



because good research needs good data

# Preparing for the UK Research Data Registry and Discovery Service

Alex Ball

DCC/UKOLN, University of Bath

31 July 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

Vision

Progress so far

What this means for repositories

Metadata

Syndication

Participation

## **Acknowledgements**

**Project team** Kevin Ashley, Alex Ball, Patrick McCann, Laura Molloy,  
Veerle Van den Eynden

**Funder** Jisc

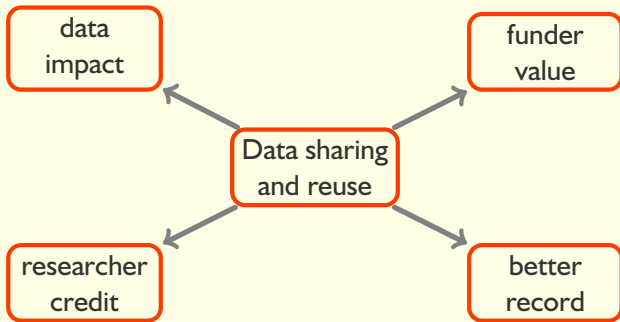


## UK Research Data

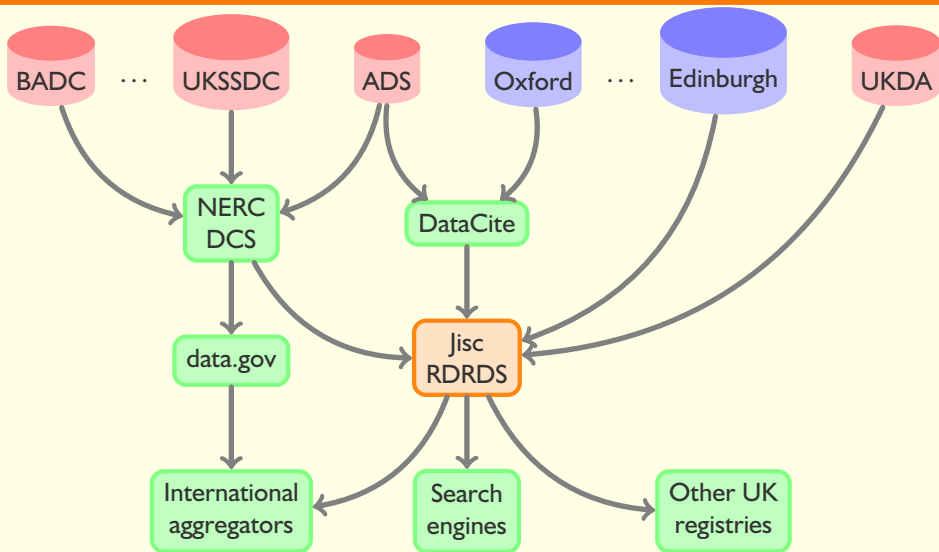
# Discovery Service for UK Research Data



## Discovery Service for UK Research Data



# Where does RDRDS fit in?



# Phase I pilot

- ▶ Registry based on ORCA (Research Data Australia)
- ▶ Participating data repositories:
  - ▶ 9 universities: Edinburgh, Glasgow, Hull, Lincoln, Leeds, Oxford, Oxford Brookes, St Andrews, Southampton
  - ▶ UKDA
  - ▶ Archaeology Data Service
  - ▶ 7 NERC Data Centres: BADC, BODC, EIDC, NEORC, NGDC, PDC, UKSSDC
- ▶ Crosswalks written to RIF-CS *from*
  - ▶ DataCite
  - ▶ DDI Codebook
  - ▶ EPrints (native + ReCollect)
  - ▶ MODS
  - ▶ OAI-PMH Dublin Core
  - ▶ UK Gemini 2



Search for Research Data



Advanced Search

Browse by Subject Area



Browse by Map Coverage



## What's in the Research Data Registry and Discovery Service



### Collections (49)

Research datasets or collections of research materials.



### Parties (36)

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



### Activities (0)

Projects or programs that create research datasets or collections.



### Services (0)

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

## Spotlight on research data

<http://rdrds.cloudapp.net/>

## Who contributes to the Research Data Registry and Discovery Service?

5 research organisations from around the UK contribute information to Research Data Australia.

[See All](#)

 Share

Research Data Australia is an Internet-based discovery service designed to provide rich connections between data, projects, researchers and institutions, and promote visibility of Australian research data collections in search engines. [Read more about us...](#)

ANDS is supported by the Australian Government through the National Collaborative Research Infrastructure Strategy Program and the Education Investment Fund (EIF) Super Science Initiative.





# Hydrographic data profiles collected by a conductivity-temperature-depth (CTD) sensor package during the Jan Mayen cruise JM4

## Full Description

This dataset comprises 73 hydrographic data profiles, collected by a conductivity-temperature-depth (CTD) sensor package, in June 1994 from stations in the North East Norwegian Sea between 69 - 71 N, 15 - 19 E. A complete list of all data parameters are described by the SeaDataNet Parameter Discovery Vocabulary (PDV) keywords assigned in this metadata record. The data were collected by the University of Tromsø Norwegian College of Fishery Science as part of the Ocean Margin Exchange (OMEX) I project.

[SHOW ALL DESCRIPTIONS](#)

## How to Cite this Collection

### Citation (Metadata):


Tande, Kurt ( 2013,2013,2013,2010,2012 ): Hydrographic data profiles collected by a conductivity-temperature-depth (CTD) sensor package during the Jan Mayen cruise JM4. British Oceanographic Data Centre. Local: CSR9662CTDR00147.

[https://www.bodc.ac.uk/data/online\\_delivery/nodb/search/](https://www.bodc.ac.uk/data/online_delivery/nodb/search/)

## Identifiers

Local: CSR9662CTDR00147

## Additional Metadata

URI: <http://csw1.cems.rl.ac.uk/geonetwork-NERC/srv/eng/csw?SERVICE=CSW&VERSION=2.0.2&REQUEST=GetRecordById&ElementSetName=full&outputSchema=http://www.isotc211.org/2005/gmd&Id=b2535c18b9d9554fa24e25e50f3b4a5> 

## Dates

Issued: 2013-07-24 01:00

Created: 2010-02-03 01:00

## Access

<https://www.bodc.ac.uk/data/o...>

### Access rights

Usage restrictions are specified in the terms of the licence

### Access rights

Data are freely available to all following agreement to the terms and conditions of the British Oceanographic Data Centre Data Licence. The licence terms and conditions are available via <https://www.bodc.ac.uk/data/documents/nodb/267795/>

## Connections

### People

Kurt Tande <sup>(P)</sup>

### Organisations & Groups

British Oceanographic Data Centre

## Suggested Links

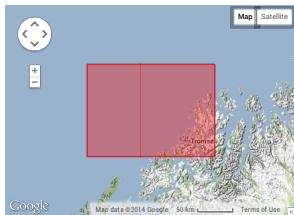
### Internal Records

9 records with matching subjects

### External Records

62 records from DataCite

## Spatial Coverage:



text: Norwegian Sea

## Temporal Coverage:

From 1994-06-13 01:00 to 1994-06-18 01:00

## Subjects

### Keywords

biota	oceans	Natural Environment Research Council Desi
Marine Environmental Data and Information Network	water col	
upper epipelagic water column	water column	bathypelagic
mesopelagic water column	epipelagic water column	
Coordinate reference systems	Elevation	Oceanographic g
Chlorophyll pigment concentrations in the water column	Densit	
Salinity of the water column	Temperature of the water column	
Vertical spatial coordinates		

## Phase 2

- ▶ Define a set of clear use cases and workflows.
- ▶ Compare different possible platforms for the service and assess their suitability.
- ▶ Establish a working instance of the system, involving all UK data centres and university data repositories.
- ▶ Establish a simple workflow for adding more data sources to the service, adapting to changes in existing data sources, and avoiding duplication.
- ▶ Test the system for usability.
- ▶ Produce recommendations for quality and standardisation of metadata records.
- ▶ Evaluate the costs and benefits of the system.



# What does this means for repositories?



# What does this means for repositories?



- ▶ Metadata
- ▶ Syndication
- ▶ Participation



# Metadata

- ▶ Title
- ▶ Description/Abstract
- ▶ Dataset identifier
- ▶ Subject
- ▶ URL of landing page
- ▶ Creator (+ID)
- ▶ Release date
- ▶ Rights information
- ▶ Spatial coverage
- ▶ Temporal coverage
- ▶ Publisher



# Metadata

- ▶ Title ← Required by DataCite
- ▶ Description/Abstract
- ▶ Dataset identifier ← Required by DataCite
- ▶ Subject
- ▶ URL of landing page ← Required by DataCite
- ▶ Creator (+ID) ← Required by DataCite
- ▶ Release date ← Required by DataCite
- ▶ Rights information
- ▶ Spatial coverage
- ▶ Temporal coverage
- ▶ Publisher ← Required by DataCite



# Syndication

## Method

- ▶ OAI-PMH
- ▶ CSW
- ▶ Atom/RSS?
- ▶ Other XML export over HTTP

## Scope

- ▶ Whole data repository
- ▶ Items of type 'dataset'
- ▶ Specific sets or collections

## Metadata format

The more detail, the better!



# Participation

## Get involved by

- ▶ letting us know about issues that concern you;
- ▶ letting us know when you have data records we could look at;
- ▶ helping us test the evolving system:
  - ▶ set up and update an account on the system;
  - ▶ harvest your metadata into the system and check it;
  - ▶ see if we handle duplicates (and non-duplicates) correctly;
  - ▶ see how your records look on the system;
  - ▶ see how easy they are to find;
  - ▶ measure the visibility of your datasets before and after inclusion.







because good research needs good data

Thank you for your attention

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

Jisc RDRDS Project: <http://www.dcc.ac.uk/projects/research-data-registry-pilot>

