Research data platforms - finding the right levels of capability for the organisation

Angus Whyte (Digital Curation Centre) Stuart Macdonald (University of Edinburgh) Janet Heslop, Florio Arguillas (Cornell University) Veerle Van den Eynden (UK Data Service)

The problem: mapping needs to platforms
Research data services in higher education institutions and subject-based data archives have a wide range of choice of platforms to meet their organisation’s research data management (RDM) needs. Appraising options can be complicated, as different types of product may offer part of the solution – repository platforms, current research information systems, preservation tools, data sharing and storage-as-a-service offerings. These offer extensive ranges of functions that overlap but can be difficult to align with the task of prioritising the organisation’s requirements.

Narrowing the field
Our high-level selection criteria aim to help Research Data Management (RDM) professionals make the case for service improvements to decision makers, by highlighting capability gaps. These may be gaps in:

-Organisational capabilities, considering changes in the RDM environment and stakeholder expectations
-Data management platform capabilities, considering technical developments and user expectations

Example: CISER
Here the social science domain repository Cornell Institute for Social and Economic Research describes the organisational characteristics that matter most to its mission of anticipating the evolving needs of researchers on campus. The Figure (top) shows current levels of expectation met by CISER’s current data management capabilities (bottom). The data archive’s aspirations for improved functional and technical capabilities are shown (centre). Case studies detailing this and other examples will accompany the DCC How-to guide.

The approach shown here is detailed further in a forthcoming DCC Guide: How-to Evaluate Options for a Research Data Catalogue & Repository. Its focus is on helping a service to narrow the scope of options to develop or enhance its data management platform capabilities. This would then inform the analysis of specific use cases and functional requirements, to match between these with available products and services.

5 Steps to ease the way
The approach involves 5 straightforward steps to identify the organisiation's capability gaps.
1. Choose any organisational criteria likely to influence management choice, from a list of 18 described in tables *
2. For each of those characteristics refer to the criteria statements selecting the appropriate level of expectation statements

<table>
<thead>
<tr>
<th>Level</th>
<th>Minimal</th>
<th>Midrange</th>
<th>Extensive</th>
</tr>
</thead>
<tbody>
<tr>
<td>e.g. Data collection</td>
<td>Service would primarily catalogue data held in external repositories, with limited online access.</td>
<td>Service would hold extensive datasets and index content, serving specialist needs of external depositors.</td>
<td>Service would hold extensive datasets and index content, serving specialist needs of external depositors.</td>
</tr>
</tbody>
</table>

3. Identify whether those expectations are currently met, or planned
4. Select the data management criteria relevant to your organisational needs, from a list of 20, using tables to identify the expected level of capability - minimal, midrange, or extensive
5. Group requirements of similar levels together, if necessary using organisational characteristics to refine use cases, and selected data management capabilities & levels to prioritise functional requirements

* Selection criteria tables derived from literature review are available for comment here: http://bit.ly/rdplatforms

Please tell us…could this approach help you?
1. Are the selection criteria relevant?
2. Are the criteria and tables useful to help scope platform requirements?

Contact Angus Whyte (a.whyte@ed.ac.uk) Stuart Macdonald (Stuart.macdonald@ed.ac.uk) Janet Heslop (heslop@cornell.edu), Florio Arguillas (flaar@cornell.edu) Veerle Van den Eynden (veerle@essex.ac.uk)