

Early Adopter Use case

Early Adopter's Name: Australian Research Data Commons

Organisation type:

The Australian Research Data Commons (ARDC) is a national and strategic infrastructure investment capability funded under a wider Australian Government initiative, the National Collaborative Research Infrastructure Strategy (NCRIS). The ARDC is a transformational initiative that enables Australian research and industry access to nationally significant, leading edge data intensive e-infrastructures, platforms, skills and collections of high-quality data.

Organisation size:

The ARDC is a medium-sized organisation according to Eurostat, but we operate a highly distributed employee model across Australia and a growing membership of 18 current institutional subscribers.

Organisation Research Field(s):

The ARDC has national scope and our members are Australian Universities, independent research institutions and other NCRIS capabilities. Collectively our members represent nearly all fields of research.

Organisation Profile:

The Australian Research Data Commons (ARDC) was formed on 1 July 2018. Presently, the ARDC is a company limited by guarantee and a registered charity with Australian Charities and Not-for-profit Commission (ACNC). We engage with publicly funded research agencies, universities and eResearch capabilities to become members of the ARDC and contribute to strategic direction and priority definition. Our current members can be viewed [here](#).

As a new initiative, our history is short. However, we bring to the eResearch sector over 10 years of experience on research data infrastructure and services. This is achieved by building on legacy initiatives of the Australian National Data Service (ANDS), the National eResearch Collaboration Tools and Resources (Nectar) and the Research Data Services (RDS).

Our organisation is defined by the following six principles

- Transforming research through better tools, by providing better software, platforms and data across the research lifecycle.
- Focusing on national scale opportunities to help develop a nationally coherent eResearch infrastructure environment in a global context.
- People are essential and we will continue to raise awareness of this and support communities in order to build skills and culture in the sector.
- Building strong partnerships and collaborations are at the heart of everything that we do. As one part of a national and international system, we work with others to inform, magnify and sustain common work objectives.
- Be a catalyst for and complement the sector by accelerating innovation through projects, infrastructure, services, consultancy and outreach.
- Commitment to sustainable expertise and services, digital infrastructure for data and tools.

Organisation website URL:

<https://ardc.edu.au/>

Suggested Use case title: Defining National Scale Data Archive Services**Problem definition:**

The ARDC considers the needs of the entire Australian research sector and aims to build a robust data commons via strategic investment, coordination and partnership. A recent national consultation exercise highlighted the absence of a national scale, discipline agnostic data archive facility. Such a data archive was recognised as a distinct service concept to the widespread instances of data repositories, large-scale storage facilities and distributed cloud service architectures. Given the widespread need for a national scale data archive capability the ARDC is investigating possible service delivery models and architectures that could fulfil this service gap. We are particularly interested in how such services are designed and implemented and made available via a federated and physically distributed community like our stakeholders and indeed the European research communities.

Our stakeholders vary in size, scope and data scale. We anticipate a data archive to scale to 20-100 petabytes over a 5-10 year period and provide services to 30-50 distinct organisational users. Any solution will need to operate a realistic business model with mature SOPs and SLAs. While presentation is a secondary concern to data preservation, a coherent and consistent view would expedite integration with other national and international infrastructures like [Research Data Australia](#), [DataCite](#), [ORCID](#), community data repositories, other international data commons movements and collaborative environments.

Is this use case new for your organisation?

Small scale data archives exist in several disciplines and with varying degree of maturity. The involvement of the ARDC in this project seeks to further understand the possible models that can fulfil national scale capabilities that can evolve into a comprehensive solution or solutions. We anticipate a period of service definition and design 2020-2023 that would specify a distributed model that mixes commercial and localised provision.

Data and metadata Characteristics:

The ARDC is building a minimal metadata requirement that specifies existing international schemas, e.g. DataCite MDS, but recognise a degree of extensibility is required for specific for defined communities.

Suggested length: 200 words

Cost requirements:

Currently costs are recognised as real and necessary but are not yet defined. Our only requirement is that they are realistic and reflect predictable and efficient investment for benefit.

Benefits and expected impact:

Benefits are recognised in different contexts.

Research Data Management. The benefits of RDM are internationally recognised and the ARDC accepts the current dogma that research data are first class research objects and are valuable evidence that underpins the scholarly record. We recognise the significant reuse benefits of cost efficiency, reputation and the value of a complete scholarly record.

Infrastructure provision: We recognise the consequences of accelerating data generation on capacity cost that to date has been shielded from efficient infrastructure management due to falling hardware costs. The approaching issue of increasing relative storage costs is significant and capacity cannot keep up with data generation. A key driver will be determining what data to keep and move to lower cost capacity and what data to delete is now as necessary for the IT Infrastructure Manager as equally as the Research Manager.

Strategic goals. In line with the foundational NCRIS principles, a national and collaborative approach to realising the above benefits is in itself a benefit to inclusiveness, efficiency with a limited but critical market appetite and delivering national solutions that supports a entire sector.

Suggested length: 600 words

Contact person & details:

WILKINSON, J Max

Linkedin profile: <https://www.linkedin.com/in/max-wilkinson-4486392/>

Twitter-n/a

Disclaimer: this form aims to gather further information on your use case for promotional purposes. It does not guarantee the deployment of your use case on the solutions resulting from the ARCHIVER project. As a reminder, the Buyers Group will approve the Early Adopters use cases that will be deployed to test the pilot solutions.