Research at Risk: Developing a Shared Research Data Management Service for UK Universities

Rachel Bruce, Deputy Chief Innovation Officer, Jisc – CNI Fall Meeting, December 2015
Jisc is the UK higher, further education and skills sectors’ not-for-profit organisation for digital services and solutions.

We...

Jisc

Operate shared digital infrastructure and services

Provide trusted advice and practical assistance for universities, colleges and learning providers

Negotiate sector-wide deals and conditions with IT vendors and commercial publishers
Our vision and mission

Vision
To make the UK the most digitally advanced education and research nation in the world

Mission
To enable people in higher education, further education and skills to perform at the forefront of international practice by exploiting fully the possibilities of modern digital empowerment, content and connectivity.
Co-design challenges

Research at risk (R@R)  Implementing FELTAG  Digital learning & capabilities

Prospect to alumnus (P2A)  Learning analytics  Business intelligence

Jisc

December 14th - 15th 2015 - Jisc Shared Research Data Management Service 2015
1) Data are a Public Good

Publicly funded research data are a public good, produced in the public interest, which should be made openly available with as few restrictions as possible in a timely and responsible manner.

2) Data should be managed...

3) Data should be discoverable...

4) There may be constraints...

5) Originators may have first use...

6) Reusers have responsibilities...

7) Data sharing is not free...
Royal Society Science as an Open Enterprise

» Recommendation 1
» • “Scientists should communicate the data they collect and the models they create, to allow free and open access, and in ways that are intelligible, assessable and usable for other specialists in the same or linked fields wherever they are in the world…”

» Recommendation 11
» • “…There are legitimate boundaries of openness which must be maintained in order to protect commercial value, privacy, safety and security.”

» Recommendation 9
» • “…proportionate governance…”
» • “…personal data is only shared if it is necessary for research with the potential for high public value”
European Open Science Cloud

» Open in two senses: (Open-Science) Cloud & an Open (Science-Cloud)

» Science = research in all domains, public and private

» Cloud = virtualised, transparent, service oriented,…


» European Free flow of data initiative

» Seamless, interoperable digital services

» Digital ecosystems of hardware, software, applications and data

» Removing technical, legislative barriers to data driven science

» Enabling new services for data-driven science through open systems, services and cross-border flow of data

» Data as a catalyst for economic growth and innovation

» A transition towards more efficient Open Science
“There is typically very little, if any, formal recognition for data outputs in key assessment processes – including in funding decisions, academic promotion, and in the UK Research Excellence Framework.”

**Expert advisory group on data access**

**May 2014**
There is a stick ...

- Awareness of regulatory environment
- Data access statement
- Policies and processes
- Data storage
- Structured metadata descriptions
- DOIs for data
- Securely preserved for a minimum of 10 years from last use
It’s fragmented
We don’t want 150 solutions
Different rates of progress
RD storage capacity

<table>
<thead>
<tr>
<th>Category</th>
<th>&lt;100 TB</th>
<th>100-499 TB</th>
<th>500-999 TB</th>
<th>1PB+</th>
</tr>
</thead>
<tbody>
<tr>
<td>Current active</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Planned active (3yrs)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Current archive</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Planned archive (3yrs)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Research Systems

- Institutional Repository
- CRIS
- File sync & share
- Data registry/catalogue
- Archival preservation systems
- Software repository
Realising a robust and sustainable research data management infrastructure and services to enrich UK research.

Challenge
There is currently an inefficient and fragmented approach to research data management throughout the UK. This leads to underperformance and cost inefficiencies across universities. There is a real risk to universities in failing to comply with government and funder requirements.

Who it affects
>> Researchers need to comply with government and funder requirements
>> IT directors need to store and manage research data
>> Librarians need to ensure other researchers can access the data

Why it matters
>> Ability to attract and retain researchers
>> Reputation of universities can be affected by mismanagement of research data
>> Support better research collaboration (internationally)
>> Comply with requirements of research funders

Success criteria
>> Cost effective national brokered infrastructure as a service
>> Research outputs discoverable and reusable
>> Fewer impediments to doing research
>> Research Data Management is business as usual

Jisc audience
Which members of Jisc's audience will be most interested?

Timescale
How long until we can expect to see the benefits?
1 year - Quick wins, requirements and plan
2 years - Implementation

Commitment
Who in the steering group pledged to commit two days of their time?
Mike Fraser - Director of Infrastructure Services, University of
High level RDM Architecture

CNI Fall Meeting, December 14th -15th 2015 - Jisc Shared Research Data Management Service
Pilot Shared Service

Pilot Shared Service Area

Other R@R Work Areas

Existing Jisc Services/Agreement Areas

Institutional Internal Systems

National and International Systems

Data Management Planning

DMP Creation Tool UI

DMP Registry

Projects Database

Research Outputs Database

Researcher Database

CRIS is sometimes used as the data catalogue/registry

Institutional Data Catalogue/Registry

Archived Data Storage (For Access)

Archived Data Storage (For Preservation)

Data Ingest UI

Active Data Storage

Active Data Management System

Institutional Data Repository

Key

Data Flow

Metadata Flow

UI/Researcher Interface

Research Data Storage

Metadata Store/Database

CNI Fall Meeting, December 14th-15th 2015 - Jisc Shared Research Data Management Service

Credit for Architecture concepts: John Lewis (Sheffield) & Stuart Lewis (Edinburgh)

http://dx.doi.org/10.6084/m9.figshare.1202230
Making it easy for researchers

https://en.wikipedia.org/wiki/Data_sharing
A **key** requirement

**CAUTION**

**PRESERVATION AT WORK**

**MAINTENANCE HELPS**

**PRESERVE OUR HERITAGE**
Interoperability

CNI Fall Meeting, December 14th -15th 2015 - Jisc Shared Research Data Management Service
Researchers shouldn’t need to think (too much!) about Research Data Management

- Provide researchers intuitive, easy functionality to publish, archive and preserve their research outputs.
- Provide interoperable systems to allow researchers and institutions to fulfil and go beyond policy requirements and adhere to best practice throughout the RDM lifecycle
- Preservation!!
Links to research at risk portfolio

Research at risk inputs into RDM shared services pilot

**Shares services**
- UK research data discovery service
- UK ORCID consortium
- Usage statistics for research data
- Journal research data policy registry
- DMP online

**Policy practice and standards**
- Funder policy guidance (EPSRC etc.)
- RDM business case and costing
- Research data metadata
- Research data spring prototypes
Open Access Shared Services Portfolio

Current & Future services
What we need

Front End/ User Interface

<table>
<thead>
<tr>
<th>Ingest UI</th>
<th>Data discovery UI</th>
<th>Data Publication UI</th>
</tr>
</thead>
<tbody>
<tr>
<td>Basic Metadata Entry</td>
<td>Registry/Catalogue search function</td>
<td>Landing page with DOI, Discovery Metadata, and metrics</td>
</tr>
</tbody>
</table>

Middle Layer

- Data Registry/ Catalogue/ Repository
- Archival Management

API’s CRIS, DataCite, ORCID, LOD, funders Etc.

Storage Layer

- Access Data Storage
- Access Data Storage
- Archive Data Storage
- Archive Data Storage

Preservation Layer

- Preservation/ Curation Metadata
- File Format Identification tools
- File/ media migration / transformation tools
- Emulation tools
- Other Preservation/ Curation tools

CNI Fall Meeting, December 14th -15th 2015 - Jisc Shared Research Data Management Service
Lots

1 Research Data Repositories

2 Repository Interfaces

3 Research Data Exchange Interface

4 Research Information and Administration Systems Integrations

5 Research Data Preservation Platforms

6 Research Data Preservation Development and Tools

7 Research Data Reporting

8 User Experience enhancements
Working with a range of institutions based on:

- size, type of institution
- types of data
- degree of RDM readiness
- current institutional systems

Deadline for formal EOI’s 14th December [Today!]
What universities want from the Jisc shared service

- Reduce IT burden
- Preservation
- Integration across systems
- Interoperability with CRIS
- User friendly deposit
- Automation with ELN
- Improve discovery
- Move to centralised services
- Cost effective storage
- Share & develop practice
- No service yet...this will enable that

CNI Fall Meeting, December 14th - 15th 2015 - Jisc Shared Research Data Management Service
» What are your strategies for:
» engagement of researchers?
» preservation?
» software?
» storage?
» text & data repositories combined?
» costing?
» sharing services?
Some further information

» Our research data (etc.) blog - http://researchdata.jiscinvolve.org/wp/

» The shared research data management service requirements:

» Related workshop:

» The expression of interest form for universities and colleges:

» The related tender document for systems & developers can be found here:
https://tenders.jisc.ac.uk/

» Latest research at risk update:
http://researchdata.jiscinvolve.org/wp/2015/12/10/research-at-risk-progress-update/
Acknowledgements

Images:

Clouds & rapeseed: https://www.flickr.com/photos/scandal-princess/4815191034/in/dateposted/

Dog on lead: https://www.flickr.com/photos/cogdog/19666844558/in/photolist-61tiJQ-vXTHGS-5nftYp

Fix the process: https://www.flickr.com/photos/cogdog/19666844558/in/photolist-61tiJQ-vXTHGS-5nftYp

Lot 88: https://www.flickr.com/photos/cogdog/8347605913/in/photolist

Preservation at work: https://www.flickr.com/photos/cogdog/7127177209/in/photolist-bRNDtF-75WVnY

Pottery fragments: https://www.flickr.com/photos/cogdog/9927993383/in/photolist

Share the road: https://www.flickr.com/photos/cogdog/5910598311/in/photolist

Text for European Open Science Cloud from Juan Bicarregui ‘s Westminster Forum December 2015 slides, (STFC Head of Data Division)

Chart of progress for Research Data Management areas on slide 11 from Joy Davidson, DCC, University of Glasgow, 8 May 2014
Rachel Bruce
rachel.bruce@jisc.ac.uk

Twitter
@rachelbruce