

## Show me the data

### The pilot UK Research Data Registry

Alex Ball<sup>1</sup>   Kevin Ashley<sup>2</sup>   Patrick McCann<sup>3</sup>   Laura Molloy<sup>3</sup>  
 Veerle Van den Eynden<sup>4</sup>

<sup>1</sup>DCC/UKOLN Informatics, University of Bath

<sup>2</sup>DCC, University of Edinburgh

<sup>3</sup>DCC/HATII, University of Glasgow

<sup>4</sup>UK Data Archive

26 February 2014



Except where otherwise stated, this work is licensed under the Creative Commons Attribution 4.0 International licence:  
<http://creativecommons.org/licenses/by/4.0/>

Supported by



# Outline

Motivation

Project overview

Architecture

Collaborators

Metadata

Evaluation

Future

# Repository landscape



# Motivation

- ▶ More funders require research data sharing.
- ▶ ... EPSRC requires that universities ensure research data is preserved and disseminated.
- ▶ ... Universities need to run their own data repositories.
- ▶ Interdisciplinary and multidisciplinary research requires data drawn from diverse sources.
- ▶ Data outputs will contribute to research assessments, so need to be tracked.



Browse by Subject Area



Browse by Map Coverage



Advanced Search

## What's in Research Data Australia



### Collections (92633)

Research datasets or collections of research materials.



### Parties (25467)

Researchers or research organisations that create or maintain research datasets or collections.



### Activities (40674)

Projects or programs that create research datasets or collections.



### Services (184)

Services that support the creation or use of research datasets or collections.

## Spotlight on research data

### N.C.W. Beadle Herbarium

The N.C.W. Beadle Herbarium (NE) at University of New England contains around 90,000 pressed, dried, incorporated and databased plant specimens. The collection includes more than 150 TYPE specimens that anchor scientific names as cited in the original publication of those names. This rich resource contains many collections that are of great interest to local and international researchers. The specimen sheet collection of the N.C.W. Beadle Herbarium is databased and available to registered users for online data entry and data query.



[Explore the N.C.W. Beadle Herbarium Collection through Research Data Australia >>>](#)

# Research Data UK?

## Attractions of the Research Data Australia software:

- ▶ Familiar to project team
- ▶ Proven technology
- ▶ Plays nicely with search engines
- ▶ Displays sample citations and access/rights information up front

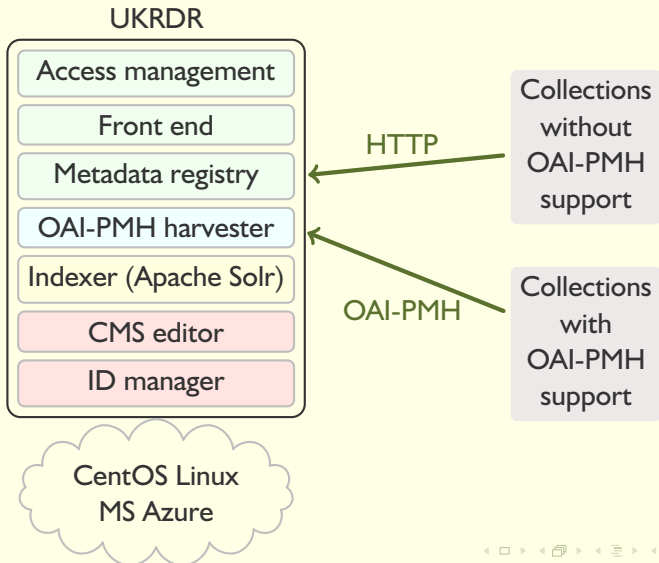
## Challenges of using the software in the UK:

- ▶ Not used before outside Australia
- ▶ Uses uncommon metadata standard (RIF-CS) internally
- ▶ Original implementation only harvests in RIF-CS
- ▶ No UK data centre can output RIF-CS metadata

# Project overview

1. Implement a working instance of the ANDS software.
2. Assemble a group of contributors and establish how their metadata will be harvested.
3. Write crosswalks for transforming contributed metadata into RIF-CS.
4. Harvest metadata from contributors.
5. Reports on
  - ▶ using the Research Data Australia software;
  - ▶ how harvesting from data centres went;
  - ▶ how harvesting from university repositories went;
  - ▶ the value of continuing to develop the registry.

# Architecture





# Collaborators

## 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

# Metadata crosswalks

## 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

# Evaluation

## Evaluation questions

- ▶ Does the software work as intended?
- ▶ Do the harvested records look useful and accurate?
- ▶ Is the system straightforward to use?
- ▶ What might be improved?
- ▶ What additional functions would be desirable?

## For a future evaluation?

- ▶ ROAMEF = **R**ationale, **O**bjectives, **A**ppraisal, **M**onitoring, **E**valuation, **F**eedback

# Questions for a future phase

- ▶ 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?



because good research needs good data

Thank you for your attention

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

UKRDR Pilot Project: [http://www.dcc.ac.uk/  
projects/research-data-registry-pilot](http://www.dcc.ac.uk/projects/research-data-registry-pilot)

This project is funded by Jisc.