Research Data Culture

Origins

- Monash University approached some of the other research intensive universities to test interest in holding a multi university roundtable on sustainable institutional approaches to research data.
- Led to a Birds of a Feather session at eResearch Australia 2018. The topics were:
    - Exponential growth in data creation and use => institutions will coordinate large-scale infrastructure to support research data
    - Efficiency depends on the availability of appropriate information and metadata, and on the recognition of constraints, that depend on the involvement of researchers.
- And a meeting at Sydney University. The topics were extended to include:
    - The role of the five ‘pillars’ == library, records, archive, IT and eResearch activities
    - The fact that efficiencies, scalability, quality improvements and opportunities for strategic prioritisation require a coordination of a researcher's data through these pillars

Key Elements from Sydney Meeting

- **Research Data Management**
  - Involves a Focus on the P-Dimension (People - Policy - Process)
  - More than, the T-Dimension (Technology - Tools - Transformation)
- **Research Data Management is a Multi-Group, Multi-Responsibility, Multi-Support, activity within the university**
  - Needs collaboration from the 5 pillars - eResearch, IT, Library, Records and Archives
    - In the research data intensive institutions
    - Possibly at a national level (eg CAUL, CAUDIT, RIMPA, ASA, UA)
  - There is a need for engaging the leadership within these areas and enabling the leverage between these areas (this is a horizontal activity).

Key Elements from Sydney Meeting

- **Governance first / best practice:**
  - The purpose is to accelerate the adoption of the change in practice across many universities, using drivers found in those facing the problem at scale.
- **Researcher centric**
  - Has low barriers / no barriers to entry, low / no effort to perform.
  - Understands how to motivate the researcher to contribute, on research terms.
  - Is tolerant and enabling to those researchers doing something new / not well understood.
  - Is efficient and enabling to the majority of researchers.
  - Is actively and iteratively tested on real research use cases.
  - Provides potential for individual curation.

Key Elements from Sydney Meeting

- **Automate where possible**
  - Focus the use of human capital (to support research / research domain specific needs).
  - Reduce / limit the amount of orthogonal human capital to achieve the goals.
- **Plan for capture and access**
  - Improve the capacity to capture from the outset (instead of at end of life).
  - Request retention parameters from the outset and automate the capture of metadata.
  - Provision access to the right people for the right amount of time.

Concept Arising

We need a reference model for implementing Research Data Management Plans 2.0 (RDMP-2.0).

The goal is that RDMP-2.0 should be automatically generated as a result of research being done in the university environment and should help automate subsequent decision making.

Proposed actions:

- Prepare ‘Showbag’ outlining existing workable solutions for researcher’s problems.
- Form a sub-committee containing members from all representing institutions (requires name).
- Meet delegates to consolidate views and develop a position to put forward at a proposed ARDC sponsored Data Summit in July 2019.
- Prepare reference document outlining architecture (outlining research data culture mission).
- Find a suitable patron organisation to support the initiative.
- Visit heads of all community organisations to enable engagement

ARDC Interest

Discussion with the ARDC has produced an interest in exploring some additional important questions

- What is the role of institutions in a national research data commons.

More specifically

- How should national and institutional data collections infrastructure integrate into a more coherent national data system?
- How to best articulate the role of an institution in a national data commons and the institutions' expectations vis a vis the ARDC?
- What ongoing arrangements/structures might be required for that purpose?
- How investments in FAIR (and other national priorities) be aligned with investment in institutional priorities such as GDPR, sensitive data, disposal and capacities at large.

Context - ARDC Interest

The ARDC recognises that institutions and research organisations play a vital role in any national data framework.

An alignment of institutional and national infrastructure arrangements can drive better data management practices and outcomes.

The ARDC is seeking input from the sector on the alignment that would best assist progress with:

- Distributed national collections
- A national FAIR data safety net

ARDC Interest

National FAIR Safety Net - a data governance capability

As part of an open co-design process ARDC seeks input on

- The role of institutions in establishing research discipline led national commons
- The role ARDC can play in promoting collaboration, coordination, and coherence

Distributed National Collections

As part of an open co-design process ARDC seeks input on

- How this idea can align with institutional ambitions and how it can be best aligned

Further Events

- Two meetings are being considered
  - The first in the second week of December 2019
  - The second in late February 2020
- Discussions on the agenda for those meetings include
  - Roundtable input from the participants from the RDC Sydney meeting
  - Zoom briefings in the week of 14 October
  - An introductory discussion at the Data Summit on Monday 21st October in Brisbane
  - A BoF at the eResearch conference on Wednesday 23rd October