

9 April 2019



Australian Research Data Commons

ARDC Strategic Plan 2019 - 2023



1. The Australian Research Data Commons.....	6
1.1. Definition	6
1.2. Principles	6
1.3. Values	7
2. Landscape Analysis.....	8
3. ARDC’s Unique Role	11
4. Strategic Impact	12
5. Organisational Structure	13
6. Engagement and Communications	14
7. Implementation Plan.....	15
7.1. Theme 1 - Coordination and Coherence: Facilitating an Australian research data commons.....	16
7.2. Theme 2 - People and Policy: Transforming culture and community	16
7.3. Theme 3 - Data and Services: Maximising the value of Australia’s data assets	17
7.4. Theme 4 - Software and Platforms: Enabling research insights	18
7.5. Theme 5 - Storage and Compute: Providing foundation infrastructure	19
8. Investment Profile.....	21
9. Appendix 1 - ARDC Impact Pathway	22

LIST OF FIGURES

Figure 1: Contributions to realising a commons (based on an original idea from osf.io).....	5
Figure 2: CSIRO Impact Framework	12
Figure 3: Contributions to realising a commons (based on an original idea from osf.io).....	15

Vision

Transforming digital infrastructure to support leading edge research and innovation

Mission

ARDC is a transformational, sector-wide initiative, working with sector, government, and industry partners to build a coherent national and collaborative research data commons. This will deliver a world-leading data advantage, facilitate innovation, foster collaboration and enhance research translation.

Executive Summary

The Australian Research Data Commons (ARDC) was formed by the Federal Government in July 2018 through the merger of three existing National Collaborative Research Infrastructure Strategy (NCRIS) eInfrastructure capabilities (the Australian National Data Service (ANDS), National eResearch Collaboration Tools and Resources (Nectar) and Research Data Services (RDS)) with the vision of “transforming digital infrastructure to support leading edge research and innovation”.

The ARDC has an operating budget for the period of this strategic plan of \$110m and a capital investment budget for the same period of \$72m. The ARDC will strategically invest these funds to deliver a step change in sector capabilities to complement and build on existing infrastructure. This will ensure resources, skills, and workflows are discoverable, interoperable, accessible to all researchers, sustainable and robust, and provide access to the necessary skills and workforce development programs to enable competitive and high impact Australian research.

ARDC will play a key role in transforming Australia’s research culture through its strategic vision, leadership, programs and resource provision, and through its mandate to provide national coherence via the Digital Data and eResearch Platform capability (identified in the 2016 National Research Infrastructure Roadmap, and including the National Computational Infrastructure (NCI), Pawsey Supercomputing Centre, Australia’s Academic and Research Network (AARNet) and Australian Access Federation (AAF). Building on the achievements of its three predecessor capabilities, ARDC is uniquely positioned to drive collaboration with partners sector-wide, nationally, and internationally to develop a coherent and collaborative national research data commons for the benefit of all Australian researchers.

The ARDC organisation will, working in concert with others, deliver the Australian Research Data Commons (the commons), which brings together people, data, skills, and resources to enable researchers to conduct world class data-intensive research. The commons, potentially made up of component commons revolving around domains, data types, or other priorities, will be a transformational change in the research data ecosystem, increasing the coherence and interoperability of existing investments and thereby increasing the effectiveness and efficiency of the system for researchers and producing more impactful outcomes for the nation. By enabling increasing numbers of researchers, communities, and institutions to engage with the resources and capabilities a commons delivers, a transformational change in research capability and impact can be achieved.

ARDC recognises the critical positions that partnerships occupy in the delivery of a commons. It is crucial that the commons is developed collaboratively by the sector, for the sector, and taking into account the sector’s partners - government, industry, society, both nationally and internationally. As a driver and enabler within the Australian eResearch ecosystem, ARDC will leverage its expertise and partnership networks to enable collaborations across the sector that support the development of a sectoral vision for a nationally coordinated commons.

The ARDC priorities in the development of this commons with partners focus on delivering four transformational outcomes for the research sector:

1. enabling a world-leading data advantage;
2. facilitating accelerated innovation;

3. fostering collaboration for borderless research; and
4. enhancing researchers' ability to translate their research into benefits for society.

In pursuit of the four organisational priorities, the Strategic Plan describes the five strategic themes that will frame the implementation of ARDC's vision:

- 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 & supporting collaboration
- Theme 5 - Storage and Compute: Providing foundation infrastructure

These complementary and interdependent themes guide ARDC's activities and outcomes, and act as an organising structure for activities. Illustrated diagrammatically, the themes will contribute to realising the commons as shown in Figure 1:



Figure 1: Contributions to realising a commons (based on an original idea from osf.io)

Theme 1 will deliver Coordination and facilitation between all themes and across the sector, industry, and government in pursuit of the commons. Themes 2 & 3 will concentrate on Communities and Culture, Policy and Incentives, Skills and Training, Frameworks and Standards - these themes have overlaps but different emphases. Theme 4 is directly focussed on Platforms and Tools. Theme 5 will provide the enabling storage and compute infrastructure.

1. The Australian Research Data Commons

1.1. Definition

The ARDC was formed on 1 July 2018 as a result of the 2016 NCRIS Roadmap activity with the aim of working with sector-wide partners to build a coherent national and collaborative research data commons.

There are several definitions of what constitutes a commons. For the purposes of ARDC's activities it is:

Information and knowledge resources that are collectively created and owned or shared between or among a community. Such a commons includes the community itself together with the skills of the individuals building the resources.

The overarching goal of the commons is to accelerate Australian research by developing, testing, and supporting platforms where investigators can store, discover, share, access, and interact with digital objects (data, software, etc.). To support this goal the commons will enable workforce capability and policy development across stakeholders. By connecting digital objects and making them accessible and reusable, the commons is intended to allow collaborative and novel scientific research that was not possible before, including hypothesis generation, discovery, and validation.

A research data commons therefore brings together people, skills, data, and related resources (storage, compute, software, models) and partnerships to enable researchers to conduct world class data-intensive research. This concept arises from the increasing interest in the last decade in the economics of commons in general, and in the ability of digital goods and services to be part of a digital commons. In addition to enabling improved research impact, a research data commons can be transformational in facilitating change in the way research is conducted, in terms of both research culture and the availability of research data and analytical tools.

The 2016 National Research Infrastructure Roadmap provides ARDC with a mandate to provide national coherency, catalyse an improved eResearch ecosystem in the pursuit of the commons, and support NCRIS capabilities. While supporting the needs of NCRIS, the ARDC strategy aims to engage and align with the strategic aims of the broader research community, both nationally and internationally, as well as industry and government priorities. ARDC will advocate for policies, strategies and resources that help to create an enabling framework and build capability that drives high-quality, accessible, and competitive research outcomes.

1.2. Principles

ARDC will be guided by the following principles in delivering on this mission.

Transforming research through better data and tools

Digital data, tools and methods are changing everything, including the way research occurs and address the grand

challenges of society. ARDC works to transform research and its impact through the innovative application of FAIR data, software and platforms across the research lifecycle

Focus on national scale opportunities

ARDC's goal is to work with others to realise a nationally coherent eResearch infrastructure environment in a global context. ARDC facilitates coordination across the system as a whole by connecting resources, cross-pollinating, driving interoperation, joining forces and sharing best practice across multiple stakeholders.

People are essential

People must always be part of the solution. ARDC will continue to raise awareness of this and support communities in order to build skills and culture in the sector.

Partnerships and collaboration are central

Partnerships and collaboration are at the heart of everything that ARDC does. As one part of a national and international system, ARDC strives to connect up with, empower and learn from others in the sector. ARDC works with others to inform, magnify and sustain shared work.

ARDC catalyses and complements

ARDC accelerates innovation through projects, infrastructure, services, consultancy and outreach to help the sector do what cannot be done otherwise. ARDC always looks towards the next big challenge.

Committed to sustainable outcomes

National and international collaborative research needs dependable digital infrastructure for data and tools with sustainable long term underpinnings. ARDC work with and enable partners to develop sustainable expertise and services.

1.3. Values

ARDC is developing a shared set of values as part of a program of work on cultural change that is already underway.

2. Landscape Analysis

ARDC is located within both national and international eResearch ecosystems, and its unique positioning must be analysed, alongside the achievements of its three predecessor organisations, to identify the niche from which ARDC can provide maximum value and impact.

The value of national infrastructure in digital data and eResearch platforms is well understood in Australia, as evidenced in the work of a number of recent reviews:

- Facilities for the Future: Underpinning Australia’s Research Innovation (Government Response to the 2016 National Research Infrastructure Roadmap Research Infrastructure Investment Plan) (Australian Government, 2016)¹
- 2016 National Research Infrastructure Roadmap²
- eResearch Framework³
- Status Report on the NCRIS eResearch Capability⁴

These reviews each conclude that the ability to perform complex computations rapidly, coupled with data management and storage, and data mobility, is essential if Australia is to effectively provide and efficiently take advantage of an evolving data-intensive research environment. The collaborative emphasis of NCRIS capabilities is also seen as crucial. For example, the Status Report on the NCRIS eResearch Capability states: “The national eResearch investments have been timely and critical to the advance and competitiveness of Australian research. ... The deliberate strategy of fostering and implementing collaborative approaches has been a hallmark of the development as acknowledged across a range of community responses.”⁵

The three predecessor organisations to ARDC have contributed to significant transformations in the research sector for more than ten years. ARDC is now positioned to lead the second stage of this journey. ANDS championed the enduring value of data and data management and changed how institutions see their fundamental responsibilities for the data their researchers generate; and spearheaded highly successful data initiatives that ensure data are

¹ Australian Government, 2016, *Facilities for the Future: Underpinning Australia’s Research Innovation*. Viewed 12 February 2019, https://docs.education.gov.au/system/files/doc/other/ed18-0069_-_he_-_research_infrastructure_investment_plan_-_public_version.pdf

² Australian Government, 2016, *2016 National Research Infrastructure Roadmap*. Viewed 12 February 2019 https://docs.education.gov.au/system/files/doc/other/ed16-0269_national_research_infrastructure_roadmap_report_internals_acc.pdf

³ Francis R, 2016, *Towards an eResearch Framework* (powerpoint presentation). Viewed 12 February 2019, <http://aero.edu.au/wp-content/uploads/2015/03/Rhys-eRFramework-AeROForum2016.pdf>

⁴ Cochrane, T, 2014, *Status Report on the NCRIS eResearch Capability*. Viewed 12 February 2019 https://docs.education.gov.au/system/files/doc/other/abridged_eresearch_status_report_-_web.pdf

⁵ Cochrane, T, 2014, *Status Report on the NCRIS eResearch Capability*. Viewed 12 February 2019 https://docs.education.gov.au/system/files/doc/other/abridged_eresearch_status_report_-_web.pdf

findable, accessible, interoperable and reusable (FAIR). One of the major outcomes of the ANDS activity was that research institutions now see data as assets, not just a responsibility - principally in order to build research strength and prosecute research partnerships. Nectar enabled researchers to work together in new ways, able to collaborate over data, tools and computation across the country and internationally, providing an international competitive advantage for Australian researchers who were therefore more attractive as international research partners. RDS partnered with other eInfrastructures to enable a suite of reliable, high quality data services, including high capacity data storage and identity services with the result that data services became both more reliable and less the responsibility of individual research groups, which substantially lifted the professional partnership between research and eResearch.

The three predecessor organisations played significant leadership roles nationally through various initiatives, and internationally through roles in significant global initiatives such as the Research Data Alliance. With the combined background, ARDC is well positioned to both continue and expand upon this leadership.

ARDC functions within a broader eResearch ecosystem in Australia that includes significant national investments by NCRIS capabilities, national service providers including Australia's Academic and Research Network (AARNet) and Australian Access Federation (AAF), Publicly Funded Research Organisations (PFROs), research institutions and others. These organisations recognise the value of digital data and eResearch platforms and continue to invest heavily in eInfrastructure to support the achievement of their strategic research aims. Australia's state-based eResearch service providers also have significant investment in infrastructure and services.

The diversity and number of organisations involved in the digital data and eResearch platforms in Australia provide both richness and complexity to the landscape. Figures from the Council of Australasian University Directors of Information Technology (CAUDIT) reveal that Australian universities and CSIRO spend \$2,313.4 billion annually on IT in higher education and research (this figure includes \$980.8m of staff investment and \$1,332.6m of non-staff investment)⁶, highlighting the need for collaboration across the sector to maximise the benefits of these investments. While some of these organisations focus on institutional, regional or research community interests, many have national interests, with significant prior collaborations with ANDS, Nectar and RDS. These include NCRIS capabilities that have invested heavily in data-focussed infrastructure.

There are a range of other research sector organisations that are important in this landscape, including: federal and state governments; research communities; funders such as the Australian Research Council (ARC) and National Health and Medical Research Council (NHMRC); peak bodies including the Council of Australian University Librarians (CAUL), Council of Australian University Directors of Information Technology (CAUDIT), and the Australian research academies; and a range of other organisations that affect policy (including publishers) and workforce development.

The impact of commercial services entering this sector, where once they were absent, is also relevant to this landscape analysis. For example, anecdotal evidence suggests an increased utilisation of commercial cloud, storage and analysis tools by research stakeholders; it would seem prudent to engage with such an emerging market.

⁶ Figures are available in CAUDIT, *2018 Benchmarking Report*. Viewed 12 February 2019. The figures cited here can be accessed using AAF or Tuakiri login credentials at <https://www.caudit.edu.au/benchmarking-reports>

Australia's digital data and eResearch platforms initiatives can also benefit from increased alignment with a range of leading international initiatives, such as the Research Data Alliance. Whilst different parts of the Australian research sector engage with many international initiatives, alignment with national interests and coordination across national initiatives could benefit from more coherent facilitation.

An increasing number of international programs are also forming to coordinate specific geographic or research discipline communities. Significant examples include the European Open Science Cloud (EOSC) and USA's National Institutes of Health (NIH) Data Commons.

EOSC is envisioned by the European Commission as:

*a supporting landscape to foster open science and open innovation: a network of organisations and infrastructures from various countries and communities that supports the open creation and dissemination of knowledge and scientific data.*⁷

The USA's NIH Data Commons is being created to:

accelerate new biomedical discoveries by developing and testing a cloud-based platform where investigators can store, share, access, and interact with digital objects (data, software, etc.) generated from biomedical and behavioral research. By connecting the digital objects and making them accessible, the Data Commons is intended to allow novel scientific research that was not possible before, including hypothesis generation, discovery, and validation.⁸

In conclusion, the landscape analysis indicates that national coordination of eInfrastructure is the most effective mechanism to maximise the investment impact, particularly given the diverse stakeholder community. Any such national coordination must also build strong alignment with emerging commercial interests and international eResearch initiatives.

⁷ Andreozzi, S, et. al, 2017. *Report on the governance and financial schemes for the European Open Science Cloud*. Viewed 12 February 2019 https://ec.europa.eu/research/openscience/pdf/ospp_euro_open_science_cloud_report-pdf

⁸ National Institute of Health. *Data Commons Pilot*. Viewed 12 February 2019 <https://commonfund.nih.gov/commons>

3. ARDC's Unique Role

ARDC has a national coordination role to play in this environment, able to facilitate the involvement of the sector as a whole in the development and implementation of a national research data commons.

One of the long-term considerations for ARDC will be to what degree ARDC investments support the priorities of the commons as a whole, or are more focused on enabling the ARDC to address aspects of the commons where ARDC has particular expertise.

ARDC is recognised by the Government as a cornerstone of the research ecosystem, as one of the four identified priority areas under Digital Data and eResearch Platforms in the 2016 National Research Infrastructure Roadmap:

- tier 1 high performance computing - National Computational Infrastructure (NCI) and Pawsey Supercomputing Centre;
- research networks - Australian Academic Research Network (AARNet);
- access and authentication - Australian Access Federation (AAF); and
- an Australian research data cloud [sic] (ARDC)

ARDC can also play a role in facilitating alignment with other commons initiatives internationally, as part of the international emphasis inherent in all of the ARDC's themes. The two Research Data Alliance (RDA) plenary meetings held in 2018 included sessions to advance collaboration between different international initiatives⁹, including the African Open Science Platform, ARDC, USA's NIH Data Commons and the Canadian approach. There is interest in convergence on a framework that will facilitate alignment, efficiency and interoperability where possible and desirable. In parallel to these discussions, the European Commission has identified as one of the five priority deliverables for the EOSC to ensure EOSC activities align and coordinate with others globally. Greater coherence between eResearch environments in Australia and overseas will strengthen the ability of researchers to contribute to international research collaborations and drive success in the four transformations.

⁹ Towards a Global Open Science Commons - RDA 11th Plenary BoF meeting. <https://rd-alliance.org/towards-global-open-science-commons-rda-11th-plenary-bof-meeting>. Viewed 1 March 2019; and Delivering a Global Open Science Commons. Viewed 1 March 2019 <https://www.scidatacon.org/IDW2018/sessions/252/>

4. Strategic Impact

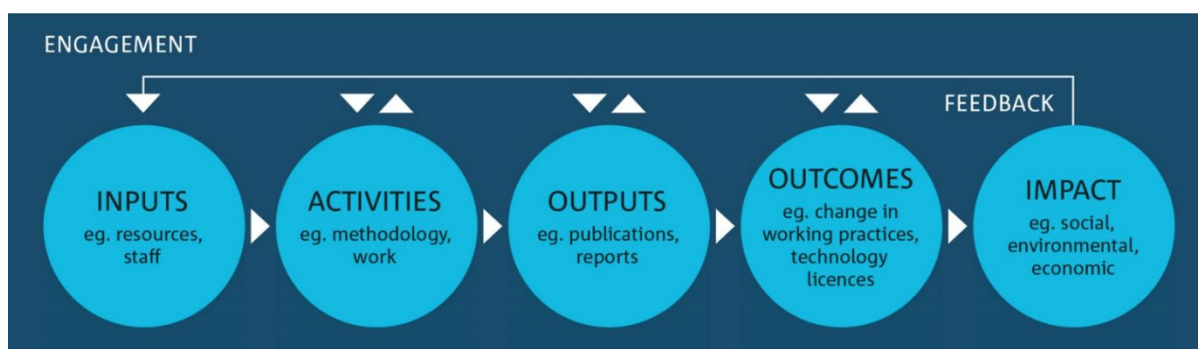


Figure 2: CSIRO Impact Framework

ARDC utilises the impact framework methodology developed by CSIRO to tie activities to achievement of societal impact, as represented in Figure 2¹⁰.

This includes articulation of the impact pathways that enables the four transformational outcomes for the research sector that ARDC prioritises in the development of this commons with partners:

- enabling a world-leading data advantage
- facilitating accelerated innovation
- fostering collaboration for borderless research
- enhancing researchers' ability to translate their research into benefits for society

Ongoing work within ARDC is utilising this impact framework to identify key performance indicators that will guide implementation planning. The current version of the ARDC impact pathway is included in Appendix 1.

¹⁰ CSIRO, 2018, *Impact methodology*. Viewed 12 February 2019 <https://www.csiro.au/en/About/Our-impact/Our-impact-model/Ensuring-we-deliver-impact>

5. Organisational Structure

ARDC has an organisational structure that meets its current workforce requirement, while also remaining flexible to support future growth and change. Retaining and securing talent within key fields of excellence are also critical, and ARDC aims to work with the broader community to ensure the best minds are contributing to ARDC's work.

6. Engagement and Communications

ARDC's engagement strategy aims to facilitate effective engagement at the appropriate level with strategic stakeholders of all types throughout the sector, to ensure the ARDC vision and programs are informed by, and responsive to, stakeholder interests.

Collaborative relationships will be key to delivering on ARDC's vision and in achieving the transformational goals shared by both ARDC and the larger community. ARDC's engagement strategy aims to facilitate effective engagement at the appropriate level with strategic stakeholders of all types throughout the sector, to ensure the ARDC vision and programs are informed by, and responsive to, stakeholder interests.

The ARDC engagement strategy includes a communications plan, which tailors messages to different stakeholders, ranging from consultative approaches for partners, to information provision to the broader community.

Key stakeholders include:

- Peak bodies, funders
- Research organisations
 - Universities
 - PFRA's
 - CRCs
 - ARC Centres of Excellence
 - MRIs
- Infrastructure providers
 - NCRIS
 - Other national and state infrastructure providers (including the current node community)
 - Digital Data and eResearch Platform members (AARNet, AAF, NCI, Pawsey)
 - Disciplinary archives
- Commercial infrastructure providers
- Government (state and commonwealth)
 - Science agencies
 - Data-rich departments
 - Policy agencies
 - Galleries, Libraries, Archives and Museums

7. Implementation Plan

ARDC has identified five strategic themes to frame the implementation of its vision:

- 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

These complementary and interdependent themes guide ARDC’s activities and outcomes, and act as an organising structure for activities. Illustrated diagrammatically, the themes will contribute to realising the commons as shown in Figure Figure 3:



Figure 3: Contributions to realising a commons (based on an original idea from osf.io)

Theme 1 will deliver Coordination and facilitation between all themes and across the sector, industry, and government in pursuit of “the commons”. Themes 2 & 3 will concentrate on Communities and Culture, Policy and Incentives, Skills and Training, Frameworks and Standards - these themes have complementary foci and emphases. Theme 4 is directly focussed on Platforms and Tools. Theme 5 will provide the enabling storage and compute infrastructure.

The ARDC Priority Criteria will be utilised to assist with more detailed planning across the five themes to guide investment decisions:

- **Transformation** - will the investment contribute to transformation of research practices or outcomes?
- **National Scope and/or Scale** - is the investment truly national, across multiple research disciplines?
- **Impact** - will the investment contribute to impact both directly by ARDC and also down-stream by the research sector? How will impact be measured?
- **Sustainability**, which in terms of measurable criteria, for the purposes of the ARDC is made up of:

- **Leverage** - how will the proposed investment amplify the impact of existing funding or activities?
- **Responsibility** - is there an entity who will take long term responsibility for sustainable continuation of the results of the investment?

7.1. Theme 1 - Coordination and Coherence: Facilitating an Australian research data commons

A fundamental role of ARDC is to facilitate sector-wide discussion regarding the vision, governance structure and implementation program that will create an Australian research data commons. This must occur in partnership with a wide range of stakeholders, to ensure coherence across organisational, institutional and national efforts to maximise the impact of the investments.

Implementation of ARDC's Strategic Plan will include two aspects: to determine what is necessary to create an Australian research data commons; and identification of the niche roles to which ARDC is uniquely positioned to contribute.

Areas of initial focus for this theme will include:

- identification of relevant international initiatives, emerging technology and future trends, as inputs for consideration with relation to a developing Australian research data commons
- gathering key stakeholders and thought leaders in discussions to identify the vision and mission of an Australian research data commons
- engagement with a wide range of stakeholders, to encourage communication and participation in the development of the Australian research data commons
- coordination in identifying and implementing governance and operational structure
- coordination and identification of key parts of the Australian research data commons, and discussions around who will take responsibility for enabling these

7.2. Theme 2 - People and Policy: Transforming culture and community

Meeting the challenges required for Australia to excel in the new digital economy necessitates a focus on the policy environment and social infrastructure needed to realise a vibrant data commons. In order to create effective and sustainable digital infrastructure, social, organisational and cultural issues must be addressed, in addition to technical solutions. Cultural change, workforce development programs and policy frameworks, supported by sustainable communities, are needed to create an environment that support and capitalise on the FAIR data collections, collaborative platforms and underlying infrastructure; collectively these are the key elements of an Australian research data commons.

ARDC will establish partnerships with key stakeholders to facilitate collaboration and coordination of this theme, with a particular focus on:

- advancement of cultural change through policy and funding frameworks
- skilled workforce planning for the sector
- development of key communities of practice, including connections with international communities and initiatives

Areas of initial focus are likely to include:

- policy initiatives (funder, publisher and institutional) that address the incentives for action
- awareness raising and research value propositions to support required cultural change
- analysis of supply and demand, to identify what is required to meet future need
- education and training for data and eResearch professionals
- alignment with sector-wide strategies for training researchers and the broader context of digital skills in Australia
- development of strong communities of practice
- credit and attribution models and career paths for eResearch and data professionals
- workforce diversity

A set of national summits will be convened with key stakeholders in mid 2019, to further focus aspects of this theme.

7.3. Theme 3 - Data and Services: Maximising the value of Australia's data assets

The 2016 National Research Infrastructure Roadmap noted that “All areas of research are increasingly dependent on data and eResearch infrastructure”¹¹ and that “a prominent trend is the rising dependence on digital data at new scales, and the complexity and diversity of data being generated. This unprecedented growth in data volume and complexity places increased demand on infrastructure and the skilled staff needed to support it.”¹²

The 2016 National Research Infrastructure Roadmap and the other reviews surveyed in the Landscape Analysis have emphasised the importance of data to the research enterprise. Recurring elements in these reviews are the need for

¹¹ Australian Government, 2016, *2016 National Research Infrastructure Roadmap* (page 5). Accessed 12 February 2019 , <https://docs.education.gov.au/system/files/doc/other/ed16-0269-national-research-infrastructure-roadmap-report-internals-acc.pdf>

¹² Australian Government, 2016, *2016 National Research Infrastructure Roadmap* (page 27). Accessed 12 February 2019 <https://docs.education.gov.au/system/files/doc/other/ed16-0269-national-research-infrastructure-roadmap-report-internals-acc.pdf>

an ongoing impact and responsibility framework for data, as well as the importance of putting in place national FAIR data assets and the infrastructure required to support these assets and services to provide sustained access to them.

Two strategic objectives will be pursued in the Data and Services theme to transform the research system as a whole:

1. Increase innovation in research by enabling new research from existing data (and even new kinds of research from improvements in the quality and quantity of data assets).
2. Increase the integrity and reproducibility (and reduce duplication) of research across the whole research system by increasing the FAIR-ness of the data arising from research.

The first objective will see ARDC partner with research communities, facilities, government agencies, and research organisations to target data collections with high re-use value, strong community ownership, and a national scope (ie, beyond the normal remit of single institutions). ARDC's role will be to provide seeding resources to develop these national and community collections and their long-term sustainability.

The second objective will see ARDC partner with the sector as a whole to promote institutional data curation and management capacity as well as greater coherence of data collections and services across different sectors of the national research data ecosystem. ARDC's role will be in supporting, facilitating, and communicating best data practice, interoperability, capacity, and capability across the sector. This will include a commitment to ensuring that Australia is aligned with international standards and initiatives and has more data that is FAIR.

The capacity of the sector to manage sensitive data emerged from sector consultations as a current priority issue.

Areas of initial focus for the sustainable data assets investment will include:

- transformational data collections
- sensitive data and approaches, platforms and services to manage, collaborate over and share these data
- institutional roles and approaches in the data commons

A national data summit will be convened with key stakeholders in mid-2019 to further focus this theme.

7.4. Theme 4 - Software and Platforms: Enabling research insights

The effective use of continuously growing and increasingly complex data assets brings an associated need for reliable and trustworthy software to deal with the volume and complexity. The volume dimension means that much data can only be effectively analysed by machine and the complexity dimension often means that the data needs to be transformed into secondary products in order to be understood by humans. These two requirements mean that ARDC will need to continue to support the delivery of world-leading informatics capabilities for Australian researchers and their collaborators. This increase in complexity carries considerable reproducibility risks. Data assets need robustness, reliability, and accessibility. Similarly, software and process assets need provenance, reliability,

support, accessibility and transparency. Finally, software needs to be recognised as a first class research output in the same way that data now is.

Priorities for the software and platforms program include (but are not limited to):

- Supporting communities of computational technique users to develop and grow
- Supporting training and skills development across all career stages and levels of expertise (in close coordination with the People and Policy theme)
- Maximising the impact of research software by promoting software sharing (along the same lines as data sharing)
- Supporting the development of Australian research software that is high quality, accessible and sustainable
- Supporting the use of transformational technologies such as software and platforms for big data analytics and machine learning
- Providing access to research platforms for a wider range of Australian researchers

Initial activities for the software and platforms program will include:

- Defining a comprehensive software-as-infrastructure strategy for the next five years
- Shaping a next generation platforms investment, including determining how best to draw on the results of existing national and international investments in platforms
- Investigating the creation of an Australian equivalent to the Software Sustainability Institute in the UK

7.5. Theme 5 - Storage and Compute: Providing foundation infrastructure

The 2016 National Research Infrastructure Roadmap notes that “In terms of established infrastructure, the demand for computation and connectivity will continue unabated” and the “unprecedented growth in data volume and complexity places increased demand on infrastructure and the skilled staff needed to support it.”¹³ The underlying physical infrastructure needed to support this growth, including data storage and compute, are key elements of

¹³ Australian Government, 2016, *2016 National Research Infrastructure Roadmap* (page 27), Viewed 12 February 2019 https://docs.education.gov.au/system/files/doc/other/ed16-0269_national_research_infrastructure_roadmap_report_internals_acc.pdf

ARDC strategy. This infrastructure provides the base on which collaborative platforms can be used to analyse data and to maximise the value of data assets.

ARDC aims to foster an environment that will support delivery of reliable, agile and sustainable underpinning infrastructure to Australian researchers. This infrastructure will align and integrate with international, national, regional, institutional, and domain-focused compute and storage infrastructures and services as part of the broader digital data and eResearch platform capabilities agenda. ARDC seeks to enable strategic infrastructure planning that provides key features identified through national consultations:

- Flexibility for storage and compute services
- A value model to maximise impact of these investments
- Robust governance
- Capacity and capability build on design and need
- Integration with other services
- Innovative solutions to market failures

This infrastructure will support the provision of FAIR data, tools and resources, while enabling infrastructure agility. Initial foci include:

- Managing potential challenges around co-investment and resource allocation in the face of new delivery models (such as consumption based infrastructure)
- Agreeing on short, medium, and long term priorities
- Identifying funding models for the future
- Addressing the urgent need for infrastructure refresh
- Approaches that avoid future capital cliffs for infrastructure refresh
- Infrastructure that supports ARDC activities

8. Investment Profile

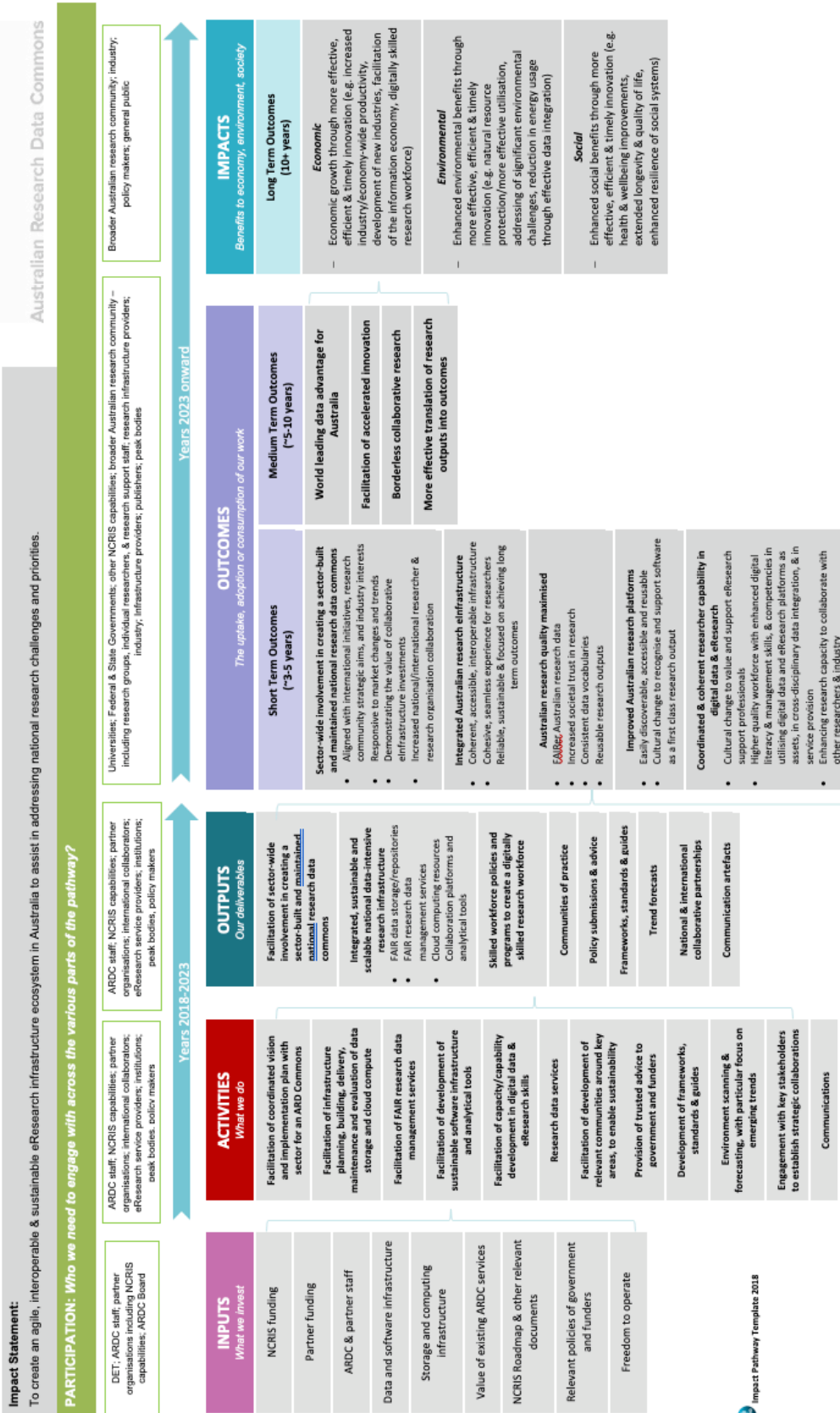
This ARDC Strategic Plan applies to a total investment of \$182 million, reflecting ARDC's operational allocation from NCRIS for 2018-2023 of \$110m, and the \$72m capital investment included in Facilities for the Future: Underpinning Australia's Research Innovation (Government Response to the 2016 National Research Infrastructure Roadmap Research Infrastructure Investment Plan). It should be noted that the allocations for 2018/19 have already been implemented.

The capital investment case, which was funded in the 2018 Budget, was based on the provision of reliable, agile and flexible underpinning resources including sustainable access to compute and storage. It is therefore anticipated that the capital investment will be primarily focused around theme 4 (Software and Platforms: Enabling research insights) and theme 5 (Storage and Compute: Providing foundation infrastructure). The timing and nature of this investment is detailed in the "ARDC Capital Plan 2019-2023".

9. Appendix 1 - ARDC Impact Pathway



Australian Research Data Commons (ARDC) – Impact Pathway V1.0





Australian Research Data Commons

Contact us:

+61 3 9902 0585

contact@ardc.edu.au

www.ardc.edu.au

The Australian Research Data Commons (ARDC) is a transformational initiative that enables Australian researchers and the research community access to nationally significant, leading edge data intensive eInfrastructure, platforms, skills and collections of high-quality data.

The ARDC is supported by the Australian Government through the National Collaborative Research Infrastructure Strategy.

