Who's paying?

Breaking down the costs of shared services
Do you know how much your Research Data Management (RDM) is currently costing you?

http://sli.do

# SSDP
»Are you sure?
Who we are

Jisc is the UK higher, further education and skills sectors’ not-for-profit organisation for digital services and solutions.

We...

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

Negotiate sector-wide deals and conditions with IT vendors and commercial publishers.

Operate shared digital infrastructure and services.
Who We Are - Jisc

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
Who We Are – Jisc – Futures

“Who pays?”

Technology foresight
Exploring which new technologies offer promise

Discovery
Exploring an idea to see if it is feasible and desirable

Alpha
Developing initial prototypes

Beta
Developing robust prototypes and quickly testing them with Jisc members

Service growth
Delivering the service to Jisc members while scaling it and finalising business model

Service enhancement
Improving an existing experience

Service decommission
Removing a service from the service catalogue
Today’s RDM cottage industry

- c.200 UK public sector institutions perform around £6b of funded research per year in public purse every year
- Only 20% of research data from the 1990s is still usable
- Stand-alone solutions need every institution to manage and maintain on limited budgets
- No sector-wide visibility of RDM effectiveness
- Majority of research institutions want a solution which cuts the pain

Research Funders require data to be managed and EPSRC have put onus on institution to preserve and make research data available as a condition of funding

Jisc shared service

- Share and optimise cost and benefit across the sector
- Grow value of research data and increase research productivity
- A single, cost-effective, pre-integrated solution available to all
**Key researcher RDM issues**

- **Filling a gap**: 75% of respondents look first to their institution to preserve their data.
- **Uptake of RDM**: Only 40% of respondents have a Research Data Management plan.
- **Advocacy**: Only 16% of respondents are currently accessing university RDM support services.
- **Metadata**: Only 18% of respondents say they follow established metadata guidelines.
- **Public datasets**: >70% recognise that research is a public good and should be publicly released.
- **Sensitive data**: 41% of respondents have some form of sensitive data.

**Source:** Jisc DAF Survey results 2016 (https://tinyurl.com/y7ro9pbx)

---

**Advice & best practice**

**Capture & reuse**

**Preserve**

**Report**

---

20/Mar/2018  Who pays?
EPSRC and RCUK have both set out clear expectations that Institutions must take responsibility for Research Data Management, Preservation and Sharing. Institutions must act now - a ‘burning platform’.

### Costs & risks

<table>
<thead>
<tr>
<th>Risk to research funding</th>
<th>Reduced income</th>
</tr>
</thead>
<tbody>
<tr>
<td>Loss of research data</td>
<td>Lost value of research work (17% lost key data – DAF survey)</td>
</tr>
<tr>
<td>Leakage of sensitive data</td>
<td>Legal threat and cost (Unlimited fines with GDPR)</td>
</tr>
<tr>
<td>Researcher reputation</td>
<td>Key staff leave (75% EXPECT HEI to do this)</td>
</tr>
<tr>
<td>Institution reputation</td>
<td>Defensible integrity of research, responding to FOI etc. (e.g. Climategate)</td>
</tr>
<tr>
<td>Cost and risk of delivery</td>
<td>Inefficient research and over-expensive IT</td>
</tr>
</tbody>
</table>

### Consequential loss

- Reduced income
- Lost value of research work (17% lost key data – DAF survey)
- Legal threat and cost (Unlimited fines with GDPR)
- Key staff leave (75% EXPECT HEI to do this)
- Defensible integrity of research, responding to FOI etc. (e.g. Climategate)
- Inefficient research and over-expensive IT
RDSS functional solution at a glance

Jisc Research Data Shared Service

Capture & reuse
- Deposit
- Describe
- Store
- Publish

Preserve
- Assure Integrity
- Normalise
- Transform
- Curate
- Flag at-risk data

Report
- RDM planning
- Costs
- Service performance
- Audit trails
- Compliance and benchmarking

Research data
- Secure outputs
- Accelerate research
- Grow reusable data value
- Managed storage included

Secure managed storage

API’s and Member Dashboards

Open Standard Interfaces (APIs)

Integration with local systems

3rd party tools

Advice & best practice
Research data management toolkit

Cost-effective improvement in research data access and reuse

Shared standards-based technology framework
Who pays?

Pilot Alpha MVP

Jisc alpha repositories
- figshare
- Samvera

User capabilities
- Researcher User Interface
- Deposit Open Data
- Describe
- Cloud Storage
- Publish with DOI and ORCID
- Curate

Initiation support and training

Jisc alpha preservation
- Preservica
- Arkivum

User capabilities
- Research Manager Interface
- Set policies
- Assure Integrity
- Normalise
- Transform
- Curate
- Flag at-risk data
- Archival Storage

Automated data flow

Fully managed pilot service with cloud hosting and archival storage for your data provided by Jisc

20/Mar/2018 Who pays?
Customer side Beta MVP

Jisc

Core RDSS offering

Data storage packages available

Available framework options

Jisc beta repository

Jisc beta preservation

Jisc beta reporting

Fully accredited infrastructure with UK Data centre cloud hosting for your data by Jisc

Jisc research systems framework repositories

Jisc research systems framework preservation

Jisc research systems framework CRIS

Preservica®
Digital Preservation

Arkivum

Full reporting capabilities or APIs available

20/Mar/2018

Who pays?
Service workflow summary

1. Researcher deposits data

2. Data added to aggregation

3. Data is automatically preserved

4. Use of data and service is monitored

5. Other services are updated

6. Researchers find and reuse data

Repository

National research data aggregation

Preservation service

Reporting and analytics

Institutional or external services
Post-Beta Developments

» Integration with Jisc Safe Share
» Integrations with active data storage e.g. Institutional storage area network, file sync and share services.
» Integrations with additional 3rd Party systems
» Integration with Jisc Tiered Storage
» Integration with updated DMP Online
» Certification
3 standard service options

End-to-end service

Repository service

Preservation service

Approach supports pilot investments in a range of components, direction of travel towards standard managed solution for all over time.

All 3 options include:

✓ Financial benefits
✓ Standards
✓ Advisory
✓ Network membership
A shared service can ...

- Save you money
- Reduce risk
- Reduce burden
- Ensure re-use

If you don’t have a plan for RDM now, we can help you to develop one.

If you do have a plan, we can help you to validate, optimise and implement it.

Background...

Jisc has worked with the sector for 2 years to develop a shared service for Research Data Management to enable re-use of research outputs.

16 HEIs and some of the world’s leading vendors joined the Jisc to pilot and develop the service, and our team regularly speaks and engages at global conferences on RDM.

120 UK HEIs attended our Research Data Network meeting in June 2017 – all are investing time and energy in understanding their future obligations and options for managing research data efficiently.
Some of the problems

» New funder mandates

» 10 years from last access

OR
Some of the problems

The Cambrian Explosion...of Data

- Exabytes (billions of GB)
- Years: 2006 to 2020
- Graph showing the increase in Exabytes of data from 2006 to 2020, with a significant increase in 2019 and 2020.

Source: Patrick Cheesman
A big part of the problem

» No existing provision and no budget
» No concept of the true cost
» No inter-communication
The sustainability headache

New need/mandate + no budget + lack of understanding

= sustainability headache
Problems for sustainability

» The funding cliff
» What can I put in my budget?
» Can I recharge?
» How?
» Double dipping
Problems for sustainability

» How much is it going to cost in the future?

» How much does it cost NOW?

1958
“I BET IN THE FUTURE THEY WILL HAVE FLYING CARS”

2018
Problems for sustainability

» What it’s worth?
» What are the risks and benefits?
» How can we measure intangibles?
Problems for sustainability

» Lack of understanding
Problems for sustainability

» The Sustainability Chasm
A big part of the problem

Its **ALL** about the money, money money

(though what Jessie J says)
http://sli.do

# SSDP

Who currently pays for your Research Data Management (RDM)?
In an ideal world...

- Project proposal with realistic RDM budget
- Approved
- Spent
- (pay once keep forever)
- Tracked
- Claimed
What actually happens

- No budget / unrealistic Approved
- Spent
- Over budget
- Under claimed
» Commissioned high level business case
» Commissioned research re costing framework and methodology analysis
» Commissioned research re funder mandates
» Financial X-Ray for Research
Some of the numbers....
Activity Based Costing

1. Determine ALL activities associated with a product/service (direct and indirect)
2. Establish costs for those activities
3. Add it all up and divide by the number of units to get the unit cost
Traditional Costing

1. Determine the direct costs (e.g. Salary)
2. Add the company “standard” overhead (e.g. salary x %)
3. Add it all up and divide by the number of units to get the unit cost

Costing
ABC
More accurate
BUT
Harder to do

versus

Traditional
Easier to do
BUT
Less accurate
What activities?

[Diagram showing the process of Costing, including stages such as Preservation Planning, Data Management, Ingest, Archival Storage, Access, Administration, and interactions with Producer and Consumer.]
Breakdown of RDM by costing activities

Managing active data
- First time R&D, development of best practice
- Acquisition
- Ingest
- Data management and DMP
- Access
  - **Active data storage**
- Generate metadata

Pre-archive phase
- Outreach
- Initiation
- Creation

Archive phase
- First time R&D, development of best practice
- Acquisition
- Disposal
- Ingest
  - **Archive Storage**
- Preservation Planning
- Data Management
- Access
  - **Data publication**
- Generate metadata

Support Services
- Administration

Training and advocacy
- Common services
- Publicity and information dissemination
- Training

Dimensions to consider for each activity
- Cost type(s)
  - (staff/hardware/software/capital costs)
- Nature of cost
  - (fixed/variable/semi-variable/step costs)
- Anticipated or realised costs?
- Activity fulfilled by
- Recurrent costs?
- Timing of activity

20/Mar/2018  Who pays?
All you really need to know for the moment is that the universe is a lot more complicated than you might think, even if you start from a position of thinking it’s pretty damn complicated in the first place.

Douglas Adams
What costs...?

» Staff
» Licencing
» Storage
» Procurement
» Compute
» Set-up
» Training
A shared service helps how...?

» Fully managed
» Economies of scale
» Standards based
  › Certification and API
» Uses cost effective optimised storage
» Integrated
» Reporting
» Network of shared expertise
Range of Alternatives

» Do nothing
» Do it yourself
» Use a fully managed, end to end service
» Costs less than 1% of total research income
» Shared service between 50% and 80% cheaper than DIY
» Practical and auditable cost recovery
Who currently pays for your Research Data Management (RDM)?

Who **should** pay for your Research Data Management (RDM)?

(and who funds the chasm?)
http://sli.do
# SSDP
Who **should** pay for your Research Data Management (RDM)? (and who funds the chasm?)
http://sli.do

# SSDP

Which of the following statements........?
Any questions....