Australian Government
Department of Education

National Digital Research Infrastructure Strategy

Draft for Public Consultation

Contents

Introduction 4

Purpose 5

Implementation and Review 6

NCRIS Investment 6

Vision for future NDRI ecosystem 7

Outcomes 7

1. Underpinned by training frameworks for researchers and NRI workforce 8

2. Responsive to disruptive technological and societal shifts 9

3. Consistent in its standards for data collection, curation and access. 11

4. Integrated across levels of computing and data 14

5. Cybersecure, particularly for national-scale data and computing 16

6. Maximised by openly available research software tools 18

Introduction

In an age of increasing complexity, future technological progress will be largely driven by research that occurs across disciplinary boundaries, aiming to solve grand-scale problems.

This is not possible without the data management, computational, networking, visualisation and analytics tools that are essential for transdisciplinary collaboration at scale. These digital capabilities underpin modern research and national and international priorities, such as climate change.

Australia's system of national digital research infrastructure (NDRI) is fundamental to the research and innovation effort and provides critical support to government decision-making[1]. This is further emphasised by the need for Australian researchers to maintain and grow international research collaboration networks and to uphold the nation's reputable global standing.

> National Digital Research Infrastructure (NDRI) encompasses those digital research infrastructure components that are so nationally significant or large in scale, complexity or cost that they cannot be offered by a single institution or facility. - 2021 National Research Infrastructure Roadmap

Examples of digital research infrastructure include:

- Suitably skilled, expert NDRI personnel
- Data generated both for and through research
- Computing, including national (Tier 1) and institutional (Tier 2) high performance computing and cloud computing
- Data management platforms, including mechanisms for sharing data, storage and long-term management
- Analysis and visualisation software tools, in particular integrated into virtual research environments
- Cybersecurity research infrastructure
- Trust and identity services, frameworks and persistent identifiers.

The growth in user demands for computational and data tools and services[2] is fuelling a rapid expansion of an increasingly complex NDRI landscape.

All research communities employ digital tools of some form to generate, share and analyse data. The nature of this data can vary widely, depending on the application and the needs of their user communities. Technology is rapidly evolving to cater for these demands, creating novel opportunities – but also posing challenges in maintaining cohesion and shared strategic intent across the Australian research and innovation system.

To manage these challenges, Australia needs a coordinated strategy for planning and expanding Australia's digital research infrastructure system. This NDRI Strategy positions Australia to identify and prioritise NDRI needed for researchers, innovators and policymakers to conduct high quality research, and translate it to improve the lives of all Australians.

[1] *2021 National Research Infrastructure Roadmap*, Department of Education; *Outstanding research infrastructure powering innovation*, Defence Science and Technology Group.

[2] *The rise of "Big Data" on cloud computing: Review and open research issues*, Information Systems Journal

Purpose

This strategy sets out the desired features of Australia's future NDRI ecosystem to ensure it meets the rapidly growing demands for increasingly complex digital tools and services that NDRI provides. It describes key challenges facing NDRI users and providers, and sets strategic direction to address them. Achieving the aims of the strategy will require coordinated action by all existing and future NDRI stakeholders, and will be driven by National Research Infrastructure (NRI) and government.

NRI whose central objective is to meet the foundational digital needs of researchers, are well positioned to drive nationally consistent approaches to NDRI by providing expertise and leadership in policy and best practice, as well as guiding frameworks and long-term planning.

NRI which focus on the needs of particular research disciplines or communities of practice, are increasingly providing digital capabilities, in line with advances in technology and instrumentation, and therefore have important insights into the specific needs of different research communities.

Government departments and agencies involved in the research sector will need to provide coordinated messages to their stakeholders. These challenges require a coordinated whole-of-government approach, beyond the remit of the National Collaborative Research Infrastructure Strategy (NCRIS) program. The NDRI Strategy will act as a vehicle to promote and anchor broader discussions across government agencies, state and territory governments, research institutions, public sector organisations and industry.

These stakeholders will be essential in driving the national change and high-level outcomes necessary for an optimised and a nationally coordinated NDRI network that is aligned with international developments and best practice.

The NDRI strategy should be used by:

- government to inform future investment decisions and cross-sector planning
- research infrastructure providers to inform strategic planning of data assets, computing resources, digital services and training
- research communities to ensure it reflects their user needs, and to help their users understand the NDRI available, so that digital assets and resources are easily discoverable, accessible, reusable, and interoperable whenever possible and relevant
- other organisations, institutions and agencies in the digital infrastructure ecosystem to inform co-investment and coordinated delivery of NDRI services.

The Strategy complements other government digital strategies, including the Digital Economy Strategy and the Data and Digital Government Strategy, which aim to maximise access and use of data assets by the government, research and innovation sectors.

Implementation and Review

The Government will drive implementation of the NDRI strategy by:

- directing its future investment in NDRI to expectations that promote the strategy's aims
- building existing leadership, governance and policy activities within the current NDRI network into a coordinated response with shared strategic intent and
- promoting close collaboration and shared planning among stakeholders and their wider network.

The Government will also leverage this strategy to catalyse broader conversations towards establishing a national approach to digital research and innovation across government, academia and industry.

Development of the NDRI system must include engagement with Indigenous Australians and acknowledgement of the value Indigenous knowledges can offer to this process.

The Strategy will be reviewed by the NRI Advisory Group annually to reflect the fast-paced nature of technological change in NRI and research sector developments.

NCRIS Investment

This Strategy will provide a guiding framework that shapes future NCRIS investment in NDRI, including:

- as part of general NDRI activities undertaken by NCRIS projects and
- through the Step change identified in the 2021 Roadmap.

The Step change process, expected to commence in 2024, will include focussed consultations and discussions with expert stakeholders to:

- address the challenges and approaches identified in the Strategy and
- identify specific NDRI investment options that implement the NDRI Strategy's vision.

Vision for future NDRI ecosystem

User-centric design must lie at the heart of Australia's NDRI system.

NDRI user communities are increasingly diverse. In addition to researchers, they include a wide spectrum of research beneficiaries, such as policy makers, industry partners and community groups. This increasing diversity of user expertise, digital literacy and backgrounds, creates a need for varying levels and types of NDRI support. This creates challenges for the NDRI system that must be met through the design of NDRI tools, services and frameworks that meet the genuine needs of these different communities, ensuring that the impact and value of investment is maximised. In working towards enhancing Australia's NDRI, all approaches must be delivered in a manner that is as practical as possible for the Australian researcher community to adopt.

Australian NDRI holds a significant role in promoting ethical and responsible Indigenous research across disciplines and methodologies (AIATSIS Code of Ethics for Aboriginal and Torres Strait Islander Research; CARE Principles for Indigenous Data Governance) and improving outcomes for Indigenous Australians.

Due to the underpinning nature of digital capabilities, NDRI also has a central role in making the entire NRI ecosystem more accessible and connected.

Research infrastructure that caters to specific research disciplines and communities of practice are an important part of the digital research infrastructure landscape. Close collaboration with user communities is critical to deploying fit-for-purpose solutions and ensuring responsiveness.

Outcomes

The following six priority outcomes define priority features of an Australian NDRI ecosystem that meets this Vision. These outcomes have been designed to address major challenges that NDRI users and providers will face in the coming years. All these challenges ultimately stem from the growing demands of Australian researchers for advanced computation, data storage and software complexity that are needed to retain global competitiveness.

By 2030, Australia's NDRI system will be:

1. underpinned by training frameworks for researchers and NRI workforce
2. responsive to disruptive technological and societal shifts
3. consistent in its standards for data collection, curation and access
4. integrated across levels of computing and data
5. cybersecure, particularly for national-scale data and computing
6. maximised by openly available research software tools.

These six outcomes are not listed in order of priority and are not mutually exclusive.

1. Underpinned by training frameworks for researchers and NRI workforce

Challenge – Australia's NDRI cannot reach its true potential for impact without access to a suitably skilled workforce by digitally skilled researchers and other users.

The increased importance and demand for NDRI by diverse users is also creating growing demand for a diverse but highly specialised expert workforce that is able to maintain and operate NDRI, maximising the value of capital investment. Global shortages are driving a highly competitive global environment for digital experts, such as research computing specialists, research data specialists and research software engineers, leading to critical reliance on appropriate training and career pathways within Australian NDRI.

The widening user base of NDRI is creating a growing need for training opportunities in the digital skills necessary to ensure that NDRI resources are used to their full potential. Though many training opportunities already exist, they are often not findable and/or accessible to the user base.

Approach - NDRI providers continue and expand training opportunities in suitable digital skills for their wide user base, and providers and government will work to address staff shortages.

Many NDRI providers currently support training opportunities in digital skills for digital specialists and their users. Yet the training community is often disconnected and unable to find best practice training material. Building on existing activities such as Digital Research Skills Australasia (DReSA), NDRI providers must pursue the development of a common framework that aims to uplift digital researcher skills through national and international collaborations, and training opportunities that are clearly communicated to digital specialists and potential users.

Investments in the NDRI workforce should align with delivering research agendas of national significance whilst also supporting domain-specific training requirements that utilise specialised NDRI. Existing Australian Government initiatives, such as the Australian Universities Accord process and the Diversity in STEM (DiSTEM) Review, will also complement this outcome. NDRI investments to support sustainable career pathways should also consider complementary education and training qualifications that already exist.

In developing Australia's NDRI workforce, considerations should extend to incorporating cultural awareness training programs. These may include, but are not limited to, Indigenous cultural awareness training and training around the application of the CARE principles for Indigenous data. With the anticipated expansion of Australia's NDRI workforce, efforts should also be focused towards promoting the involvement of Indigenous Australians, providing training and up-skilling as necessary to ensure the collective benefit from data whilst also improving data collection, management and minimising any potential impact on community cultural practices.

NDRI providers are also encouraged to ensure the provision of career development opportunities for the NDRI workforce to attract and retain staff with the required expertise. This is expected to be a key issue in the NRI Workforce Strategy. The Government will leverage existing and planned activities in skills sector reforms to address shortages and support the training of a digitally skilled future workforce.

2. Responsive to disruptive technological and societal shifts

Challenge – Rapid technological advances and societal shifts can quickly lead to obsolete and ineffective digital resources.

Rapidly evolving digital technologies represent great opportunities to progress research discovery and innovation, but also a profound risk of obsolescence if not appropriately managed.

Known drivers of technological change include:

- The speed of development in High Performance Computing (HPC), which is accompanied by increasingly diverse hardware architectures that create massive demands for data storage, reductions in code portability and other software requirements, as well as increased capital costs and power usage
- The increasing adoption of digitally-driven applications, such as artificial intelligence (AI) and machine learning (ML), is leading to exponential growth in the potential of research to solve complex problems – and the demand for computing, data, software and networking capabilities
- Digitally-driven applications and software are also creating new needs that will be met through diverse computing setups (high-throughput, cloud and edge computing)
- New advanced computing technologies such as quantum computing, and new instruments and sensors which produce ever larger volumes of data of increasing complexity, and the increased availability of these capabilities to a wider range of users
- Trust and identity solutions that are globally aligned to prepare for future cybersecurity risks and technology disruptors
- Climate and environmental challenges will drive the demand for sustainable and energy efficient systems, particularly in accordance with the Australian Government's Net Zero greenhouse gas emission target by 2050

Technological advancements also create societal shifts, and corresponding new public policy, that must be considered to ensure impact of NDRI-supported research and innovation.

Known examples of societal shift and public policies include:

- public acceptance of AI
- UNESCO open science recommendations
- OECD recommendations on data access
- society's attitude towards the collection, use and disclosure of sensitive information
- privacy policy reforms
- government digital and data strategies and requirements.

It is imperative that Australia's NDRI ecosystem is prepared to adapt to these changes and capitalise on opportunities presented. It is also critical that the NDRI strategy is aligned with the broader system of relevant government activities, including the List of Critical Technologies and the National Science and Research Priorities. A whole of Australian Government approach will greatly support the delivery of this outcome and the broader NDRI strategy.

Approach – Nationally coordinated, expert-informed strategic planning of an agile NDRI system.

Coordinated, well-formed planning is essential to ensure Australia's NDRI is prepared to address and adapt to current and emerging technological developments. This includes providing the flexibility for a system that is agile and well-geared to respond to unforeseen shifts, including those not yet realised in the space of AI/ML as well as the ever-growing demand for cloud-based capabilities (commercial or otherwise) that may help address the scalability and plasticity required for Australia's growing NDRI landscape. NCRIS-funded organisations should not be passive consumers of the expressed NDRI needs of the communities they support; they should be active contributors to the discussion, providing expert input where it is needed. As such, the individual and collective expertise of NCRIS projects should be directed towards supporting or spearheading domestic and international activities that respond to disruptive shifts.

The Government should utilise NCRIS investment to drive expectations for shared planning among funded organisations, leveraging existing relevant groupings. It should also have clear expectations that funded organisations remain informed about the NDRI needs of the research communities they support, including users from multiple domains. This will ensure a shared strategy in fore-sighting and responding to emerging and future changes in the interests of Australian NDRI users.

Government must also utilise its contacts, including the NRI Advisory Group and government stakeholders, to ensure that the NDRI Strategy remains current, that investment is sound and cognisant of emerging and future developments, and alignment across NDRI-related government programs and policies.

Further considerations should also extend to Indigenous Australians and any additional or alternative data they may wish to see collected in delivering this outcome.

3. Consistent in its standards for data collection, curation and access.

Challenge – Rapidly increasing rates and volumes of data are being generated across all research fields in formats that are not findable, accessible, interoperable or (re)usable.

Data is at the heart of 'digital' and is both a critical enabler of research and a crucial asset to empower and inspire future discovery. Data is also an intrinsic element of research itself, from shaping the science questions, designing focussed experiments and monitoring programs, through to conduct of the research itself, including testing theories, development of algorithms and models, and presentation of the evidence that will drive research translation. At all of these steps in the scientific process, access to and use of data, together with full contextual metadata, stimulates and empowers collaboration, within and across disciplines. Consequently, access to high-quality data is a fundamental driver of excellence in research and innovation3. This enabling power of data also represents an investment now in future research opportunities, both to extend the scope and timeline for responding to as yet unposed science questions and collaborations, and to test the validity of future hypotheses.

The strategy provides an important opportunity to recognise the intrinsic value of data as an integral component of digital infrastructure, not just through the mechanisms to access, share and process it, but through the policies, tools and skills needed to manage, store and sustain this vital asset for collaboration and inspiration in the present and in the future. Key identified barriers are:

- poor data management practices, including restrictive or local access conditions
- incompatible metadata standards and formats
- incompatible or siloed information systems
- a lack of archiving mechanisms that ensure that data is available for reuse (including in new ways)
- challenges faced in discovering relevant research.

These issues are compounded by the exponentially rising volumes of data created through modern research practices. Large volumes of data can be easily and rapidly generated from diverse applications. This creates immense potential for value through its use, reuse and new application across domains and disciplines. However, this potential is only achievable with careful consideration of the data's whole-of-lifecycle needs.

Many research infrastructure facilities rapidly generate large datasets (e.g. from high-resolution instruments), which increases demand for storage that are unsustainable without data management frameworks – and clear prioritisation of data assets. In addition, it is often unclear with whom the responsibility for the long-term storage of this data lies.

In the research sector, collaboration often maximises returns on investment, which means that digital research infrastructure needs to provide access to diverse creators and users of data across research, government and industry. This creates challenges in differing access needs between the different types of datasets – for example, between open access research and commercially confidential projects.

[3] *2021 NRI Roadmap*, Department of Education

Approach – A sector-wide data management framework that considers the entire data lifecycle and strategies for storage and accessibility.

These issues require a sector-wide data management framework that considers the entire data lifecycle from creation to deletion and includes clear metadata and access standards. The framework will be employed by NCRIS facilities and adopted by members of the wider research community, including (but not limited to) university-affiliated research institutions and government-associated agencies. To be effective, this will require broad consensus and encapsulate best practice without being overly restrictive, considering the diverse nature and obligations of the organisations that create and process research data. Major data producing research infrastructure have an important role to play by being integrated into the data lifecycle from the point of capture. Standards (including those for data collection, curation and access) often have international origins, are well established and contextualised within specific domains of research, hence maintaining a level of consistency with any existing international standards is important for delivering this outcome.

This should also address the importance of curating nationally significant data collections, ensuring the right data is retained in accordance with FAIR principles to enable reuse in research, public policy, and industry, and CARE principles, where applicable, to ensure that Indigenous governance over Indigenous data and its use are respected. Shared data governance arrangements must navigate differing user needs to improve data sharing among government, industry and academia, through the provision of safe data spaces and workflows that attract commercial engagement and that enable the ethical use of sensitive data. Sensitive data must only be used in applications that are lawful and in keeping with community expectations. Importantly, NDRI data practices must uphold Indigenous data sovereignty through the employment of appropriate Indigenous data governance protocols and principles.

This framework will review the guidance in the 2017 ‘Data Availability and Reuse’ Productivity Commission Inquiry Report to enable NDRI providers and research communities to identify and prioritise nationally significant datasets.

The development and adoption of a consistent data management framework across NCRIS and the broader research sector should aim to deliver:

- suitably trained, expert digital specialists needed for implementing and sustaining standards, especially for NDRI
- a unified approach toward data interoperability protocols, that accommodates the distinct requirements of different data communities and international obligations
- easy access to the community-wide tools for data upload and migration
- data access in complex collaborative environments to be as open as possible, as closed as necessary
- consistent use of appropriate informatics standards for data structures, semantics, persistent identifiers (alignment with the Australian National Persistent Identifier (PID) Strategy and Roadmap) and metadata
- coordination of data storage capacity with appropriate levels of redundancy
- planning for the entire data lifecycle, including
    - national data governance policy to meet wider policy needs
    - capture data and meta-data, and appropriate pre-processing

- long-term archiving of nationally significant datasets in accordance with the FAIR/CARE principles, to ensure that Indigenous governance applies to Indigenous data and its uses are respected
- efficient curation of nationally significant datasets
- end-of-life cycle data disposal (accounting for data of varying research value, ranging from transient to longer-term referential or sovereign importance)
- clear delineation of responsibilities across the data life cycle.

Importance should also be placed upon establishing formal partnerships and shared decision-making with Indigenous Australians to develop measures of data governance and capability, including data readiness and preparedness of Indigenous communities and organisations. Where possible, Indigenous identifiers and/or metadata (such as AustLang Codes, Traditional Knowledge and BioCultural Notices and Labels) should be utilised to improve findability of Indigenous data assets for all researchers.

Creating a single consistent framework across all research disciplines will be challenging. The Government will work to focus existing NDRI efforts on the development of research data management policies and frameworks, to ensure a unified approach and common strategic intent across different stakeholders, organisations and projects. NCRIS projects with expertise in this area and that are already leading national and international efforts should also play a leadership role in its design and implementation. It should also be noted that a data management framework alone will not deliver a real-world business model for the storage and archiving of significant research datasets. In implementing the NDRI strategy the Government will need to work with stakeholders to identify the most appropriate solutions to the long-term storage of significant research datasets.

4. Integrated across levels of computing and data

Challenge – Research communities and other NDRI users have varying and rapidly expanding computing, data and software demands but may lack the expertise to identify suitable resources, leading to inefficient and ineffective use.

The increasingly diverse nature of problems that NDRI are helping solve creates growing need for seamless access between different computing, data and software offerings, for users with varying levels of expertise, experience and digital literacy.

In addition to NCRIS-supported facilities, there is a growing array of institutional, international and commercial offerings available to digital research infrastructure users. Selecting the most appropriate resource for the task at hand can be challenging for users who are not experts in NDRI. There is also often need to move between levels of computing and data – for example, across levels of computing when a successful trial needs to be scaled up, and accessing compute for analysis of data. Moreover, many communities require compute that is operated in a specialised manner.

Lack of interoperability and integration between all these resources creates barriers for user access, inefficiencies in resource utilisation, and further compounds redundancies with data and coding.

Approach – Integrated access to different tiers of computing capability and shared data.

To drive the transdisciplinary approaches that will extract the full value of NDRI investment, different research and broader user communities need to easily access different tiers of computing capability and shared data.

Linked national digital infrastructure is needed to allow research communities and other users to move between these offerings at the national and institutional levels seamlessly.

This requires:

- suitably trained, digital specialists to support and optimise NDRI-related research tools,
- scalable computing capability to support large-scale NDRI projects
- frameworks that encourage institutional resources to interface with national capabilities to support a seamless experience for users
- NDRI cloud capabilities that alleviate demand for on-premises HPC and interoperability with the broader cloud ecosystem
- flexible access schemes that enable a variety of users, including periodic and early users, to access HPC and other computing capabilities
- co-location of data and computing, where necessary
- federated data and access platforms, where multiple databases can be accessed as one
- consistent and national approach to services, tools, workflows, and governance
- integration between major data producers and data processing and storage capabilities that allows automation
- specialised and fit-for-purpose computing solutions that leverages national-scale capabilities
- a more connected governance framework for national-level (Tier-1) HPC and data to support efficient shared planning of upgrade and replacement needs

- trust and identity services that enable seamless and consistent access across the system, supporting movement between levels of computing and infrastructure, and preventing inefficiencies in resource utilisation and redundancies with data and coding.

The technical challenges associated with integration of such a wide array of resources are considerable.

A staged approach will be necessary, building upon existing NDRI efforts to develop locally seamless user experiences that can then be elevated nationally. These local efforts must be well connected to ensure compatibility and better integration in the future. Extensive coordination with other stakeholders in the research ecosystem will also be necessary, in particular those that have developed their own digital research infrastructure capabilities. This further highlights the need for better connected NDRI governance frameworks.

Systematic trust and identity services and access management frameworks will be fundamental to delivering secure access to high value digital resources. They will assist to standardise access to national and institutional compute and storage resources to simplify transition between tiers. They will also promote the secure connection of Australian datasets with related international datasets.

5. Cybersecure, particularly for national-scale data and computing

Challenge – Cybersecurity threats are becoming increasingly sophisticated while the sensitivity of research data is increasing.

Recent years have seen the targeting of valuable digital resources and sensitive data through increasingly sophisticated cyber-attacks. The far-reaching consequences of such threats that impact personal privacy, innovation, national security and sovereignty, are driving the need for more coordinated cross-sectoral responses.

As a central component to Australia's research community, NDRI are a treasure trove of highly valuable resources, due to their immense potential for value creation and the increasing strategic importance and sensitive nature of the types of research data handled. This strategy also recognises that the higher security treatment of data may conflict with the principles of an open science agenda. In accounting for this potential conflict of principles, a selective treatment approach tailored towards specific data sets (for example, sensitive data) could be considered as well.

Without a holistic approach to cybersecurity that builds resilience across Australia's NDRI ecosystem, our digital infrastructure and their users will become increasingly vulnerable to such threats.

Approach - NCRIS provides tools, frameworks and resources for assessment and mitigation of risk posed by cybersecurity threats – underpinned by a system-wide solution for trust and identity.

As leaders in NRI and custodians of highly valuable digital resources, NCRIS facilities should prioritise the development of tools, frameworks and resources that enable the assessment and mitigation of risk posed by cybersecurity threats. Activities will include:

- benchmarking and baselining NRI cybersecurity posture
- frameworks designed to mitigate risk to resources, including alignment to institutional policies and procedures that set out responsibilities and reporting requirements
- mechanisms to train staff and raise awareness and best practice among their diverse user base
- a community of best practice across NCRIS that supports coordinated efforts; and
- implementation of technologies that protect NDRI resources and their users, including (but not limited to) source code management, reproducible software builds and curated software libraries.

These activities will be grounded in best practice approaches to combat threats. A better collective understanding of these threats will be essential to ensuring responses are proportionate and aligned. These should align with broader national efforts, such as the Australian Higher Education Cyber Security Service, and government expectations, such as in the Guidelines to Counter Foreign Interference in the Australian University Sector. Additionally, a whole of Australian Government approach towards establishing, certifying and maintaining an accredited cybersecurity network could serve to further protect Australia's NDRI. For example, this could see the adoption of advice and leveraging of expertise offered through the Information Security Manual published by the Australian Cyber Security Centre within the Australian Signals Directorate.

A critical step will be to augment current activities to embed trust and identity capabilities across the current NCRIS system. This work should evolve into an ongoing NDRI capability that ensures:

- provision of an NDRI-systemwide solution for trust and identity, which may consider a collaborative and domain-orientated approach
- alignment with the national and international cybersecurity environment
- support for the development of trust and identity incubators that ensure trust and identity approaches mature along with NRI technologies; and
- good privacy practices that are built-into the design of NCRIS resources.

A holistic trust and identity capability will standardise and simplify access for the research community to deliver a more secure, efficient and usable research infrastructure. Trust and Identity services and frameworks will help strengthen security for access to sensitive and open data and support FAIR data.

In addition to supporting the NCRIS sector, these activities will create convenient on-ramps for other institutions and organisations to engage, driving wider coordination and a more unified approach to the constantly evolving challenges and threats.

6. Maximised by openly available research software tools

Challenge – Software is critically important for impactful research, but its place in the NDRI system is not well defined.

Software has become an essential component of modern research, enabling insights and collaborative practices that underpin scientific progress and innovation[4] - but its creation is often an invisible output of the research process. In the NDRI ecosystem, research software is an inextricable component of digitally supported activities, yet its position within this system is not well defined or acknowledged.

Research software needs are as varied as the research projects supported by NDRI. Access to professional research software developers differs across the sector, depending on research funding and university support systems. The software itself can take many forms, from simple single line commands to complex platforms linking multiple software packages together. All research software must be cybersecure by design. This includes the use of trust and identity services and frameworks as a foundational component.

Software engineering is a specialised field and researchers may not have the ability to develop software themselves. In some cases, software development requires both software expertise and a deep understanding of the particular needs of a research community, such as climate or geoscience research. Although not all needs can be met by the NDRI system, for some nationally significant research, the scale and complexity of software demands call for support at a national level. Maintaining the attractiveness of non-commercial research institutes to software developers and IT professionals in Australia is critical for the delivery of this outcome.

The rapid evolution of computing and data technologies compounds these problems, often leading to complex needs for coding portability and the curation of outdated research software. Other limitations to the use of research software include rules of use for commercial software, particularly on Tier-1 HPC or cloud compute and IP considerations.

The diversity in need and the 'invisibility' of software already created results in software developers and researchers needing to create custom software solutions that take significant time and effort to maintain. The need for researchers to reprogram code to use HPC capabilities has been flagged as an inefficient use of time and resources. NDRI has an important role to play in supporting the maintenance and availability of software to maximise the impact and value of NDRI services.

Approach – Research software tools should be recognised as critical NDRI of equal importance to computing, data and networking services.

As with computing and data, the NDRI system is not able to support all researcher software demands but should enable dedicated software engineering support for nationally significant research projects of large scale and complexity. An example of such a project is the ACCESS-NRI NCRIS facility. NDRI support should also be provided for the maintenance and curation of critical software that is needed to ensure access and usage of significant digital assets. Software engineering capabilities must include best practice that integrates security at each phase of the development

[4] *Recognising the importance of software in research*, European Commission Publications Office

cycle. In addition to supporting the development and maintenance of research software, NDRI can also facilitate Australian research communities in accessing suites of already available, configured and connected software, for example by providing local configuration and portability on relevant Australian computational platforms. Due to the scale of demand software effort should be deployed where it has the most impact – for example, a pre-processing workflow alongside an instrument will benefit all users of that instrument.

To ensure research software is recognised as critical NDRI, it must be visible, shaped for wider re-use, while critical research software infrastructure must be sustained (as detailed in the National Agenda for Research Software)5. NDRI should support strategies that enable this three-pronged approach to improve software availability for researchers. Providers of software tools should also be guided by the principles described in the UNESCO recommendation on open science. Making software as openly accessible as possible will enable its reuse, reducing duplication of work and optimising software usage and maximise the societal value derived from the broader NDRI investment.

[5] *A National Agenda for Research Software*, Australian Research Data Commons