A research data catalogue tool for Newcastle University

iridium project, Newcastle University

Outline
• How does a research data catalogue (RDC) facilitate research data management (RDM) policy implementation?
• How was a proof-of-concept RDC developed?
• What is required to support institutional culture change in managing research data at Newcastle University?

Background
• Funding Councils increasingly expect institutional management of research data and for that data to be made discoverable, while the institution recognises a need to manage research data assets
• the RDC supports the draft policy principles (such as responsibilities, data location, data access mechanisms linked to publications, internal availability/discoverability, curation mechanisms and embargo periods) and their implementation
• existing institutional research information management (RIM) functionality and metadata was utilised to provide an integrated RDM system with a searchable web-based interface

Technical development of proof-of-concept
• resulting from institutional and funder requirements; prototyping, iterative development from tester feedback, further testing and follow-up development work was conducted
• the RDC proof-of-concept and roadmap is outlined in Figures A-D

Discussion
• policy, technical and support implications identified to date are reported in Table 1
• RDM policy implementation can be supported by a RDM tool, however significant human support factors are associated with this

Table 1. Implications of implementing a research data catalogue.

<table>
<thead>
<tr>
<th>Policy</th>
<th>Technical</th>
<th>Support</th>
</tr>
</thead>
<tbody>
<tr>
<td>Delegation of editing permissions</td>
<td>Inheriting roles from existing RIM</td>
<td>Validation &amp; 'record cleansing'</td>
</tr>
<tr>
<td>Approval process for updated records</td>
<td>Extensible metadata</td>
<td>User completion rates</td>
</tr>
<tr>
<td>Priorities in connectivity with research output systems</td>
<td>Pre-population/re-use metadata</td>
<td>How much free tagging/QA vocabularies/taxonomies</td>
</tr>
<tr>
<td>Business case</td>
<td>Interoperability</td>
<td>Report generation</td>
</tr>
<tr>
<td>Integration with existing processes</td>
<td>Evidence for ‘real-world’ use of fields</td>
<td>Business case</td>
</tr>
<tr>
<td>Navigation of multiple projects</td>
<td>Metadata options (Triplestore/NoSQL)</td>
<td>Interoperability</td>
</tr>
<tr>
<td>Publication/data set records</td>
<td>Production system performance, security</td>
<td>User uptake/embedding</td>
</tr>
<tr>
<td>Metadata options (Triplestore/NoSQL)</td>
<td>Training</td>
<td>Report generation</td>
</tr>
</tbody>
</table>

References
1. http://research.ncl.ac.uk/iridium (accessed October 2012) and a collaboration of the University Research Office, the Digital Institute, the University Library, Information Systems & Services and MEDEV, School of Medical Sciences Education Development