

# Archiving research data at the University of Bath

Alex Ball

University of Bath

28 September 2017



# University of Bath

- › Established 1966
- › Three Faculties and one School:
  - Faculty of Engineering and Design
  - Faculty of Humanities and Social Sciences
  - Faculty of Science
  - School of Management
- › Roughly 17 300 students
  - of which about 1600 (9%) postgraduate researchers



# Timeline

## 2011 Research 360

- › Active research data storage provision
- › Data management plan (DMP) templates
- › Online guidance

## 2013 Research Data Service, Library

- › University-wide training programme
- › Research data helpdesk
- › DMP checking service

## 2014 Adoption of Research Data Policy

## 2015 Launch of Research Data Archive



## A tale of two systems

---

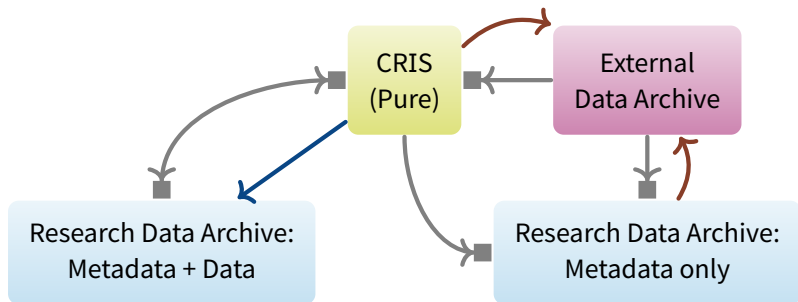
### **CRIS (Pure)**

- ✓ Used for managing REF returns
- ✓ Used for generating CVs
- ✓ Handles links with people, projects, papers, funders, etc.
- ✗ Handles only basic dataset details
- ✗ Limited scope for configuration

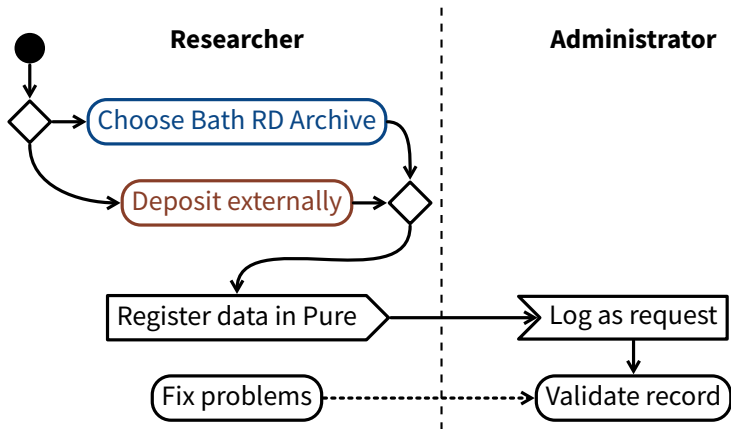
### **Research Data Archive (EPrints)**

- ✗ Not hooked up to University data sources (currently)
  - ✓ Handles rich dataset details
  - ✓ Highly configurable
  - ✓ Archival quality backend storage
-

## Network of dataset records



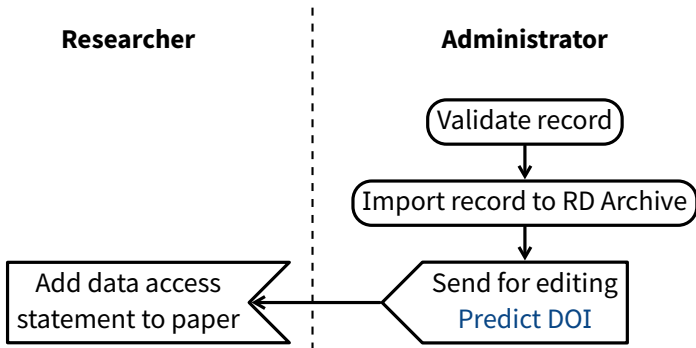
## Archiving process (1)



# Initial validation

- > Title
- > Creators
- > Dates
- > Responsible archive ('Publisher')
- > Project (thereby funder)

## Archiving process (2)



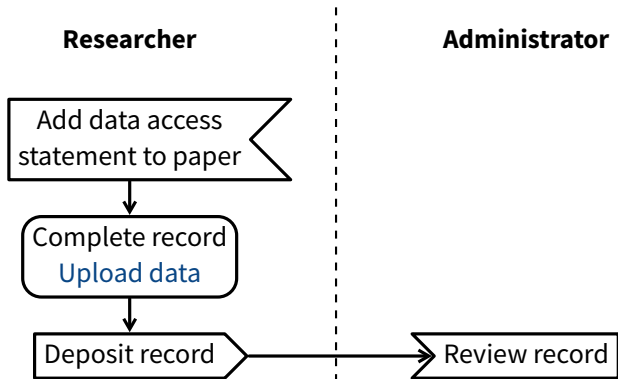


# Using DOIs

We prefer our datasets to have DOIs:

- › Recognized and understood by researchers
- › Seen as ‘mark of quality’
- › Makes datasets easier to cite
- › Makes informal references easier to track
- › Participation in wider infrastructure
  - DCI and OpenAIRE harvest metadata via DataCite
  - Sharing metadata with CrossRef, ORCID

## Archiving process (3)



## Detailed record review

Abstract Is it detailed enough?

Documentation Is it detailed enough?

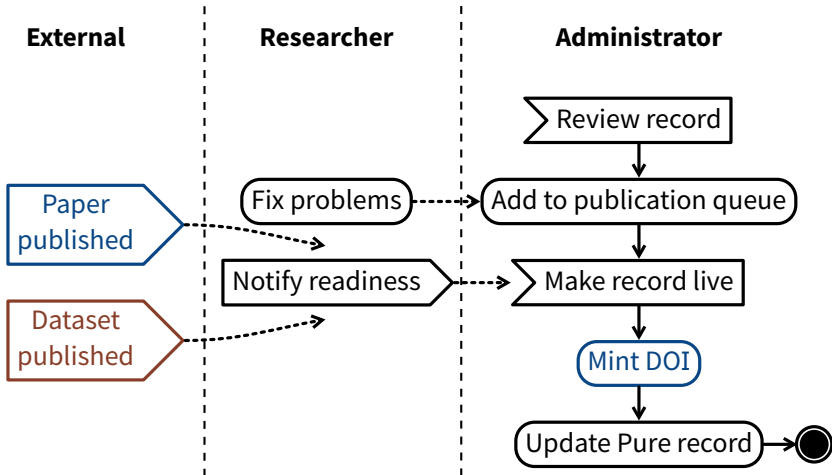
Software Are all specialist packages adequately described?

Sensitivity Do the access settings match the level of sensitivity?  
Is the level of anonymization adequate?

Licences Have they been applied?



## Archiving process (4)



## Providing metadata to DataCite

Creator

Subject

Alternate ID

Title

Description × 3

Format

Publication Year

Contributor

Version

Publisher

Date × 1

Size

Identifier

Geo-location

Rights

Resource Type

Related ID × 5

Language

Funding Ref



# Sharing data

- › Most of our datasets are open access.
- › Datasets with non-digital data or a restricted file have a request button.
- › We can handle embargoes.
- › For handling access requests, we follow instructions from depositors:
  - for some, we have criteria for granting access;
  - for others, we refer the request onwards.

Request Access  
to restricted data

## Advocacy and training

- › We discuss the archiving process at departmental meetings.
- › Several of our training courses deal with archiving:
  - Introduction to Research Data Management
  - Research Data Management: Planning
  - Research Data Management: Archiving and Sharing
- › Our Web pages have detailed instructions for filling out the forms.
- › We provide help by email and face-to-face.



# Outstanding issues

## Near term

- › Allowing researchers to predict their own DOIs
- › Tracking where researchers archive externally
- › Discovering when linked papers are published
- › Checking for data access statements

## Long term

- › Tracking impact of datasets
- › Supporting discipline-specific metadata fully



# Recommendations

We find these work well:

- › Insisting on a good abstract → **Findable, Assessable**
- › Insisting on a good methodology → **Reproducible**
- › Encouraging use of licences → **Reusable**
- › Minting DOIs → **Citable, Traceable**
- › Providing full metadata to DataCite → **Findable, Accessible, Interoperable**

Thank you for your attention