Embedding research data management behavioural change within policies, systems and human support infrastructures

Newcastle University
Niall O’Loughlin, Lindsay Wood, Megan Quentin-Baxter, Phil Heslop, Simon Kometa and Janet Wheeler

Introduction
- Management of research data has significant academic and financial implications for the University. Direct research income was £88 million in 2010-2011 and REF associated Quality-Related funding returns £35 million annually.
- Public funders are now mandating that their research data are available for verification and re-use.
- The University is addressing these challenges through the JISC-funded 18 month cross-department iridium project, due for completion in March 2013.
- The initial audit of current research data management status and staff attitudes is complete. The next stage is to incorporate these into outputs i.e. policies, research systems, and training together with final report recommendations and a costed business case.

Methods
RDM requirements gathering: An online survey was distributed to all active Principal Investigators (932 staff), Representative academics and research staff across the institution were invited for a face-to-face interview (163 staff). Transcription and thematic analysis were carried out.

RDM policy framework development: A review of external/internal policies and guidance documentation was conducted. This was mapped to the Digital Curation Centre RDM lifecycle model.

RDM systems development: Current research business systems were reviewed to assess their suitability and capability as data management and catalogue systems. Required functionality was documented.

Results
RDM requirements gathering: The online survey was completed by 128 research projects, representing 15% of active academic staff. The key findings were that:
- 31% of projects’ data storage needs were satisfied by the institutional free allowance (4GB)
- 64% of projects’ data location was serviced by the institution.
- 51% of projects were happy to release their publication-associated data within 1 year.
- 23% of projects had a formal data management plan.

Data retention up to 10 years was most common. Data retention up to 10 years was most common. Additional findings from project responses are reported in Figure 1a-d. Face-to-face interviews have been conducted for 27 staff to date with emerging themes reported including:
- clarifying RDM institutional expectations and training opportunities, future policy/systems integration with existing structures, archiving guidance and data collaboration tools.

RDM policy framework development: The policy analysis resulted in two outputs: 11 general high level principles constituting the Research Data Management Policy and supporting the Code of Good Practice. This brought together all relevant institutional guidance into one accessible document.

RDM systems development: High quality project, person and publication data existed in the University’s current systems, however none dealt specifically with research data. In response, a data catalogue system was developed; this used existing collected data to provide rich derived metadata (Figure 2), thereby minimising workload duplication and increasing discoverability and the likelihood of re-use. The metadata records can be exposed (internally or externally), as a human readable website and/or in a research business system machine readable format.

Discussion
- Moderate storage quota increases would allow most projects data needs to be satisfied.
- Data retention up to 10 years was most common and is in line with funder requirements.
- Institutional storage locations (with defined operational service standards) were used, in part, by most (i.e. two-thirds) of projects.
- The Research Data Catalogue provides a metadata index (CERIF compliant) of institutional research data, in lieu of an institutional repository.
- Outputs will be trialled with exemplar projects and evaluation undertaken.

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