

# Roadmaps, Roles & Re-engineering: Developing Data Informatics Capability in Libraries

**Dr Liz Lyon,**

Associate Director, UK Digital Curation Centre,  
Director, UKOLN, University of Bath, UK

LIBER Conference Munich, June 2013



This work is licensed under a Creative Commons Licence  
Attribution-ShareAlike 2.0



UKOLN is supported by:



[www.ukoln.ac.uk](http://www.ukoln.ac.uk)

A centre of expertise in digital information management

THE SUNDAY TIMES  
UNIVERSITY  
GUIDE 2013 **3<sup>rd</sup>**



D|C|C

because good research needs good data

Photocredits: Flickr Teamstaufrenberger

# Direction of travel

- Data headlines 2013
- University as a data publisher
- What is the Library data offer?
- Future look towards re-invention

Research360

Managing data across the institutional research lifecycle



UNIVERSITY OF  
BATH



IMMERSIVE INFORMATICS

# Research Data Alliance



Research Data Sharing  
without barriers

<http://rd-alliance.org//>

Working Groups  
Interest Groups



## Community Capability Model WG

The RDA Community Capability Model (CCM) Working Group (WG) collects, validates and publishes a range of data-centric "capability profiles" to enhance inter- and intra-domain interoperability and catalyse RDA data-sharing goals.

1<sup>st</sup> Plenary Gothenburg,  
March 2013  
2<sup>nd</sup> Plenary September  
2013, Washington DC

Open Research Data ... breaking down barriers - RDA Second Plenary Meeting



16-18 September 2013 - National Academy of Sciences, Washington DC, US. Data sharing offers important benefits for scientific progress and advancement of knowledge. However, several limitations and barriers in the general adoption of data sharing are still in place.



**ODI urges government to forge ahead with Shakespeare recommendations**

UK must be "bold and ambitious" in its response to independent review says CEO, Gavin Starks [Read more](#)

COMMISSION FOR INFORMATION  
An Independent Review of  
Public Sector Information  
MAY 2013

UK government supports open data  
Independent review of public sector  
information May 2013

Endorses OA

G8UK

Open Data Charter

Policy Paper

18 June 2013



Cabinet Office

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





## Open access to scientific information: a way to get there

*Open access (OA) = online access at no charge to the user*

- to peer-reviewed scientific publications
- to research data



Policy implications for institutions

# Roadmaps

Planning frameworks



- Policy Framework on Research Data 2011
- Principles and Expectations
- <http://www.epsrc.ac.uk/about/standards/researchdata/Pages/policyfram>



### University of Bath Roadmap for EPSRC

#### Compliance with Research Data Management Expectations

20<sup>th</sup> April 2012, Version 1.1

Authors: Ed Hill, Ewan, J. O'Donnell, A. De Courtenay, P. King, J. King, M.

Status: Submitted to Research Data Steering Group	09 <sup>th</sup> April 2012
Approved, with amendments: Research Data Steering Group	17 <sup>th</sup> April 2012
Submitted to Vice-Chancellors Group (VCG)	25 <sup>th</sup> April 2012
Submitted to VCG with revisions	03 <sup>rd</sup> June 2012
Approved, with amendments by VCG	05 <sup>th</sup> April 2013

#### Acknowledgements

We would like to acknowledge the leadership of Monaghan University in the area of research data management. The Edgworth University Research Data and Storage Plan 2012-2016, published under a CC BY license, was highly influential in the development of this document.

## Institutional Roadmaps

Engineering & Design  
Humanities & Social Sciences  
Management  
Science

# Alignment with EPSRC Expectations

## Roles and Responsibilities: Who's responsible

Where we are now



### EPSRC Expectations

1. Research organisations will promote internal awareness of these principles and expectations and ensure that their researchers and research students have a general awareness of the regulatory environment and of the available exemptions which may be used, should the need arise, to justify the withholding of research data

#### Current position

Research undertaken at the University already complies with all relevant legislation, including data protection, Freedom of Information, copyright, ethical and environmental information regulations. The Research Development Support Office and the Freedom of Information Coordinator provide bespoke guidance to researchers who have reason to withhold their research data. Further attention can be drawn to the regulatory environment, and to reasons for justifying the withholding of research data, by the production of data management plans at the start of new research projects. A new research data management website<sup>2</sup> has already been produced. This contains links to external sources of guidance and further information about the research data management regulatory environment. An output of the current Research360 project is advocacy of research data management across the institution.

#### Proposed Activities

Objectives	Actions	Milestones	Roles & Responsibilities
1.1 Develop the data management skills and knowledge of bath researchers	<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>Associate Deans of Research</li> <li>Doctoral Training Centres and Graduate Schools</li> <li>Library/IKM/IRMC</li> <li>Learning &amp; Teaching Enhancement Office &amp; Researcher Development Unit</li> </ul>
1.2 Develop the skills and knowledge of key professional services staff	<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 DOC meetings (RDMF, IDCC etc.)</li> </ul>	<ul style="list-style-type: none"> <li>Library/SUCS</li> <li>RDSO</li> <li>Staff Development Unit</li> <li>Data Scientists</li> </ul>

**Objectives:**  
Where we need to be

How we're going to get there

- Bath Research Data Sustainability Group (RDSG)
- University stakeholders represented by Heads of Service/Assoc Deans
- <http://www.bath.ac.uk/rds0/University-of-Bath-Roadmap-for-EPSRC.pdf>

**Milestones:**  
When it will be done by

*Photo credits: [http://www.hdwallpapers.in/man\\_of\\_steel-wallpapers.html](http://www.hdwallpapers.in/man_of_steel-wallpapers.html)*

# Roles

New Library data roles too 😊

Transition from pilot RDM projects to sustainable and integrated data services

Articulate business case for RDM investment.....  
but first.....



## University of Bath Research Data Survey results

# Know the state-of-the-nation

- 210 respondents (3.5% response rate)
- **Some preliminary findings**
- Most have not had training (31%)
- Much data is not shared
- agree
- Data is not shared
- Research data is often shared in shared filestore ☺
- They also use USB sticks, home computers ☹
- Data loss due to accidental deletion, hardware failure, obsolescence
- Open data is not the norm – often shared informally
- Lack of recognition for data sharing and reuse is an issue

*These findings mirror most/all institutional data surveys...*

## RDMF special event: Funding Research Data Management

"A conversation with the funders"



25 April 2013

Lakeside Conference Centre, Birmingham

### Overview



Funder panel at the event

# Understand the costs

- A rather grey area.....
- UK Research funder grants
- Direct costs (during lifetime of grant)
- Funder expectations
- Long-term RDM investment an institutional responsibility

# Bath RDM Business Case



## Research 360 Business Case to Support Future Research Data Management in the University

### Revision History

Date	Version	Description	Changed by
13.10.12	1.0	Initial draft based on agreed template	ND
21.11.12	1.1	Updated sections 5 and 8	ND
14.02.13	2	All sections updated	ND
14.02.13	2.1	Additional information provided (new tracked changes)	CLP
17.02.13	2.2	All sections updated	NR
18.02.13	2.3	NU changes accepted, typo & formatting corrections. Minor revisions	CLP
19.02.13		Minor revisions	II
18.02.13	2.4	Executive Summary and minor revisions	NJ
19.02.13	2.5	Minor revisions	LL

- Benefits of investment
- Risks if none
- Options & levels
- Recommendations

## Research360

Managing data across the institutional research lifecycle



- Worked up with Neil Beagrie
- Presented to V-C Group in March 2013
- Positive outcome: 2 f/t permanent posts in Library
- Data Scientist & Technical Data Officer

# Re-engineering....



*Photo Credits: Flickr Chris Walker*

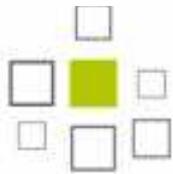
*<http://www.bmwblog.com/2013/06/17/rumor-bmw-8-series-coupe-in-works/>*

## From 300 year old print-based traditions....

## ....to data informatics services & ....data-savvy staff

# Positioning the university as a data publisher

fedora<sup>™</sup>



DSPACE

eprints



ckan



SYMPLECTIC

- Takes responsibility for its data products
- HEI stakeholders described in 2012 (*Informatics Transform*)
- Advocacy gap
- Infrastructure gap
- Reputation?
- Risk?

## Editorial

---

### Data Integrity

I write to alert the organic chemistry community to a serious problem related to the integrity of data being submitted for review and publication by *Organic Letters* and to outline steps that the Journal is taking to address this concern. Recently, with the addition of a Data Analyst to our staff, *Organic Letters* has begun checking the submitted Supporting Information more closely. As a result of this increased scrutiny, we have discovered several instances where reported spectra had been edited to remove evidence of impurities.

Such acts of data manipulation are unacceptable. Even if the experimental yields and

Data integrity: reputations on-the-line?

# Advocacy, Awareness, Training



**Managing Your Research Data**

Catherine Pink (UKOLN)  
Jez Cope (DTC)



The Research360 Collaborative Research Lifecycle Context



[www.bath.ac.uk/riso/datamanagement.html](http://www.bath.ac.uk/riso/datamanagement.html)

- Researcher awareness is low
- Advocacy needed (like OA)
- Change behaviour to “open by default” (G8)
- Open data literacy (G8)



Centre for Sustainable Chemical Technologies



RDM training: Research360 & Doctoral Training Centres

R360 RDM online learning module

# Data Management Plans



2007

2011

# 2013 DMP developments



- New radically simplified DCC DMP Checklist with funder questions – consultation and revision
- New DMP taxonomy
- DMPonline tool Vs3 road-tested with researchers
- Four User Roles with Use Cases
- Vs 4 beta version in September



- Taxonomy as basis for international DMP Profile (CERIF-compliant) development by CASRAI WG

# Bath institutional DMP guidance

## Writing your data management plan

*University of Bath DMP template guidance*

- Introduction .....
- What is data? .....
- Why plan for data management .....
- How to approach data management planning .....

- Completing the plan .....
- Administrative details .....
- Data creation/collection .....
- Data management, documentation and data security .....
- Data archival and preservation .....
- Data publication and access .....
- Roles, responsibilities and resourcing .....
- Further resources .....



### DATA MANAGEMENT PLAN

University of Bath

<b>Principal investigator</b>	
<b>Co-investigator(s)</b>	
<b>Project title</b>	
<b>Funder(s)</b>	
<b>Start date</b>	
<b>End date</b>	

*Many questions include example options for answers - use or delete as appropriate, or replace with your own response.*

### Data creation/collection

**What existing sources of data will be used?**

- |  |  |
|--|--|
| <input type="checkbox"/> Own/group previous research | <input type="checkbox"/> Publicly available database/archive |
| <input type="checkbox"/> Academic collaborators      | <input type="checkbox"/> Specialist commercial data provider |
| <input type="checkbox"/> Commercial collaborators    | <input type="checkbox"/> Other (please specify)              |

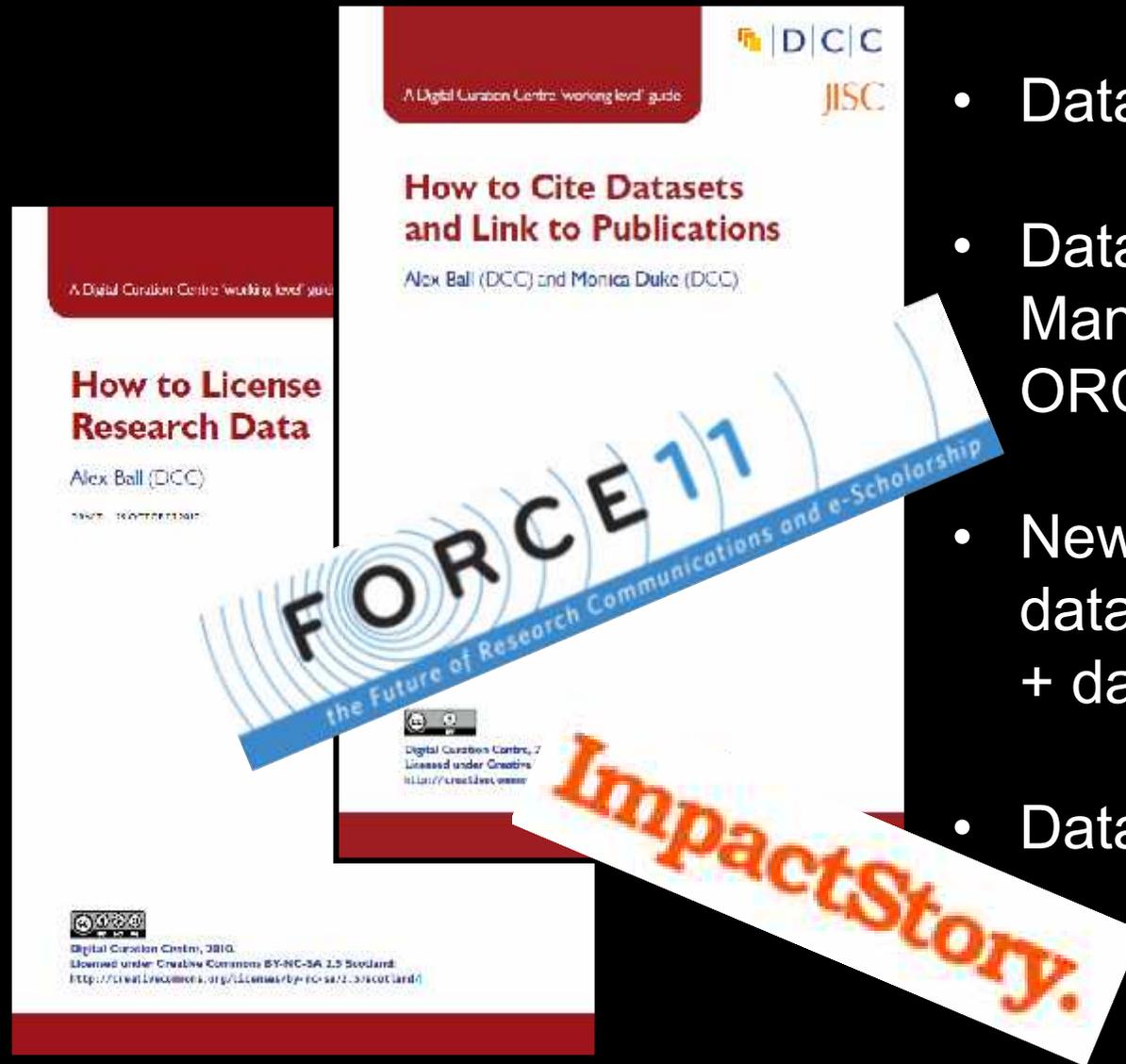
**What are the characteristics of the data?**

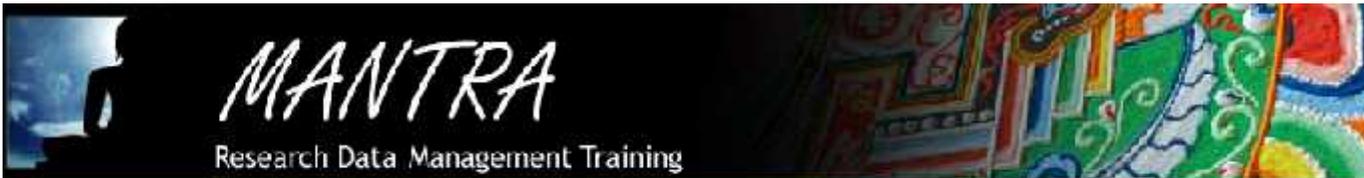
- |                                       |                                 |
|---------------------------------------|---------------------------------|
| <input type="checkbox"/> Quantitative | <input type="checkbox"/> Images |
|---------------------------------------|---------------------------------|

# Template

# Institutional data publication services

- Data repository: curation
- Data catalogue: discovery
- Data citation: Amsterdam Manifesto, identifiers, DOIs, ORCID, data licences
- New actionable formats: data papers, data journals, + data peer review....
- Data metrics: ImpactStory (UK REF2020?)





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

Do-It-Yourself Research Data Management Training Kit for Librarians

<http://www.sheffield.ac.uk/is/research/projects/rdmros>



Information School

# Up-skilling for data

Choice of RDM training  
materials for librarians



<http://dataintelligence.3tu.nl/en/home/>

# Developing data capability in professional services



<http://immersiveinformatics.org>

**Immersive Informatics** pilot at University of Bath  
Co-developed with University of Melbourne



THE UNIVERSITY OF  
MELBOURNE



UNIVERSITY OF  
BATH



# immersiveInformatics

## Questions

What is research data? Why should you share data? How do you analyse data? What is a data journal? How do you cite a data set? What is open access? How do you back up data? What metadata should you keep about data? How is data preserved? What is open data? Who owns data created at the University of Bath? What are the funder requirements for data? How do I find out what data researchers are creating? Can research data be used in the RCT? What is a consent form? How do you license data?

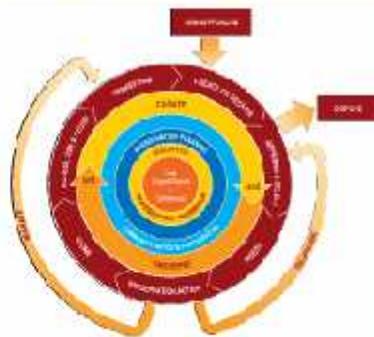


### Bath Programme Team:

- Marieke Guy (DCC)  
m.guy@ukoln.ac.uk
- Catherine Pink (Research360)  
C.J.Pink@bath.ac.uk
- Liz Lyon (DCC)  
l.lyon@ukoln.ac.uk

## Overview

immersiveInformatics is a pilot training programme which offers research data management training for Higher Education professional support services. The programme will run in parallel at the University of Bath and the University of Melbourne and is designed to be transferable to other institutions. Course materials will be stored on Moodle.



The DCC's Data Lifecycle Model will be...

### Immersive Elements

17 modules will feature Immersive Research Department Days (IRDD) which involve shadowing practising researchers. Participants will be aligned to researchers in similar subject areas. During the course of the programme participants will curate and manage a real dataset, documenting the process in a data diary using the Mahara e-portfolio tool. Selection of datasets will be arranged with the academic department in advance. Participants will learn how to describe, manage, identify, access, preserve and reuse the dataset.

Modules will comprise of

- Introductory presentations on the subject area given by immersiveInformatics trainers
- Case study presentations given by expert speakers
- Practical exercises and discussions facilitated by immersiveInformatics trainers
- Details of further resources

immersiveInformatics Website:  
<http://www.immersiveinformatics.org>

## Modular Training

The programme will consist of modular training. The modules will be delivered at approximately one per week in one-day face-to-face workshops. They will be supported with a suite of resources.



Modules 2 and 9 are Immersive Research Department Days that take place in a research department.

July 2013

10 modules (OER)

Day release

Immersive data sessions in labs

Co-curate dataset

Keep "data diary"

# Futures...



*Photo Credits: Flickr solfrost ChrisK*

## Some closing provocations...

# data intelligence hub

1. Every research project has a DMP
2. Library routinely provides DMP support
3. DMPs are harmonised with common core elements
4. DMPs are peer reviewed, stored, analysed and mined for data intelligence by the Library
5. DMP analytics inform institutional planning & validate research infrastructure investments



# Data is mission-critical for Libraries

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

1. Library delivers core RDM services
2. Regional partnerships with collaborative and shared data infrastructure services
3. Demonstrate cost-benefits and efficiency savings
4. Data scientist, data librarian, data analyst teams
5. Transformational data service delivery modes

# Librarians in the Lab



*Photo Credits: Flickr NASA HQ*

1. Co-locate data informatics support at the research coalface fully-integrated in research workflow
2. Data librarians listed as “co-authors” in citations with due attribution and credit
3. Recruit more LIS staff from STEM disciplines for data analysis, data visualisation, data stories
4. Data literacy in core LIS education curriculum
5. Re-invent the Library & re-invent LIS careers...

# Thank you !

DCC Resources

<http://www.dcc.ac.uk>



ImmersiveInformatics

<http://immersiveinformatics.org>

Research360 Project at University of Bath

<http://blogs.bath.ac.uk/research360/>

Informatics Transform article

<http://www.ijdc.net/index.php/ijdc/article/view/210>

