

Data Publishing: Challenges for HEIs and Libraries



Dr Liz Lyon,

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

Now & Future of Data Publishing, May 2013



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

UKOLN is supported by:

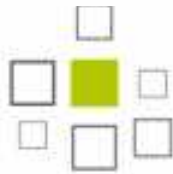


www.ukoln.ac.uk

A centre of expertise in digital information management

Positioning an HEI as a data publisher

fedora^fTM



DSPACE

eprints



ckan



SYMPLECTIC

- Take responsibility for its data products
- A data stakeholder
- Institutional buy-in
- Other priorities REF
- Leadership
- Staff resources
- Infrastructure costs

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

Who pays?

- Research funder grants
- Direct costs (during lifetime of grant)
- Long-term institutional RDM investment
- Need to articulate the business case
- Understand the benefits to the institution

Producing data fit for publication



Managing Your Research Data

Catherine Pink (UKOLN)
Jez Cope (DTC)



www.bath.ac.uk/rdsso/datamanagement.htm



The Research360 Institutional Research Lifecycle

Research360
Managing data across the institutional research lifecycle

- Researcher awareness is low
- Advocacy needed (like OA)
- Change default behaviour: from “mine” to “let’s share”
- Provide RDM training: Research360 Project at Bath
- Doctoral Training Centre role
- Who takes the lead: Library?

Institutional data publishing services needed

- Advocacy & awareness programmes

- Data publishing options: data papers, data journals, data peer review....

- Data repository: “last resort”

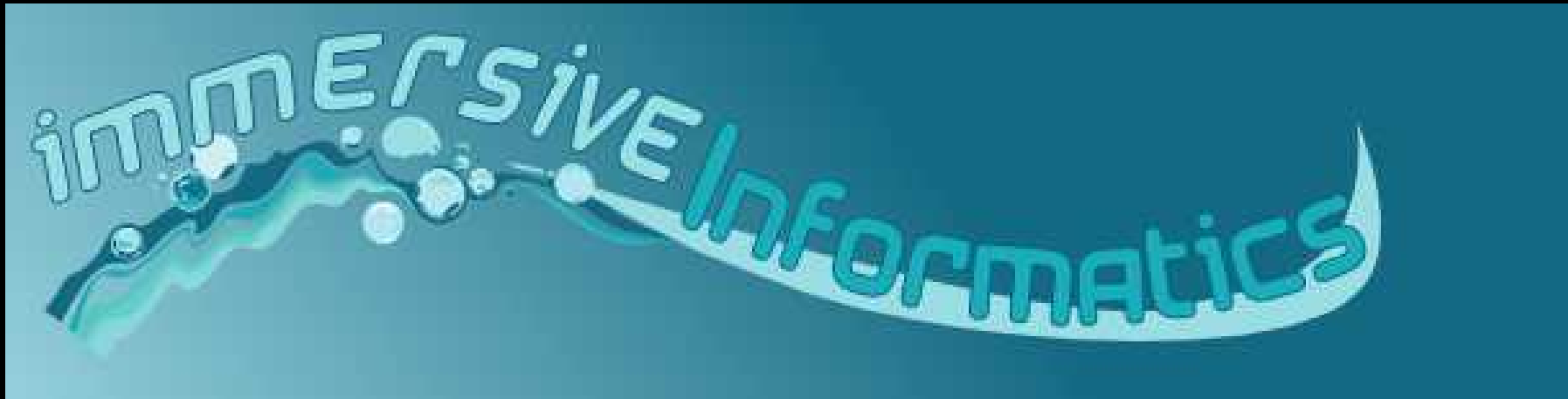
- Data catalogue: discovery

- How to cite data: identifiers, DOIs, ORCID, Amsterdam Manifesto, data licencing

- Data metrics: ImpactStory



Developing data capability in professional services



<http://immersiveinformatics.org>

Immersive Informatics Course pilot at Bath
Co-developed with University of Melbourne
Running from July 2013



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 you find out what data researchers are creating? Can research data be used in the REF? What is a consent form? How do you license data?



Bath Programme Team:

- Marilee Guly (DCC)
m.guly@ukoln.ac.uk
- Catherine Pink (ResearchIST)
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 Curator Lifecycle Model will be explored during module 1.

Immersive Element

Two modules will feature Immersive Research Department Days (RDDs), 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:

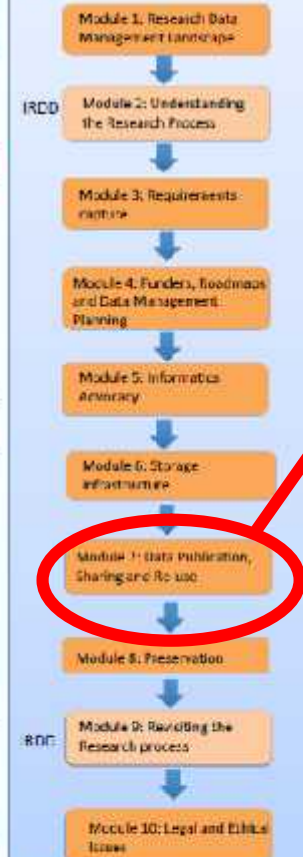
- Introductory presentations on the subject area given by immersiveInformatics trainers
- Case study presentations given by expert speakers
- Practical exercises and discussions facilitated by immersiveInformatics trainers
- Details of further resources

immersiveInformatics Website:

<http://www.immersiveinformatics.org>

Modular Training

The programme will consist of modular training. The modules will be delivered as 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.

Module 7 Data Publication will cover:

- Data sharing
- Data papers & data journals
- Data citation
- Identifiers
- Data licensing
- Metrics...

Incentivising the (open) data publishing workflow

- Researcher incentives: tangible attribution, credit, reward (prizes, promotion, tenure...)
- “REF2020” is potential game-changer: embed metrics for data products, data-sharing, data citations, data re-use
- Interpretations of “impact”: paradigm shifts (personalised medicine), societal engagement (citizen data platforms), crowd-crafting software tools, web-scale impacts
- Incentives for professional data informatics support staff: include as “authors” in citations, recognise as a career path
- Transform the (LIS) curriculum: MSc Data Science

Thank you.

DCC Resources can be downloaded from

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



Immersive Informatics

<http://immersiveinformatics.org>



Research360 Project at University of Bath

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