Introducing Jisc

Jisc Research Data Services Context

RDSS context and vision

RDSS timeline and progress

RDSS at St Andrews

Data Model & CERIFication Project

Link on https://www.jisc.ac.uk/rd/projects/research-data-shared-service
Introducing Jisc

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

We...

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.
» Drivers

» More than £5 million investment over 2 years

» Open access

» Sector defined requirements

› “R@R” co-design

› Over half the HEI sector involved

<table>
<thead>
<tr>
<th>Costs &amp; risks</th>
<th>Consequential loss</th>
</tr>
</thead>
<tbody>
<tr>
<td>Risk to research funding</td>
<td>Reduced income</td>
</tr>
<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 £10m?)</td>
</tr>
<tr>
<td>Cost and risk of delivery</td>
<td>Inefficient research and over-expensive IT</td>
</tr>
</tbody>
</table>

RDSS: How and Why?

CRIS2018
RDSS: The Challenge

» To meet the sector’s requirements we need

› True multi-tenant
› Multi content
› Shared data model
› Interoperable
› Sustainable
› Much improved user experience
» Core Architecture
  › multitenant database
  › Interoperability layer
» Data model
» Proof of concept front end API
» Initial Front end design

RDSS: Where we are now

- Jisc Repository Core Infrastructure
- Reporting / Storage
- Tenant systems
- Preservation

CRIS2018
RDSS: Roadmap

CRIS2018
» Pure is our data catalogue & repository

» RDSS pulling metadata and files from API into preservation systems

» We are trialling both Archivematica & Preservica

» Nothing passed back to Pure ... yet
Goal for St Andrews: sustainable digital preservation

» Integrate with our existing systems particularly, Pure - to keep single interface for researchers and rekeying of metadata and transfer of data to a minimum

» Provide a preservation platform/service - integrated with Pure; two-way – preservation status back into Pure

» Solution that is flexible e.g. loosely coupled integrations - based on standards, to ensure we can swap systems in/out easily

» Solution that works for other digital content e.g. university records, building plans, e-theses, digitised special collections
RDSS: Priorities

» First priority is research data
  › Research output (Article/Thesis etc.)
  › Research data
  › Research software/code
  › Provenance metadata (method)

» But also.....
  › Preservation systems tailored for multiple digital objects and data types
  › Use cases and pilots for objects beyond research data

https://creativecommons.org/licenses/by/2.0/
https://www.flickr.com/photos/cogdog/
RDSS: Modelling interoperability
RDSS: CERIFication project

University of St Andrews

CTU
CZECH TECHNICAL UNIVERSITY IN PRAGUE
Project goals

» RDSS logical data model mapped to CERIF logical data model, including full documentation.

» Specific use cases of RDSS, or related services, mapped to CERIF-XML and accompanying guidelines for use.

» CERIF model feedback to euroCRIS and consideration in enhancements to the standard.

» Engagement via workshop/webinar(s) to disseminate outcomes from project.
<table>
<thead>
<tr>
<th>RDSS class</th>
<th>CERIF entity</th>
</tr>
</thead>
<tbody>
<tr>
<td>Person</td>
<td>Person</td>
</tr>
<tr>
<td>Organisation</td>
<td>OrgUnit</td>
</tr>
<tr>
<td>Project</td>
<td>Project</td>
</tr>
<tr>
<td>Grant</td>
<td>Funding</td>
</tr>
<tr>
<td>Collection</td>
<td>Service</td>
</tr>
<tr>
<td>Object</td>
<td>ResultPublication</td>
</tr>
<tr>
<td></td>
<td>ResultProduct</td>
</tr>
<tr>
<td></td>
<td>ResultPatent</td>
</tr>
<tr>
<td></td>
<td>Event</td>
</tr>
<tr>
<td></td>
<td>Event</td>
</tr>
<tr>
<td></td>
<td>Equipment</td>
</tr>
<tr>
<td>File</td>
<td>Medium</td>
</tr>
</tbody>
</table>
### RDSS canonical data model version 2.0 mapping to CERIF

<table>
<thead>
<tr>
<th>RDSS model element</th>
<th>CERIF model element</th>
<th>Notes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Property objectUUID</td>
<td>Entity Attribute</td>
<td>ResultProduct.id</td>
</tr>
<tr>
<td>Property objectTitle</td>
<td>Entity MlAttribute</td>
<td>ResultProduct.Name</td>
</tr>
<tr>
<td>Property objectPersonRole: PersonRole [1..*]</td>
<td>Foreign key ResultProduct_Person</td>
<td>ResultProduct.id</td>
</tr>
<tr>
<td>Property objectDescription</td>
<td>Entity MlAttribute</td>
<td>ResultProduct.Description</td>
</tr>
<tr>
<td>Property objectRights: Rights [1..*]</td>
<td>Foreign key ResultProduct._class</td>
<td>ResultProduct.id</td>
</tr>
<tr>
<td>Property objectDate: Date [1..*]</td>
<td>Foreign key ResultProduct._class</td>
<td>ResultProduct.id</td>
</tr>
<tr>
<td>Property objectKeywords: String [0..*]</td>
<td>Entity MlAttribute</td>
<td>ResultProduct.Keywords</td>
</tr>
<tr>
<td>Property objectCategory: String [0..*]</td>
<td>Foreign key ResultProduct._class</td>
<td>ResultProduct.id</td>
</tr>
<tr>
<td>Property objectResourceType : resourceType</td>
<td>Foreign key ResultProduct._class</td>
<td>ResultProduct.id</td>
</tr>
<tr>
<td>Property objectValue : objectValue</td>
<td>Foreign key ResultProduct._class</td>
<td>ResultProduct.id</td>
</tr>
<tr>
<td>Property objectIdentifier : Identifier [1..*] without relationType</td>
<td>Entity FederatedIdentifier (ixed to this ResultProduct)</td>
<td>Identifier type expressed by classifying the FederatedIdentifier using the RDSS Identifier Types class scheme.</td>
</tr>
<tr>
<td>Property objectRelatedIdentifier : Identifier [0..*] with relationType</td>
<td>Foreign key ResultProduct._class</td>
<td>ResultProduct.id</td>
</tr>
<tr>
<td>Property objectOrganisationRole : OrganisationRole [1..*]</td>
<td>Foreign key ResultProduct_OrgUnit</td>
<td>ResultProduct.id</td>
</tr>
<tr>
<td>Property objectPreservationEvent : PreservationEvent [0..*]</td>
<td>Foreign key ResultProduct_Event</td>
<td>ResultProduct.id</td>
</tr>
<tr>
<td>Property objectFile : File [0..*]</td>
<td>Foreign key ResultProduct._medium</td>
<td>ResultProduct.id</td>
</tr>
</tbody>
</table>
An example of data model economy

Alignment with CERIF resulted in handling session (authentication) metadata elsewhere and keeping the CDM about core metadata fields to aid interoperability with CRISes.
RDSS Information Interchange Use Cases

CRIS

- Search for the Collection
- Add it if not found
- Read/Update context
- Organisations, Persons, Projects, Grants
  - Deposit objects
  - Deposit files

RDSS

- Usage statistics
- Notable preservation events
Thank you!

Anna Clements akc@st-andrews.ac.uk @annakclements
Assistant Director Library Services (Digital Research), University of St Andrews

Jan Dvorak jan.dvorak.2@cvut.cz
CRIS specialist, Computing and Information Centre, Czech Technical University

Dom Fripp Dom.Fripp@jisc.ac.uk @Domicus
Senior curation metadata developer, Jisc