



Citation for published version:

Research360 Project 2013, *Institutional Data Repository User Stories*. University of Bath, Bath, U. K..

Publication date:

2013

Document Version

Publisher's PDF, also known as Version of record

[Link to publication](#)

Publisher Rights

CC BY-SA

University of Bath

Alternative formats

If you require this document in an alternative format, please contact:
openaccess@bath.ac.uk

General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

Take down policy

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

Institutional Data Repository

User stories

1 Introduction

A need has been identified to develop an institutional data repository for the University of Bath. This need is driven by changes in funder and publisher policy, as well as responses from the recent Research360 data management survey sent out to all University of Bath researchers. The purpose of this would be to provide a long-term archive of our research data, with the following benefits:

- Ensure long-term availability of data to our researchers;
- Fulfil funder and publisher requirements;
- Enable and track increased impact of our research through data re-use and citation by the wider community;
- Encourage new collaborations and deepen existing relationships with industry;
- Enable new types of research, both within the university and the wider sector.

It is expected that this would be similar in overall operation to Opus, our publications repository, which is based on the EPrints software.

This document sets out our requirements as a University for how this institutional data repository should work.

1.1 Scope

The scope of the repository will be to:

- Record and link to University of Bath research data deposited in external repositories and archives (analogous to Opus);
- Archive and optionally publish snapshots of research data which cannot for whatever reason be deposited externally (usually at the end of a project, but also at key milestones during long-running projects);

The end users of the repository will be University of Bath researchers, from postgraduate students through to academics, and those external researchers and members of the public with an interest in reusing published University of Bath research data.

2 User stories

2.1 As a depositor

I want to...	So that...
Deposit and maintain datasets through a simple web interface	I don't need to install and learn new software to deposit
Have a user interface that is familiar to me	I feel like all the University systems are joined up
Deposit and maintain datasets through Pure	I have a single one-stop shop for managing

	my research outputs
Deposit and maintain datasets through Virtual Research Environments and other workflow tools	I can continue to work with tools with which I'm familiar
Deposit the files that I have	I don't have to spend a lot of time finding the right version and converting to the right format
Place data under an embargo	My right of first-use is protected I can fulfil my confidentiality responsibilities
Apply licenses to datasets	My IP rights are protected appropriately
Allow my collaborators privileged access to datasets	We continue to have a productive relationship
Deposit arbitrarily large files	I am not limited in what files I can and cannot deposit
Link datasets to publications in Opus	Both my data and publications are more easily discovered
Mint DOIs for my data	It can be discovered and cited more easily Citations can be tracked so that I can receive credit
Have metadata automatically filled from other University systems (e.g. Pure) and/or remembered from previous deposits	I don't have to waste time re-entering the same information
Link to data stored in external repositories	I can store my data in an appropriate repository but still register it with the University I don't have to deposit my data in multiple places
Specify a retention/disposal policy for my data	I do not accidentally breach laws or collaboration agreements
Track downloads of my data	I can demonstrate the impact of my work
Track citations of my data	I can demonstrate the impact of my work
Have guarantees about data integrity	I can use my data in the future I can fulfil funder requirements for archival
Attach subject-specific discoverability metadata to records	Researchers in my discipline can find my data more easily
Link datasets with the project DMP (possibly from DMPonline)	Compliance with DMP can be demonstrated Whole project workflow is linked together
Manage and share "live" research data	Whole project workflow is linked together
Manage multiple versions of the same dataset	Changes to the dataset are transparent and do not compromise research integrity
Allow others to deposit on my behalf	I can delegate research data management tasks appropriately

2.2 As a data re-user

I want to...	So that...
Search the archive through the web	I can easily find data relevant to my needs
Access the system in my native language	I am not put off re-using University of Bath data by language barriers
Examine and identify deposited files	I can make a preliminary assessment of usefulness without downloading the whole dataset
View an example citation for a dataset	I can reference it correctly
View a DOI for a dataset	I can get back to the data in future I can import the dataset into my reference-

	management software automatically
Get a persistent URL for a dataset	I can get back to the data in future
Search the archive through Primo (University of Bath library search system)	I can search books, articles and data all in one place
See different versions (including the latest) of a dataset at a glance	I can be sure I'm using the right version of the dataset

2.3 As an external collaborator

I want to...	So that...
Gain privileged access to data for projects in which I am involved	I can collaborate effectively
Have guarantees that my IP rights will not be breached	The risk of collaborating with Bath is acceptable to me
Access data from Bath collaborators off campus	I can collaborate effectively

2.4 As a research facility manager

I want to...	So that...
Deposit data from my facility directly into the archive on behalf of researchers	I am no longer required to maintain my own archive of facility data Researchers can access their own data as needed

2.5 As a Bath Data Archive administrator

I want to...	So that...
Make some checks on deposited datasets before they are made public	Consistent quality of metadata is maintained Compliance with policies can be checked Details of licensing can be checked
Require a minimum set of metadata	Consistent quality of metadata is maintained
Approve scheduled disposal of data	Data which is still required is not destroyed
Query the entire archive (including embargoed records)	I can report on particular aspects of the archive holdings
Import Bath data from an external data centre wholesale	Bath data holdings in external archives are not lost if they close down
Encourage and promote the use of open standards for deposit	Data is as reusable as possible

2.6 As a Research Information manager

I want to...	So that...
Integrate the archive with CRIS	I can analyse impact of research data publication I can link funding to all of the outputs it produces
Include records for externally-held data	The university's record of data holdings is complete
Track citation counts for published datasets	Impact of datasets within academia can be demonstrated
Segment view & download statistics by	Impact of datasets outside academia can be

country and sector	demonstrated
Have datasets linked to metadata about projects	I can report on projects depositing datasets in relation to funder requirements

2.7 As the university IT service

I want to...	So that...
Store archived data on existing storage systems	University data storage is consistent and maintainable Future availability of data can be guaranteed
Integrate the archive with existing university systems such as LDAP	The cost of administering the system can be kept low
Store archived data directly on the HCP object store	Best use of the HCP's features can be made
Be able to export all data to a different system	I am not tied into one system which may not be the most appropriate at some point in the future

2.8 As a developer/maintainer of related services (e.g. VRE)

I want to...	So that...
Deposit and maintain datasets via an API such as SWORD2	My service can interact with the archive

2.9 As an academic publisher

I want to...	So that...
Make persistent web links between my articles and underlying datasets	My journals can be seen to be filled with robust, high-quality research

2.10 As a funding body

I want to...	So that...
Be reassured (by individual researchers or an institution) that researchers I fund have robust archival plans for their data	I can be sure that funding them is a worthwhile investment
Harvest metadata on outputs from research I fund via e.g. OAI-PMH	I can analyse effectiveness of funding strategy I can encourage cross-fertilisation of research outputs (see e.g. OpenAIRE ¹)

3 Acknowledgements

The following people have provided information to support this document.

BUCS: John Howell, Assistant Director (Management Information Systems); Roger Jardine, Assistant Director (Systems & Networking); Paul Jordan, Systems & Database Manager; Simon McLeish, Research360 Systems Developer.

Library: Kara Jones, Research Publications Librarian; Katy Jordan, Faculty Librarian (H&SS) and Research Publications Librarian (maternity cover).

¹ <https://www.openaire.eu/>

RDSO: Katy McKen, Research Information Manager.

UKOLN: Catherine Pink, Institutional Data Scientist; Marieke Guy, Research Officer (DCC), Liz Lyon, Director.

Department of Chemistry: Aron Walsh, Reader; Matthew Davidson, Director (Centre for Sustainable Chemical Technologys)/Head of Department.