Australian Research Data Commons

ARDC Discovery Activities: Storage and Compute Infrastructure

Application information

Contents

Key dates

- **10 April 2019:** Applications open
- **18 April 2019:** FAQ/Q&A Webinar
- **10 May 2019 5pm AEST:** Applications close
- **17 May 2019**: Applicants notified of the outcome of their submission
- **13 September**: Projects provide reports back
- **16 and 17 September:** Presentation of findings at ARDC Infrastructure for Research Summit

Open call information

Information

All applicants are encouraged to read through all the documentation related to the Storage and Compute Infrastructure portfolio of the ARDC Discovery Activities. This includes this open call information document as well as the application forms.

Key application points:

- Applicants should use the supplied template
- Applicants can apply for one or more activities, with a separate proposal for each activity.
- Applications should address the selection criteria (both content and organisation criteria)
- Applications should be no longer than 2 A4 pages in length (excluding overview information page)

If you have any questions please contact submissions@ardc.edu.au

Eligibility

Applicants can be individuals, groups, or organisations but must have an Australian Business Number (ABN) and be willing to enter into the standard Contractor Agreement.

Submission

Applicants are asked to submit their fully completed documentation via the online submission page by 10 May 2019 5pm AEST. Applications that are received after this date will not be considered for funding.

Outputs

Projects will be expected to deliver a written report to the ARDC on the 13th September and present at the ARDC Infrastructure Summit on the 16 and 17 September.

Selection

An internal selection panel will assess and select projects against the selection criteria. Applicants will be notified by email on the 17 May 2019.

Questions

The ARDC will set up a FAQ register to answer any questions that we receive about the open call. A Q&A webinar will be run by the ARDC on 18 April 2019. FAQs and other relevant information will be shared at this event.
If you have any questions please contact submissions@ardc.edu.au

Program information and application criteria

Background

The Australian Research Data Commons has identified "Storage and Compute: Providing foundation infrastructure" as one of its strategic themes[1]. The ARDC would like to fund a number of investigative activities to identify issues, needs, gaps and requirements within the Infrastructure and Services Portfolio. The outputs of these relatively short "discovery" activities will be used as the basis of national infrastructure summit to be held in Q3 of 2019 and to inform future ARDC long-term infrastructure and services initiatives.

The ARDC will fund projects of up to $50,000 with a completion date of 13 September 2019. Funding of $500k is allocated to this portfolio activity. In addition to a written report, it is expected that each activity will present its findings at the ARDC Infrastructure and Services summit to be held on 16 and 17 September.

This open call intends to elicit broad involvement from throughout the sector as part of building collective vision and engagement in the infrastructure and services that underpin an Australian Research Data Commons. The projects are not meant to be academic reports but rather "investigating, doing and then sharing", resulting in technical and feasibility analyses that can be scrutinised and validated as realistic solutions to real issues. The projects themselves should provide meaningful outcomes for service providers, research communities and organisations as well as broader learnings for ARDC and the sector.

[1] The five (draft) strategic themes are:
Theme 1 - Coordination and Coherence: Facilitating an Australian research data commons
Theme 2 - People and Policy: Transforming culture and community
Theme 3 - Data and Services: Maximising the value of Australia's data assets
Theme 4 - Software and Platforms: Enabling research insights
Theme 5 - Storage and Compute: Providing foundation infrastructure

Aim of the funding program

ARDC has a history of innovating and supporting the Australian research sector's capabilities in cloud computing and collection development/management. ARDC is now looking at options for future investment, given a radically changed infrastructure world. To guide the ARDC investment program to build on, innovate with, and complement the strategies of key infrastructure and research partners, the ARDC has developed a set of initial questions which it invites partners to address.

Discovery activities

ARDC invites project proposals to address any or all of these questions. Each activity requires an individual application.

DAP/T5/P1: How have changes in the research infrastructure ecosystem (including commercial options) influenced the way the sector needs, uses, procures research IT resources?

Questions to be answered include:

- Snapshot: What are the needs (for compute and storage) for ARDC, NCRIS and the sector now?
- Future: What will be the needs for ARDC, NCRIS and the sector over the next few years?
- What are the pros/cons of:
    - ARDC building new infrastructure?
    - ARDC re-building existing infrastructure?
    - ARDC buying or supporting others purchasing infrastructure resources?
- Are there still market failures? i.e. Are there opportunities for innovation or transformation which will not be met without us?

DAP/T5/P2: What is the current usage (and interest/potential demand) of commercial cloud in the research sector and NCRIS capabilities?

Questions to be answered include:

- How could ARDC support this demand? What would be the model for funding/subsidising/co-investing in commercial cloud?
- How are allocations and payments handled by research organisations or programs now?
- How do they decide between commercial and on-premise resources for particular research requirements, and how should we?

DAP/T5/P3: Multi-Clouds: State of play and impact on *collaboration*

Questions to be answered include:

- What types of applications or use cases in the research sector could work well on multicloud?
- What are existing research applications and use cases where multicloud is being used?
- How could we best use multicloud and cloudbursting to support these applications? How is this being done by other organisations?
- What is the impact of multi- and hybrid-cloud on collaboration?
    - Collaboration of infrastructure providers
    - Collaboration of researchers
    - Collaboration of institutions?
- What is the impact of hybrid- and multi-cloud on interoperability of tools and data?
- How should we best support users to take advantage of hybrid/multi-cloud infrastructure?

DAP/T5/P4: What is the *total cost* of delivering a range of example workflows on local, shared, national, international and commercial platforms; an apples and apples comparison.

The intention of this program is to make an objective comparison of running workloads on a variety of underpinning infrastructure including, where relevant, whole of life costs. These costs are to be as comprehensive as possible and where infrastructure provided through mechanisms such as grants is used, the capital and operating cost of that infrastructure should be included also.

DAP/T5/P5: Exploring options for access to sensitive data sets; what authentication technologies (e.g. multi-factor authentication) are needed to access sensitive data and secure compute environments?

Questions to be answered include:

- How many services (existing and new) would find value in a higher level of assurance about the identity of a user through technologies like multi-factor authentication?
- How far does MFA support go to addressing requirements for access to sensitive data? What other components (standards, policy, process, training, technologies) are needed to satisfy requirements?
- How are other federations employing MFA within their networks? How widely is the service used?
- How might Australian infrastructure providers support MFA while still allowing identity providers to adopt any MFA technology that makes sense for them?

DAP/T5/P6: Repositories; ideas for how ARDC could support repositories in order to help drive standard approaches, interoperability, and alignment with international activity.

Questions to be answered include:

- What are the different types of repositories?
- Using examples, how are responsibilities for repositories and their data assigned and supported??
- What should be the ARDC's approach to supporting repositories of differing types?
    - Lead
    - collaborate or
    - fast follow?
    - And what approach would be appropriate for institutions/domain groups supporting repositories?
- What does a sustainable repository ecosystem look like?
    - What is an appropriate strategy for ARDC to support progression to this state?
- what's the relationship between repositories and informatics practices or platforms?
- To what extent will local activity be influenced by international domain groups?

DAP/T5/P7: How important is data locality to infrastructure strategy (should the "inertia" of data guide the deployment of infrastructure)?

Questions to be answered include:

- How important is data locality for various use cases?
- What are the drivers for locating data and compute close to or within institutions?
    - Institutional policy?
    - state/federal policy?
    - Efficiency?
    - Performance?
- What are institutions' views on whether data collections should be hosted locally or not, and what are their criteria?

DAP/T5/P8: What is the demand for machine learning tools and services and how could these be most effectively supported?

Questions to be answered include:

- What are the requirements of NCRIS capabilities and major research groups for these types of services?
- What exemplars are there in Australia and internationally of these kinds of services?
    - Commercial cloud machine learning services, hybrid cloud, specialised hardware?
- What is the best approach for supporting these services on different infrastructures (including the Nectar Research Cloud) and how can we encompass hybrid cloud including local infrastructure (clusters, local cloud, etc), and commercial cloud?
- How can we best take advantage of commercial cloud pricing options for these requirements?

DAP/T5/P9: What is the demand for high throughput computing tools and services and how could these be most effectively supported?

Questions to be answered include:

- What are the requirements of NCRIS capabilities and major research groups for these types of services?
- What exemplars are there in Australia and internationally of these kinds of services?
    - Commercial cloud services, hybrid cloud, specialised hardware?
- What is the best approach for supporting these services on different infrastructures (including the Nectar Research Cloud) and how can we encompass hybrid cloud including local infrastructure (clusters, local cloud, etc), and commercial cloud?
- How can we best take advantage of commercial cloud pricing options for these requirements?

DAP/T5/P10: What are the global experiences of infrastructure federations and what is the capacity of federation to continue to provide benefit?

Questions to be answered include:

- Costs/Benefits of federation
- International models for federation – of data and/or tools and/or infrastructure
- Allocation and priority schemes – compare and contrast locally and internationally
- Governance models – compare and contrast locally and internationally
- How best to trade off between federation-wide standards for interoperability, and local requirements for customised services and policies?

DAP/T5/P11: How can we best support analytics platforms such as Jupyter, R Studio?

Questions to be answered include:

- How is this being done already through Virtual Laboratories, infrastructure providers, universities, international activities?
- How best to support domain-specific requirements and tools as well as more generic requirements and tools?
- Should we/could we support JupyterHub and/or BinderHub or similar services for research communities? Is there demand for this? If so, how best could it be supported nationally?

Criteria for assessment of proposals

Content related criteria

1. How effectively does the proposal address the specific questions listed above?
2. How broad is participation/engagement in the proposal? (for example, are several organisations supporting the study, expressed through letters of support (expressions of support do not count towards the application page limit))
3. Is there a clear description of who will be consulted in the investigation?
4. Is the proposal competitively priced?

Organisation related criteria

1. Is there a clear description of who will deliver the project?
2. Do they have a demonstrated track record of delivering high quality project outcomes on time?
3. Does the organisation (or person) currently play a significant role in the relevant community?