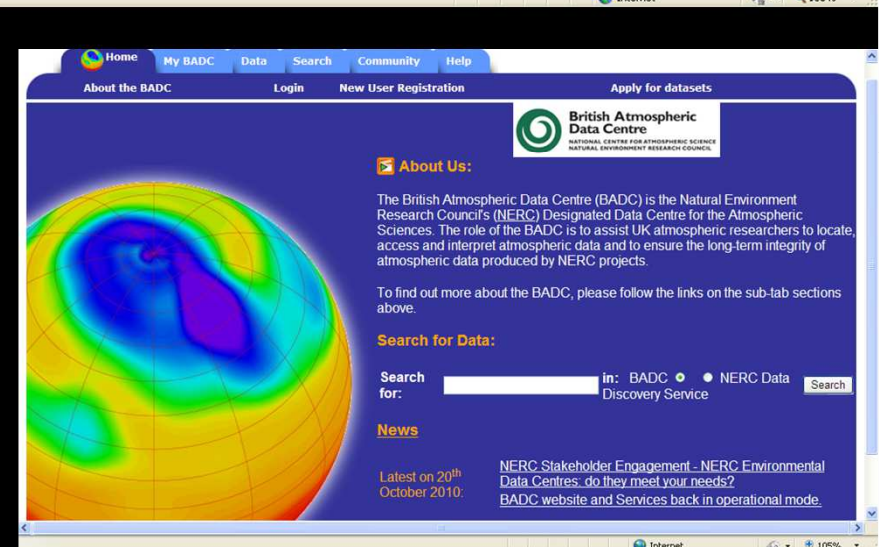
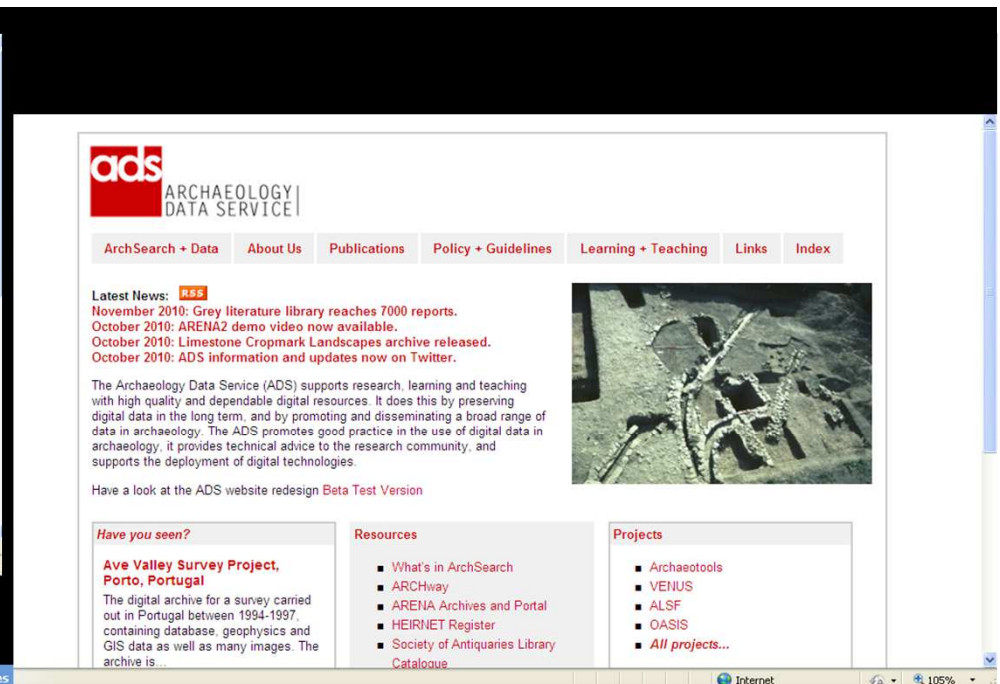
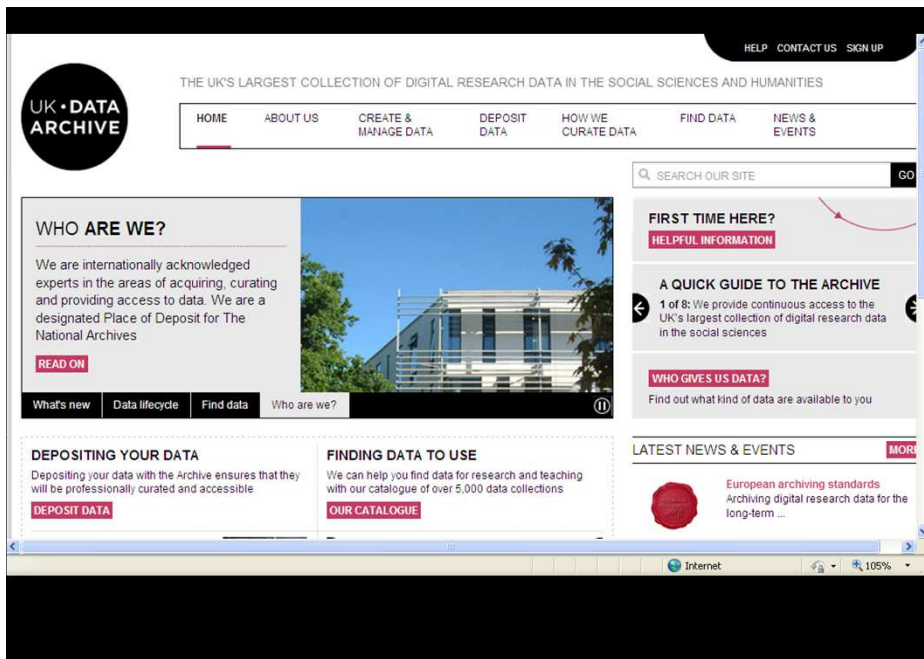
[Policy Tools and Guidance](#)
Annotated bibliography of: 1) tools and guidance for creating policies; 2) example policies; 3) publications; & 4) data management guidance.

[Preservation policy template](#)
Template to help repositories define preservation policies

[Data management plans & DMP Online](#)
Summary of what funders ask for in plans and the DCC's tool to help

DCC Policy Summary

INTERNATIONAL ACTION



Research data centres are good value!

- See Jisc reports on ADS, BADC, UKDA:
- Returns on investment between 400% and 1200%

http://www.jisc.ac.uk/whatwedo/programmes/di_directions/strategicdirections/badc.aspx

- But sustainability is an issue
- OECD funding international study to address this problem

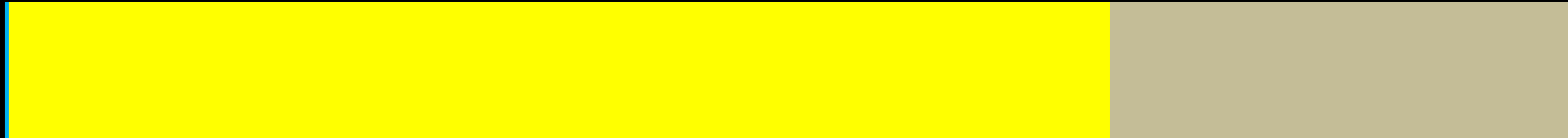
Things that are obvious, things that aren't

- Infrastructure not specific to research data – ORCID, ISNI, Fundref
- Identifiers – Datacite, Handles, etc
- Re3data
- Trans-european services: EUDAT, OpenAire, EOSC
- DMPonline
- RDA outputs – metadata catalogues, data fabric, etc.

THE UK NATIONAL MODEL

2016-10-26

14



164 universities in UK*

71 (43%) > 5% research income

115 (70%) > £1m income from research

*2011 HESA data

DCC provides access to learning resources



www.dcc.ac.uk/resources/



DCC provides access to tools

The screenshot displays the DMPonline DCC website. In the top left corner is the DMPonline logo, consisting of three stacked squares followed by the text "DMP ONLINE". In the top right corner are navigation links: "Home", "About", "Roadmap", and "Help". The main content area on the left includes a "Welcome." message, a paragraph stating "DMPonline has been developed by the Digital Curation Centre to help you write data management plans.", and a video player titled "Screencast on how to use DMPonline" showing a walkthrough of the "Create a new plan" form. The video player has a progress bar at 4:47. On the right side, there are two white panels. The top panel is titled "Sign in" and contains a form with fields for email (pre-filled with "joy.davidson@glasgow.ac.uk") and password, a "Remember me" checkbox, a "Sign in" button, a "Forgot your password?" link, and a link for "Or, sign in with your institutional credentials (UK users only)". The bottom panel is titled "Sign up" and contains the text "New to DMPonline? Sign up today." and a "+" icon.

DMP ONLINE

Home About Roadmap 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

joy.davidson@glasgow.ac.uk

[Forgot your password?](#)

☐ Remember me

Sign in

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

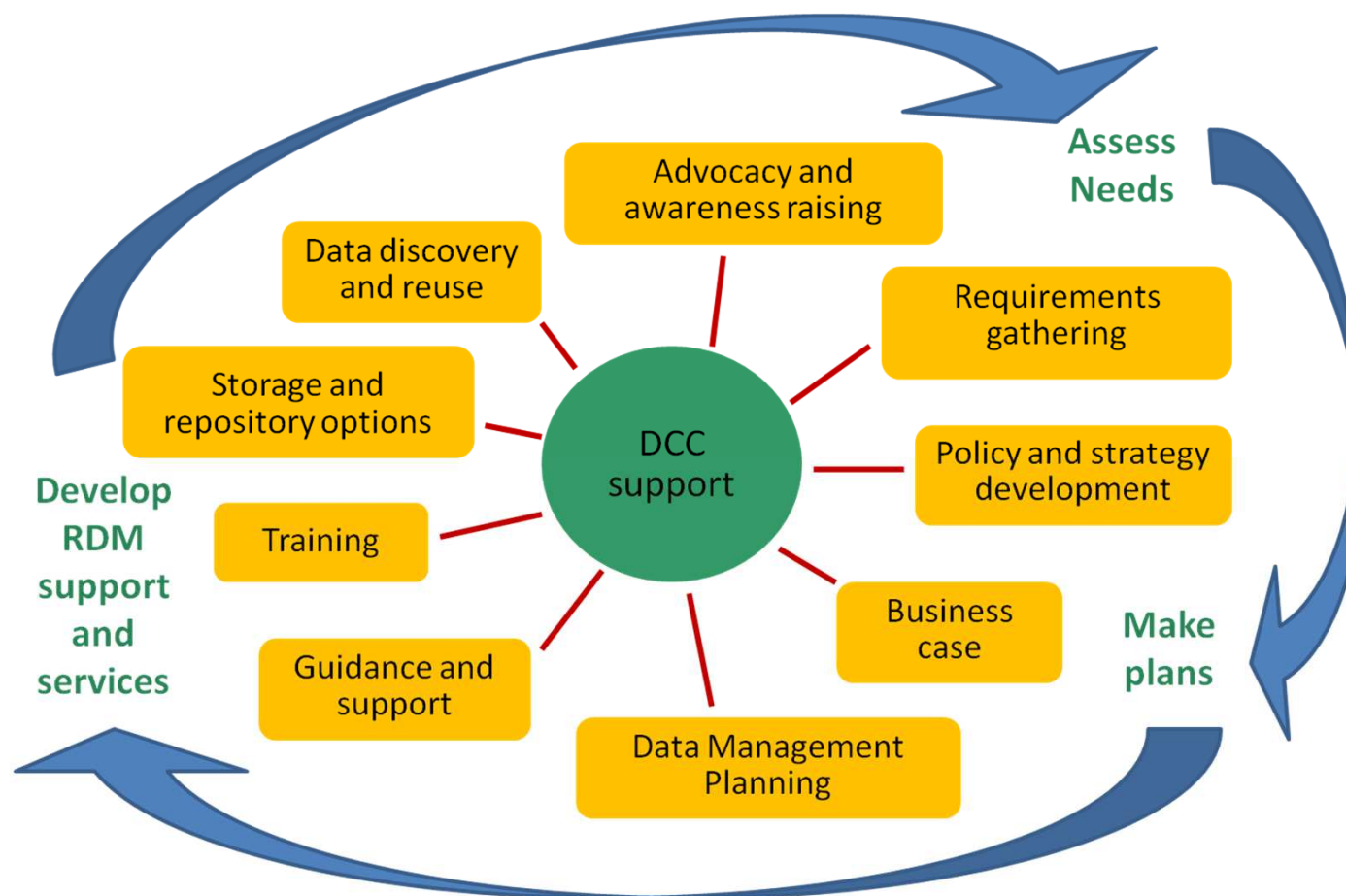
Sign up

New to DMPonline? Sign up today.

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



DCC consultancy services – the Institutional Engagement Programme



www.dcc.ac.uk/tailored-support



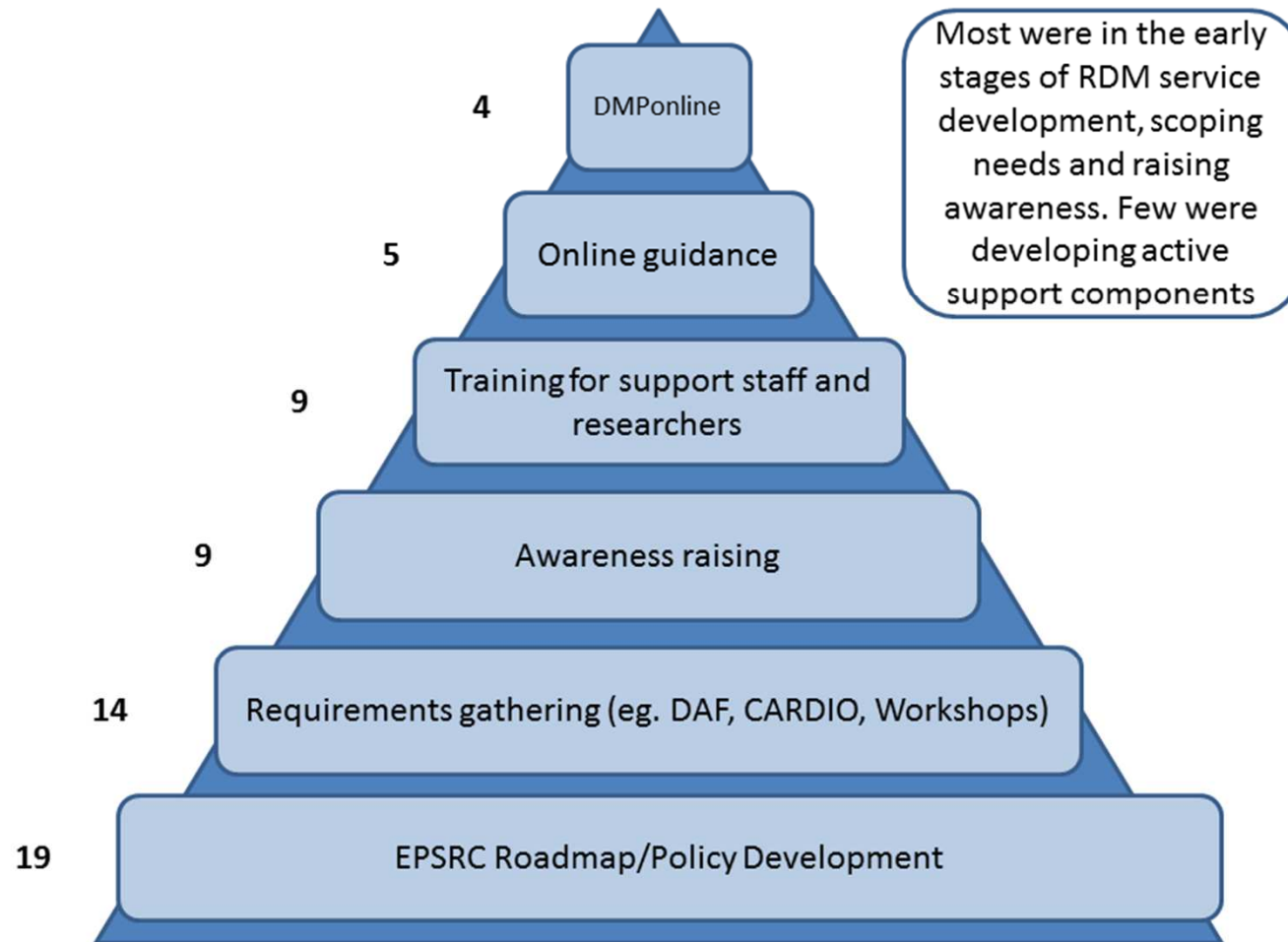
Some unis we worked with



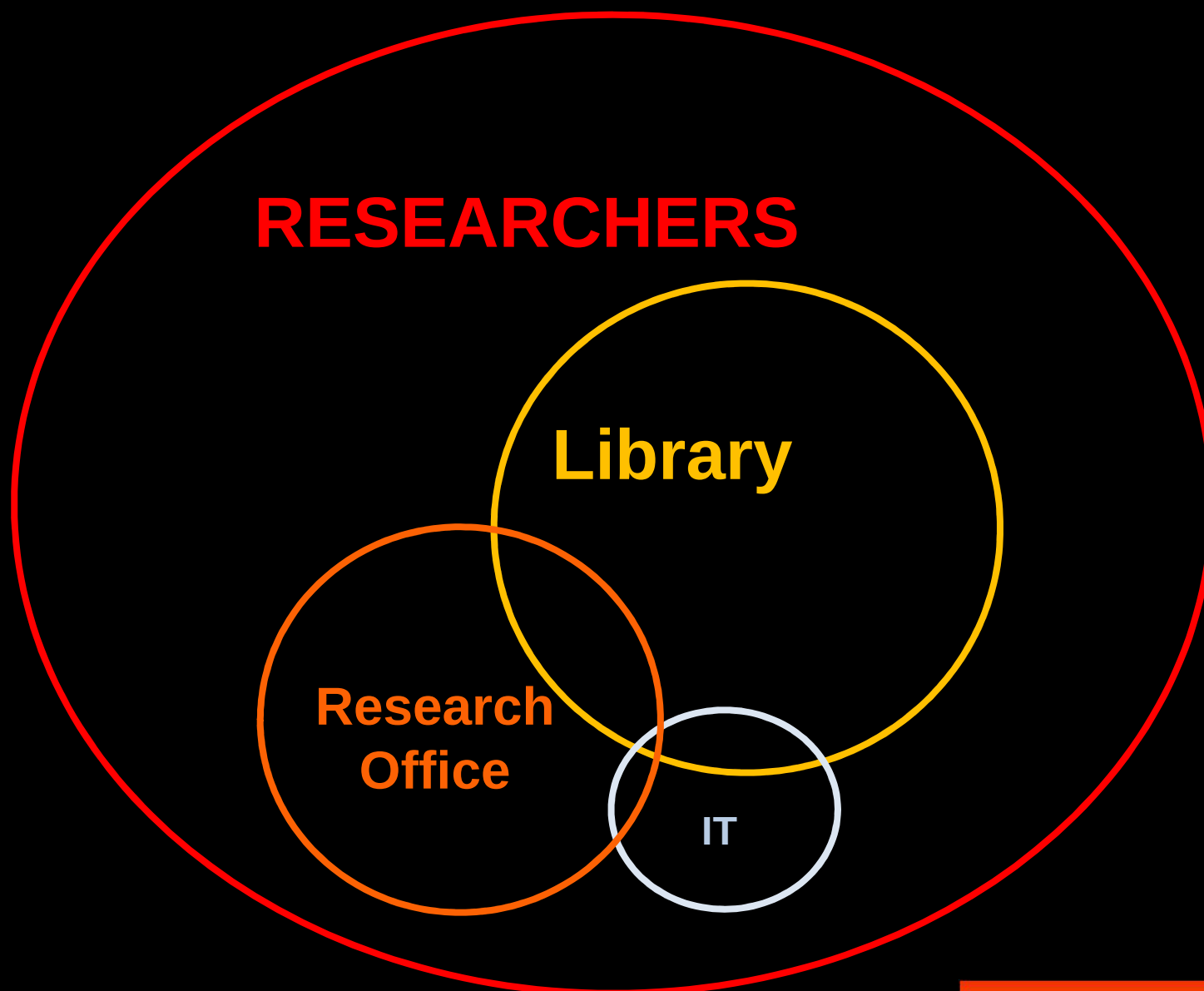
What else was happening?

- Engagement programme driven by report on UKRDS – UK Research Data Service (2010)
- Accompanied by two funded JISC programmes
- JISC funding helped organisations to start activity our advice recommended
- It also supported development of training materials
- Funder policy was being announced

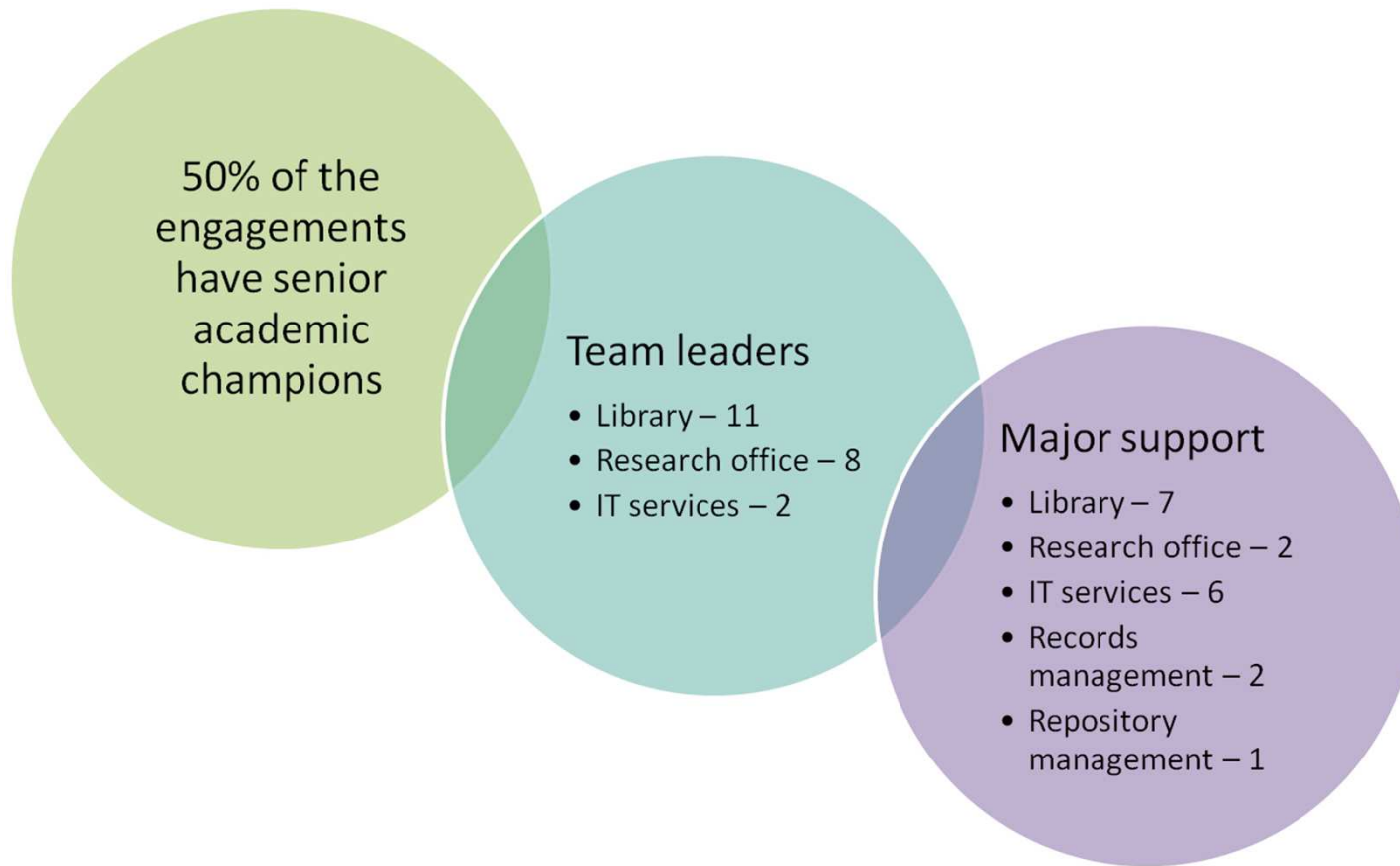
Support in the initial programme



Who (in the UK) is leading RDM work?



Roles of main uni participants

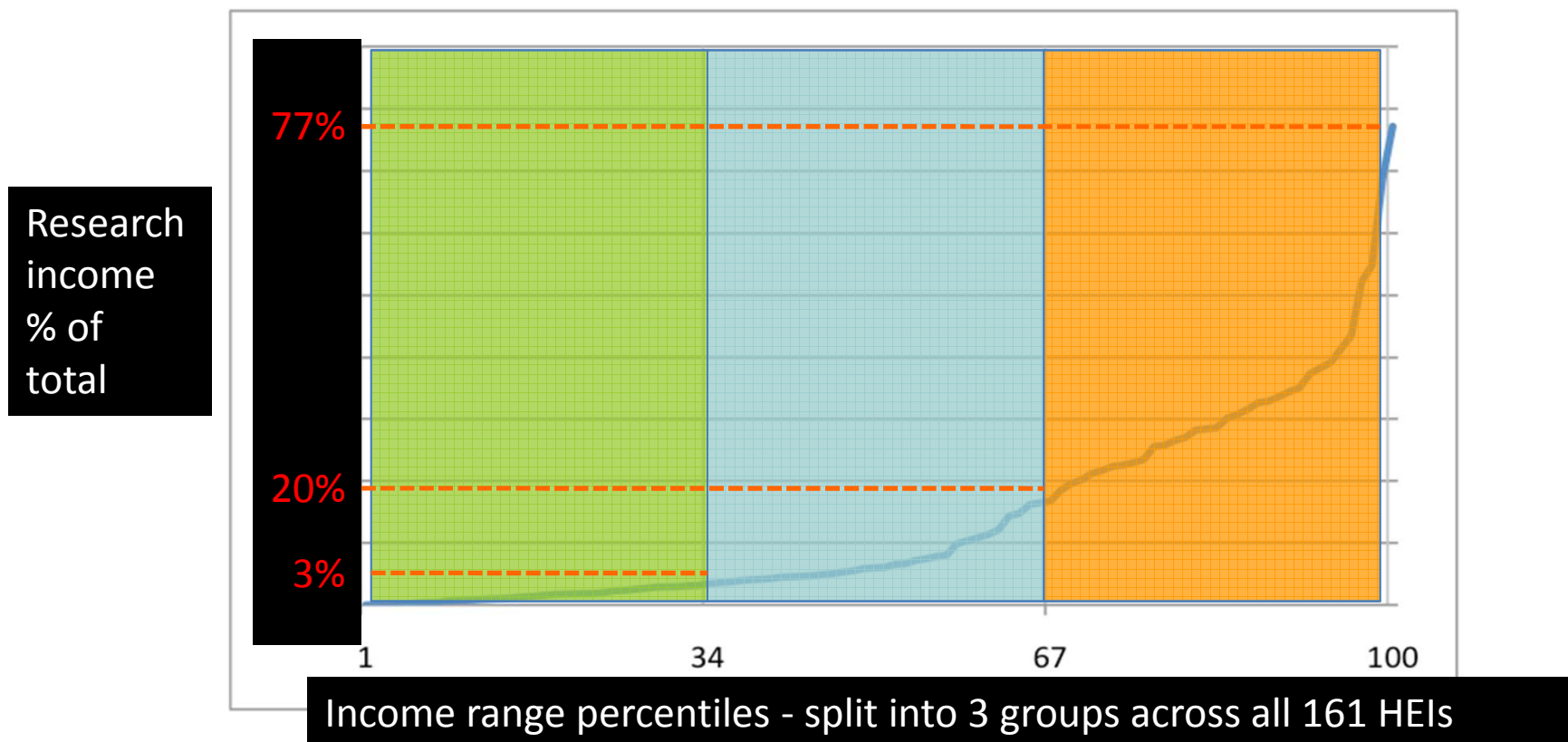
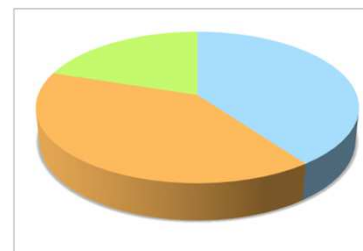


Some institutional roles

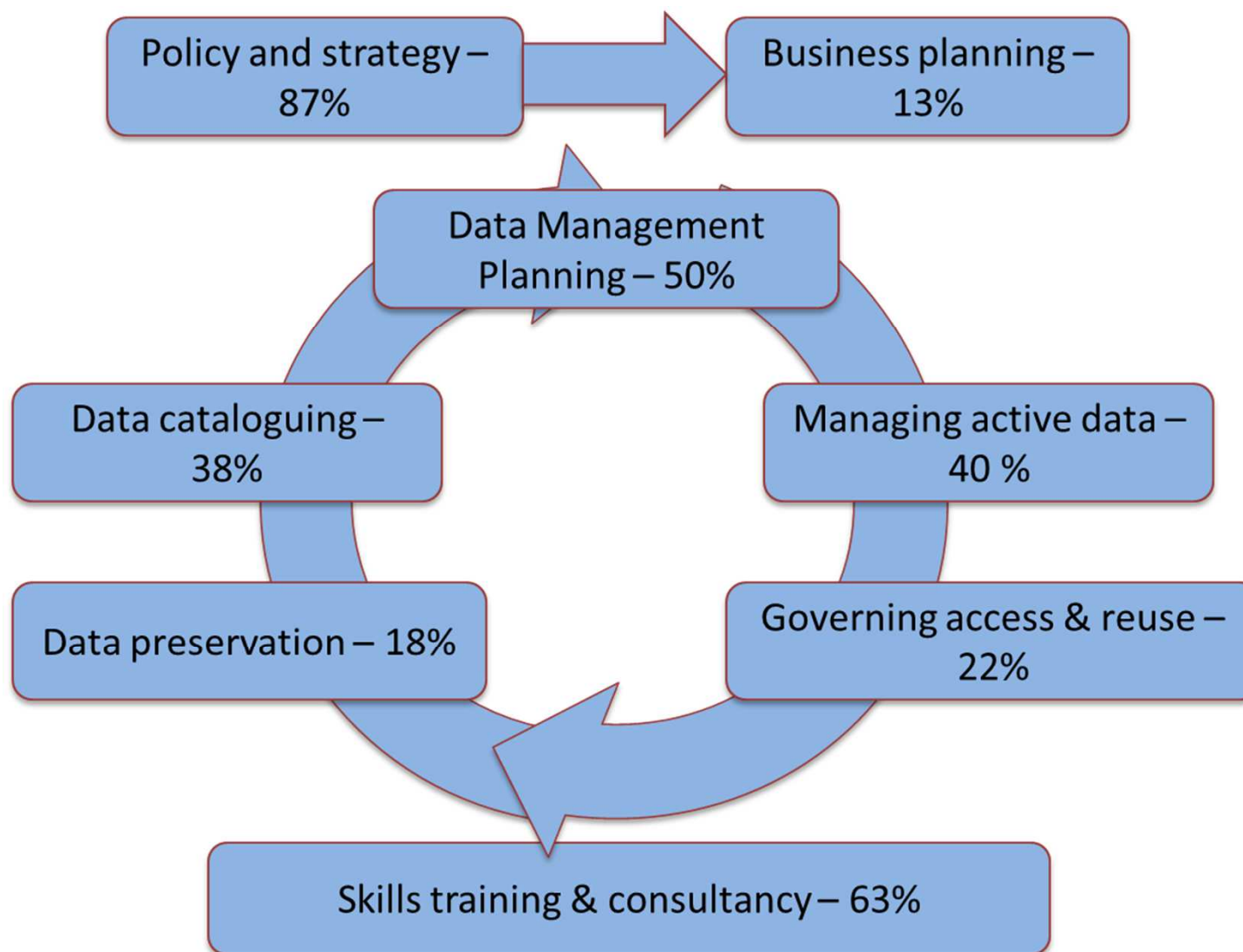
- Leadership – coordinate action
- Audit – who has what, where does it go?
- Advice on access – data, wherever it is
- Preservation – permanence
- Citability
- Data/publication linking
- Promoting data in teaching
- Selection
- Education – early career researchers

Annual RDM survey issued by DCC

- 60 UK Higher Education Institutions responded to DCC survey 2015, of 132 invited
- Research-active institutions well represented



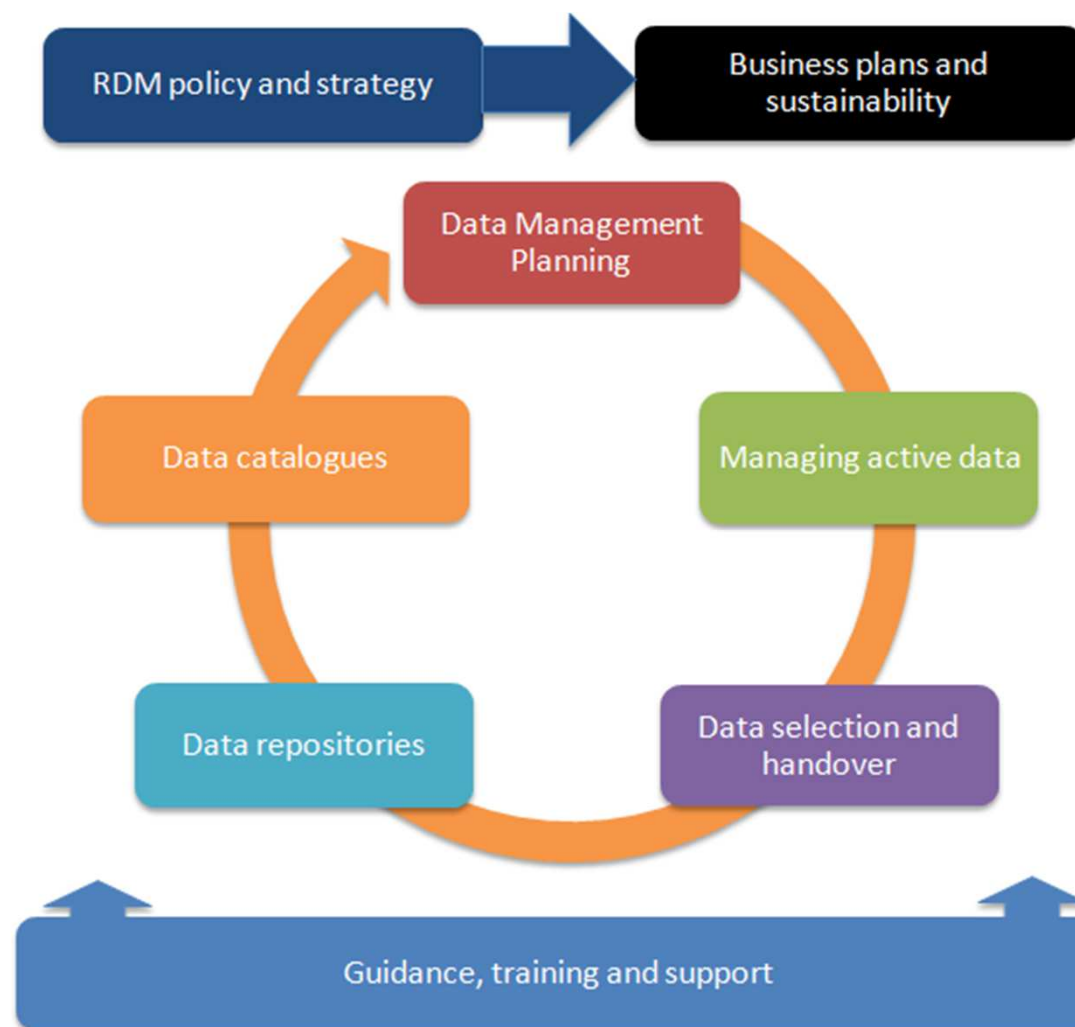
What is the national picture in the UK?



% indicating 'rolling out' or 'embedding'



Components of RDM services



These services can be...

- In-house, for one university
- Shared between a number of universities
- Shared across a country (in development)
- Purchased from a supplier – Figshare, Arkivum, etc.
- Scale is important
- Some things are hard to buy – storage is easy

What data to keep

Roles and Responsibilities

A Digital Curation Centre and Australian National Data Service 'working level' guide



How to Appraise & Select Research Data for Curation

Angus Whyte (DCC) and Andrew Wilson (ANDS)

Researcher ('data creator')

- Provide enough information for others to assess the research data's scientific and scholarly quality and compliance with disciplinary or ethical norms.
- Provide relevant information for the repository to identify who will use the data and how i.e. the 'designated community', and any specific access requirements or constraints.
- Provide the research data in formats recommended by the data repository.
- Provide the metadata requested by the repository.

Data centre or repository

- Make explicit its mission in the area of digital archiving, and its selection policy for digital objects.
- Ensure compliance with legal regulations and contracts.
- Ensure the authenticity and integrity of the digital objects and the metadata.
- Assume responsibility from the data producer for ensuring the digital objects are accessible and available to a defined 'designated community'.
- Plan for long-term preservation of the digital assets.

*"I just back everything up
onto data sticks. I didn't
even know you could
back-up to servers".*

<http://www.flickr.com/photos/mattimattila/3003324844/>



*"Departments don't have guidelines or norms
for personal back-up and researcher procedure,
knowledge and diligence varies tremendously.
Many have experienced moderate to
catastrophic data loss"*

Incremental Project Report, June 2010

SOME OTHER NATIONAL PERSPECTIVES

2016-10-26

31

Australia

- Significant long-term funding for national services & support
- ANDS – data discovery, software, skills, coordination, advice
- National storage & HPC infrastructure
- Financial incentives for universities to use common services in standard ways

Canada

- Like Australia, national research funding but province-level university funding
- Two initiatives – one from universities (PROTAGE), one from federal level (RDC)
- Data management planning, discovery, skills, repositories
- Common tools for local deployment

Netherlands

- Strong national data repository – DANS
- One cooperative service – 3TU (now 4TU)
- Combined to produce RDNL – back office tech services, front office liaison
- National Dataverse instance
- Some shared services now being proposed by SURF-SARA

Portugal

- Existing national publication repository
- Extending to cover research data
- One provider, university-branded front-ends
- Copying other aspects of UK model – e.g. regular meetings for professional staff, funders, other stakeholders

Denmark

- Multi-phase project supported by technical service providers & university libraries
- Governance by them & funders
- Training & shared services
- Still in development

USA

- Scale means many initiatives
- Much is NSF project-funded
- Some existing university collaborations
- Similar spread to UK – 3 or 4 tiers from Ivy League to small institutions
- Very complex funding model
- Technology is useful – models less so

Finland

- National initiative supports Data Management Planning service supplied by DCC
- Training also coordinated
- Central HPC & storage infrastructure for big science

EOSC

- EOSC brings together infrastructure at local, national, international, domain level
- Increasing need for standard service descriptions to enable guided choices, automation
- We all need to track this

Some questions

- What do you provide once?
- What do you develop once, deploy many times?
- What do you just provide guidance on once?
- What is compulsory, what is optional?

DCC guidance

