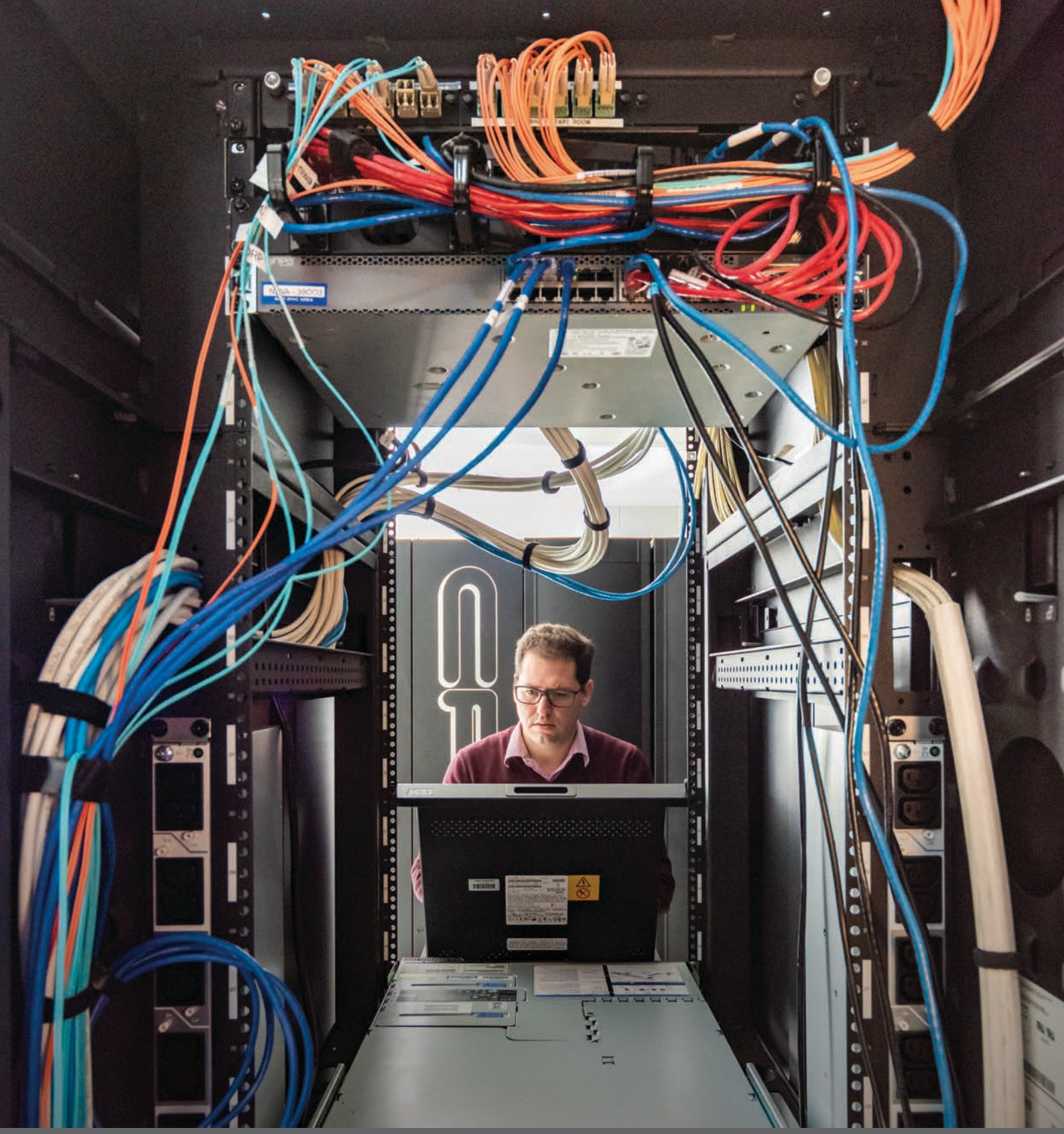


Infrastructure and Related Services to Support New Zealand's eResearch Future


New Zealand eScience Infrastructure

Contents

Executive summary

BACKGROUND

eResearch is a prerequisite capability for a world-class Research, Science and Innovation sector

1. Data- and compute-intensive research and analysis is now deeply embedded in most academic disciplines and is an essential component of modern research and national research capabilities. New technologies are enabling explosive growth in data collection, data storage and opportunities for deep data analysis. Ongoing rapid expansion in computational capabilities and capacities, allied with similarly rapid expansion in software capabilities and accessibility, opens new opportunities for the extraction of knowledge from data and modelling complex systems.

2. Internationally, enormous investment is occurring in fields such as artificial intelligence (AI), machine learning (ML) and deep learning (DL), opening up new frontiers of human progress across both the traditional scientific fields and increasingly, in non-traditional (in terms of the application of big data analysis) disciplines in the humanities and social sciences. Importantly, data sciences are increasingly being applied in cross-discipline/ multi-disciplinary fields, allowing new insights in complex systems.

3. Collaborations, both locally and internationally, go hand in hand with these new opportunities. Similarly, national scientific credibility is increasingly becoming associated with high-quality capability in these rapidly evolving fields.

CURRENT STATE

Aotearoa New Zealand has made initial, modest investment in eResearch infrastructure

4. Aotearoa New Zealand has supported the development of eResearch capabilities over the past 10 to 15 years with investment in the underpinning hardware – high-performance compute (HPC) facilities and high quality connectivity provided centrally by New Zealand eScience Infrastructure (NeSI) and Research and Education Advanced Network New Zealand (REANNZ), respectively, for the use of local research and education providers. The hardware – supercomputers and high-capacity, low loss data cables – is necessary, but to grow and sustain an active eResearch community, a range of other services is required as well, including identity management and education and training in the use of eResearch facilities, robust data curation and storage and sophisticated software engineering capabilities.

5. This commitment to providing central eResearch infrastructure and associated services is common across the governments of other developed economies. In most comparable cases internationally, direct government funding covers a higher proportion of costs than is currently the case in New Zealand.

6. There has been substantial Crown investment in eResearch infrastructure in New Zealand, as well as commitments from a small number of research and education providers. However, building capability and a community of practice in this field has been challenging. Active participation in, and funding for, eResearch remains largely the preserve of the University of Auckland, University of Otago and the National Institute of Water and Atmospheric Research Limited (NIWA). Other researchers are light users and contributors in this area, and some of them prefer to rely on commercial or in-house facilities rather than pay for and actively engage in the centralised services.

Capability shortages have limited the uptake of eResearch, despite subsidisation of services

7. The observation that a heavily subsidised set of services struggles to attract support from significant intended users is troubling. It is, however, consistent with an observation that eResearch capability across our research and education sectors largely remains at a quite immature level, outside the major system collaborators. It seems that many of our key research institutions have not seen the development of eResearch capability and its application as a strategic priority in their work.

8. These observations provide an uncomfortable backdrop to the upcoming decisions on future Crown investments in this field.

FEATURES OF AN IDEAL FUTURE STATE

eResearch infrastructure is accessible, responsive and flexible to the needs of all researchers

9. An ideal future state for our eResearch set-up would involve an active, growing and sophisticated community of eResearch practitioners across our Crown Research Institutes (CRIs), universities, independent research organisations and other tertiary education providers. It would involve active collaboration in eResearch projects, across institutions and across disciplines. New Zealand eResearchers would be sought after as collaborators in international projects.

10. The local infrastructure would be considered researcher-friendly, responsive to emerging demands and capable of meeting those demands. Access rights would be clear, with prioritisation seen to be aligned with agreed principles.

11. It would demonstrate responsiveness to Treaty of Waitangi principles. Māori interests in data sovereignty are well expressed. An eResearch platform with a leadership role in research data curation could engage constructively with Māori to further those interests. Another significant point of intersection with Treaty responsibilities relates to the development of eResearch skills. Here, in conjunction with universities and research institutions, the eResearch platform could play an effective role in facilitating Māori research interests through the development and enhanced application of data skills.

Costs are transparent to all users

12. The cost of providing eResearch services would be clear to all parties, including the Crown. The extent and form of co-funding expected of the research community would also be clear and judged by users to represent good value. These aspects of funding are essential to the viability of a national eResearch infrastructure.

Researchers are supported by an eResearch community with skills and experience in all aspects of the eResearch lifecycle

13. A strong eResearch community is based on the skills, experience and opportunities that researchers have. Building that community is a foundational task lying, in the first instance, with our universities. Core skills in statistical methods and data sciences are integral to virtually all modern disciplines. Beyond that is a set of applied skills related to the conduct of data and compute-intensive projects in an HPC environment. Development of these applied skills may be a task that should sit alongside the compute infrastructure.

14. A strong eResearch environment starts with data. The collection, cleaning, curation and storage (and structured deletion) of data, as well as ensuring efficient and safe access to it for approved research purposes, is foundational. While there are examples of high-quality research data curation across our research system, standards are patchy overall. Too much expensively collected data is left in informal or unstructured storage at the end of projects. As such, it is in danger of being lost to future researchers, unprotected in terms of future access, and missing adequate descriptions of its origins and essential characteristics. A future eResearch platform should include resource for leadership, support and guidance regarding the robust curation of research data. While the data may remain the primary responsibility of its originators, there is a role for a collaborative approach to the development and application of core curation principles, protocols and procedures. Engaging collaboratively and effectively with issues of Māori data sovereignty is integral to that function.

KEY, CHALLENGES, LIMITATIONS AND OBSTACLES

15. Obstacles to achieving that desired future state are many. Funding is a perennial challenge throughout our research system, exacerbated by the small scale of both the aggregate system and its component organisations. Other contributing factors include the relatively light connections with the industry and corresponding light private sector funding, modest income potential in New Zealand for high-performing researchers (compared to opportunities abroad) and the increasingly capital-intensive nature of many aspects of research.

16. Some of those factors are essentially contextual, given our country's size and location. However, their existence obliges us to be careful to avoid adding further obstacles by way of poor choices and poor prioritisation.

17. Specific challenges in the eResearch sphere are discussed in the following sections.

Current institutional arrangements are not sustainable in the long term

18. REANNZ is conventionally constituted and governed. Its institutional arrangements appear to be sound and functional. Its services are widely, but not universally, appreciated within the system. However, it faces a longer-term structural revenue deficit, with limited opportunities to boost income.

19. NeSI exists as a contract-based collaboration, with the University of Auckland as host. NIWA and the University of Otago are major partners and funders. It has a Board of Directors, but this is advisory only. Personnel are formally employees of the investing members.

20. NeSI's successes have come despite its institutional form. Its structure is a long way from ideal, given the need for strong strategic focus and agility in execution in a world of rapid technological innovation, as well as a client base of highly disparate needs, interests and financial capacities.

21. NeSI's current structure complicates the strategic drive and governance needs of the organisation. To function, it relies heavily on the goodwill of its primary partners. While that is generally forthcoming, it is clear that not all key parties share a common view on NeSI's future. Moreover, the reliance on in-kind contributions of personnel and other services ensures that costs and effective revenues are quite obscure.

22. It will be difficult to pursue an ambitious future while operating under the current institutional form.

The 'club' model struggles to accommodate the diversity of needs and interests

23. The current institutional arrangements are based on a 'club' model, with members essentially paying a membership fee to access the infrastructure, together with some additional user charges.

24. Club models work when the club members have near common interests, needs and ability to pay. This is clearly not the case across our eResearch community. The disparity in interests with respect to HPC capacity is wider, or more intrusive, than with the connectivity services. To be successful, future services (and associated costs) must be more closely tailored around the interests of each participating institution.

25. The awkwardness of the club model is exacerbated, in the case of HPC capacity, by NeSI's institutional form and the associated funding and pricing opacity.

Rapid technological advancement challenges the existing operating models

26. The concept of central eResearch infrastructure provision is further challenged by the very rapid progress of both compute and connectivity technology and the associated development of commercial service options for research institutions and researchers. These will squeeze the central facilities from both ends.

27. At the smaller end of the research spectrum, the cost of high-quality/high-capacity desktop and even laptop computing power continues to decrease. Similarly, commercially available connectivity options continue to gain capacity and price competitiveness. Cost-conscious researchers and smaller research institutions will continue to find those options attractive. Universities have a role to play in shaping whether their researchers conduct their work on in-house and other alternative systems, or they are directed and/or encouraged to make use of the central facilities and allied support ecosystem.

28. For large-scale research projects, options to work in the Cloud are expanding rapidly. For New Zealand, these options will come a step closer with the imminent establishment in this country of several Cloud providers. Depending on local demand, these services could come with sophisticated HPC, packaged with software and connectivity options as sought by researchers, as well as high-quality cybersecurity and other support services.

The limited scale of New Zealand's eResearch ecosystem increases costs

29. In this field, scale matters. Without scale, New Zealand research institutions face higher costs to participate in eResearch than their international peers in larger economies and have a narrower range of facilities and capabilities to work with. These factors emerge from both the lack of economies of scale in standing up large HPC facilities and the limited number of institutions across which the costs can be distributed.

30. As usual, New Zealand is obliged to be carefully selective in where it invests and how it can best collaborate to reduce the impact of our small scale. Leveraging opportunities to have New Zealand researchers engaged in international projects of scale will provide one way of remaining close to the leading edge of research.

OPTIONS

31. Options for future eResearch platforms are largely shaped by levels of ambition, and ambition is largely associated with the scale of available funding. The backdrop to the choices presented in this section is that the current arrangements are not financially sustainable and NeSI is structurally unstable. The current arrangements are not delivering the vibrant eResearch environment to which we should aspire.

32. There are options and choices with respect to future function and form in our eResearch infrastructure. The attitudes and preferences of the key participants will shape exactly which to adopt. A simple graduated range of reform options is presented in this report, as follows:

Option 1: Status quo+

33. This is a low-ambition, low-additional-spend, limited-impact option in which REANNZ remains as is and NeSI is converted into an independent Crown-owned entity structure (or a similar form with mixed shareholdings) with a renewed board that is charged with reviewing the HPC asset mix, future service offerings and pricing regime. The flexi HPC model pioneered with AgResearch is developed further, and the 'portal' model for adding access to the Cloud is explored.

Option 2: Incremental expansion

34. This involves some added cost for added services and ambition. REANNZ remains as it is and NeSI is reshaped, as in Option 1. Enhanced funding is provided for eResearch skills development, including working with universities. The flexi HPC model is developed as in Option 1, and the use of Cloud opportunities is explored. Additional funding is provided for leadership in research data curation by supporting the development of a community of practice among the pockets of expertise across the system.

Option 3: Integrated eResearch platform

35. In this option, a scaled-up Data and Digital Research Institute (DDRI) is created as a new eResearch sector platform, formed as a not-for-profit Crown-owned entity (or mixed shareholding) with a suitably skilled and experienced governance board. REANNZ and NeSI are folded together as a 'one-stop shop' for eResearchers. As in Option 2, it provides added functionality in eResearch skills development, research data curation and leadership role and research software engineering capability. The DDRI becomes the national champion of eResearch and its possibilities, working closely with the broader education and research community. Crown funding and the extent of cost recovery across the system determine the scale of this ambition.

Recommendations

36. The Report Panel makes the following recommendations:

    I. The Government pursues Option 3, the Integrated eResearch Platform (or a variant on that concept), with investments intended to create a DDRI.

    II. The DDRI should be established as a not-for-profit Crown entity (or similar form).

    III. The DDRI should be funded sufficiently to deal with the existing contractual commitments challenging REANNZ and NeSI.

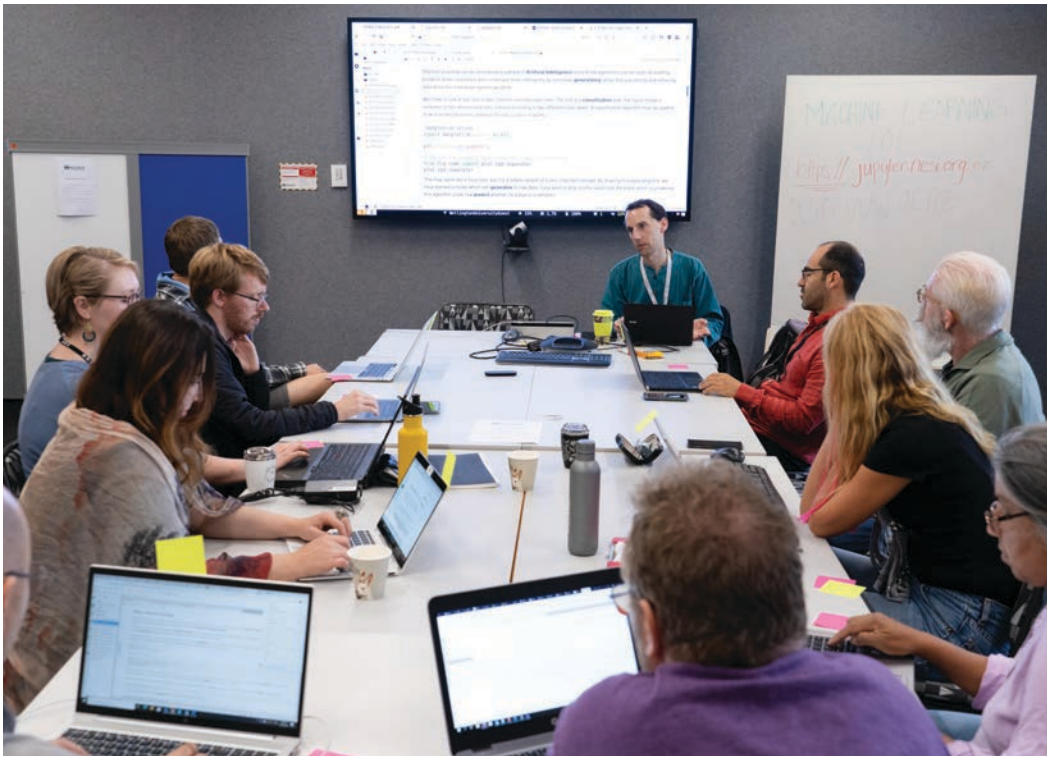
IV. The DDRI should be funded to enable it to unwind existing in-kind and other sector contributions, with the intention that the DDRI will purchase or contract services from collaborating institutions (or others) as required to conduct its business efficiently. That will likely involve some reassignment of funding that is currently going directly to collaborating partners.

V. The DDRI should be funded to support targeted skills development programmes aimed at building capability in eResearch across the education and research establishment. This should be done in collaboration with universities as they increase their commitment to the field.

VI. Technical support and advice for eResearchers is integral to the eResearch platform. The capacity to provide such support, including on research software and the associated engineering skills, should be part of the DDRI offering.

VII. The DDRI should be funded to provide support across the research community for enhanced research data curation. This would help to develop national standards and support research institutions to adopt and apply improved practice in data curation.

VIII. User charges should apply (at least at the institution level) across the connect and HPC usage, aligned with use but at a lower proportion of total costs than is currently applied. There should be no user charges for skills development or research data curation functions.



Introduction

This report has been commissioned by the Ministry of Business, Innovation and Employment (MBIE)

37. The key question this report aims to address is:

    What are New Zealand's future needs of infrastructure and related services to support eResearch, and how can the New Zealand eScience Infrastructure (NeSI) and Research Education Advanced Network New Zealand (REANNZ) best position their roles and capabilities with the system to support these?

38. The scope of the report, as laid out in its Terms of Reference document, includes:

    - the current state of play for eResearch activities
    - the strategic context, including key future trends, international comparisons, New Zealand's eResearch problems and opportunities and New Zealand's future eResearch needs
    - the future state: that is, how New Zealand should position its high-performance compute (HPC) capabilities and the related services, data management and storage, into the future
    - the roles that the Government, REANNZ and NeSI should play relative to other institutions in the eResearch system
    - any implications for the funding mechanisms that support the eResearch system.

39. To help focus the report, the Terms of Reference described the following intended end use of the report:

    - Support collaboration in developing eResearch capabilities across institutions by offering a widely engaged upon future state which is specific enough to offer a framework for decision taking.
    - Support institutions' investment planning to allow a more coordinated approach.
    - Support further policy work around role implications and related funding mechanisms for NeSI and REANNZ.
    - Support decisions on future Crown investment in eResearch.

The report is not intended as a 'scorecard' of performance

40. The Terms of Reference were also specific in ruling out of scope the following elements:

    - The report is not intended to be a 'report card' on performance – it will only surface such issues to clarify the extent of the shift needed for actors to position their capabilities for the future.
    - The report will not cover the REANNZ connect network in detail – this will be considered as part of REANNZ's network upgrade work.
    - The report will not make recommendations on funding levels, although its findings may have implications for this area.

41. The full Terms of Reference, which are attached as Appendix 1, also contained other relevant elements beyond these points.

CONTEXT

The report is intended to support a more joined-up approach to New Zealand's eResearch

42. The Terms of Reference for this report directed us to inquire into "infrastructure and related services to support New Zealand's eResearch Future". The report is intended to support:

    - collaboration in developing eResearch capabilities across institutions
    - investment planning and enabling participating research institutions "to allow a more coordinated approach"
    - policy development related to roles and funding mechanisms for the key eResearch infrastructure providers (i.e. NeSI and REANNZ)
    - future Crown decisions on investment in eResearch.

43. The term 'eResearch Infrastructure' relates primarily to the hardware on which data- and compute-intensive research is conducted and shared. In New Zealand, this includes the HPC capacity provided to the research community (via NeSI) and to the connectivity network, both domestic and international (via REANNZ).

44. The term 'eResearch' is not well defined but can be taken to encompass the increasingly wide span of data-intensive and/or complex-system modelling research, as well as related fields that rely on HPC and, often but not always, high-capacity connectivity. Much of this work is likely to be conducted collaboratively, often involving internationally dispersed partnerships.

45. It became clear when consulting on the draft of this report that the terminology in this field is not settled. The terms we have used are intended to be broadly inclusive of activity applying digital tools and techniques to advance research. The 'related services' are not defined but can be taken to include education and training in data science and the use of the eResearch infrastructure; data curation and storage; research software and its engineering; and cybersecurity issues.

Why does eResearch and the supporting infrastructure matter?

46. eResearch is rapidly becoming a core capability in most branches of modern research and science. New technology is enabling explosive growth in data collection and data storage and thus, its availability for data analysis. Allied with that is the ongoing improvement in computation capability and capacity, together with advanced software enabling ever-greater sophistication in analytical techniques for the extraction of knowledge from data, or the creation of data from complex-systems modelling. Artificial intelligence (AI) and machine learning (ML) applications across multiple fields are current leading examples of these developments.

47. While these developments have their roots in the physical and mathematical sciences, they are 'general purpose' in nature. This means the techniques and infrastructure of eResearch are not limited to particular disciplines – they can be applied wherever there is data available for analysis, or where complex systems are modelled. This can be seen in their nascent spread to the social sciences, humanities and beyond, as researcher curiosity and creativity allow the application of these evolving techniques and capabilities to cross the boundaries of traditional disciplines.

eResearch has become a core prerequisite capability to push the frontiers

48. For any nation's research community to be relevant, credible and an attractive collaboration partner, high-quality eResearch capability and capacity is now a primary coin in the research treasury. Building a cohort of highly skilled researchers with deep experience of eResearch analysis is core to an ambitious and high-performing science system, including being able to participate in, and contribute to, emerging international science projects.

49. Likewise, within our domestic research setting, eResearch capabilities have the potential to yield new insights in most fields in which contemporary challenges to our wellbeing may be found, such as precision medicine, new approaches to designing effective interventions in education and social policies, reductions in agricultural greenhouse gas emissions and better understanding of climatic or geologic processes.

Current Arrangements

NEW ZEALAND'S ERESEARCH INFRASTRUCTURE CENTRES ON TWO KEY ENTITIES

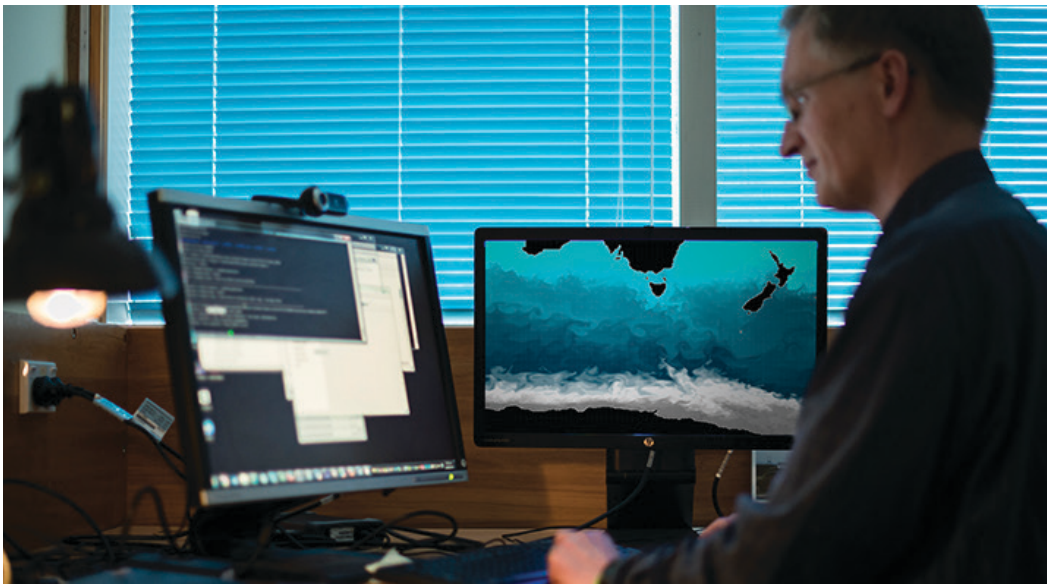
REANNZ provides high-quality international connectivity for the research and education sectors

50. REANNZ provides the 'connect' component of the eResearch infrastructure. It was established in 2007 as a Crown-owned company under Schedule 4A of the Public Finance Act. It operates as a not-for-profit company with closed membership (i.e. limited to those in the education and research sectors).

51. In effect, REANNZ is a specialist internet service provider joining education and research centres across the country with an advanced network of high-capacity cables. These enable researchers to transfer large volumes of data quickly and reliably. The specialist element of this network is its ability to cope with occasional very large bursts of data, which may extend over days. Importantly, this must occur without data loss or impairment. This service cannot be readily matched by commercial cable providers, whose business models and systems are set up to manage very large volumes of individually small and fairly predictable data flows, with greater capacity to manage loads and less emphasis on the quality of the transfer.

52. To provide these services, REANNZ maintains both 'owned' and leased cable capacity (and associated network equipment), linking the country's primary research establishments (universities and Crown Research Institutes [CRIs]), polytechnics, wānanga and independent research organisations. REANNZ was a cornerstone investor in the Hawaiki cable linking New Zealand with the United States of America, taking a 25-year lease (until 2042) on 1,000 Gpbs (gigabits per second) capacity, with 100 Gpbs currently in use. REANNZ continues to be liable into future years for substantial payments in relation to the Hawaiki cable.

53. REANNZ is New Zealand's designated National Research and Education Network (NREN) (see Box 1). Participating countries typically have one designated NREN operating on behalf of their education and research institutions in their respective countries.



BOX 1: THE NATIONAL RESEARCH AND EDUCATION NETWORK

A NREN is a specialised internet service provider that connects the research and education sectors of a country to other researchers and science infrastructure, both domestically and across the world. They achieve this by undertaking the following roles:

- **Specialist network operators:** supporting a high-speed backbone network to provide a seamless, real-time transfer of critical research data on a global scale.
- **Global connectors:** over 120 NRENs globally working together to connect the global research community through the provision of high-speed networking, access and identity services so that researchers can connect and collaborate wherever they are.
- **Technical experts:** providing advice and services to enable researchers to maximise their use of global research networks and access key science infrastructure.

NRENs operate differently from commercial internet service providers. They have a collaborative ethos, allowing them to form a global network linking national connectivity infrastructures to provide seamless international data connectivity for researchers. The NREN network can collaborate to shepherd very large data transfers quickly, safely and efficiently from point to point globally.

54. Given New Zealand's small scale and geographic isolation, high-quality connectivity, both domestically and internationally, is an essential element in our education and research platforms. Facilitating ready interaction with researchers across the globe and participation in large-scale research projects based in other countries is integral to ensuring our researchers can stay close to the leading edge in their various fields.

55. REANNZ provides a suite of ancillary services supporting its core network infrastructure. These include:

    - Tuakiri and eduGAIN: Federated identity management systems that allow researchers to access resources (e.g. academic journal subscriptions, log-in to the NeSI infrastructure), using their home institution's identity.
    - Eduroam: Provides secure global roaming access for researchers, enabling them to access internet services at any participating institution worldwide, using their 'home' log-in procedures and without incurring fees.

NeSI provides New Zealand's HPC platforms and skills development programmes

56. NeSI provides HPC capacity at a national level. It was established in 2010 as a contract-based collaboration among the University of Auckland (as host) and Manaaki Whenua Landcare Research Limited (Manaaki Whenua), the National Institute of Water and Atmospheric Research Limited (NIWA), and the University of Otago (as investing institutions). The University of Canterbury was an investing partner initially but has since left the arrangement.

57. NeSI currently provides two main platforms:

    - Mahuika, a *capacity* HPC resource (i.e. a cluster) that can run many small compute jobs simultaneously
    - Māui, a *capability* HPC resource that can run complex simulations and calculations.

58. These platforms are housed at NIWA's Greta Point facility in Wellington and operate as an extension of NIWA's dedicated supercomputer capacity. A third supercomputer, Kupe, is housed at the University of Auckland's Tamaki Data Centre and acts as back-up for Mahuika and Māui.

59. While close to 40 institutions and around 600 users are engaged as users of NeSI HPC facilities, utilisation is overwhelmingly dominated by three core collaborators: the University of Auckland, the University of Otago and NIWA (see Figure 1). This presents a key challenge for New Zealand's eResearch future. New Zealand has few groups working on the larger scale that this infrastructure is intended to service. These are projects involving sophisticated analysis and modelling of complex systems and utilising very large data sets. Despite this, the NeSI HPC facilities are currently operating at near full capacity and face strong growth in aggregate demand (see Figure 2).

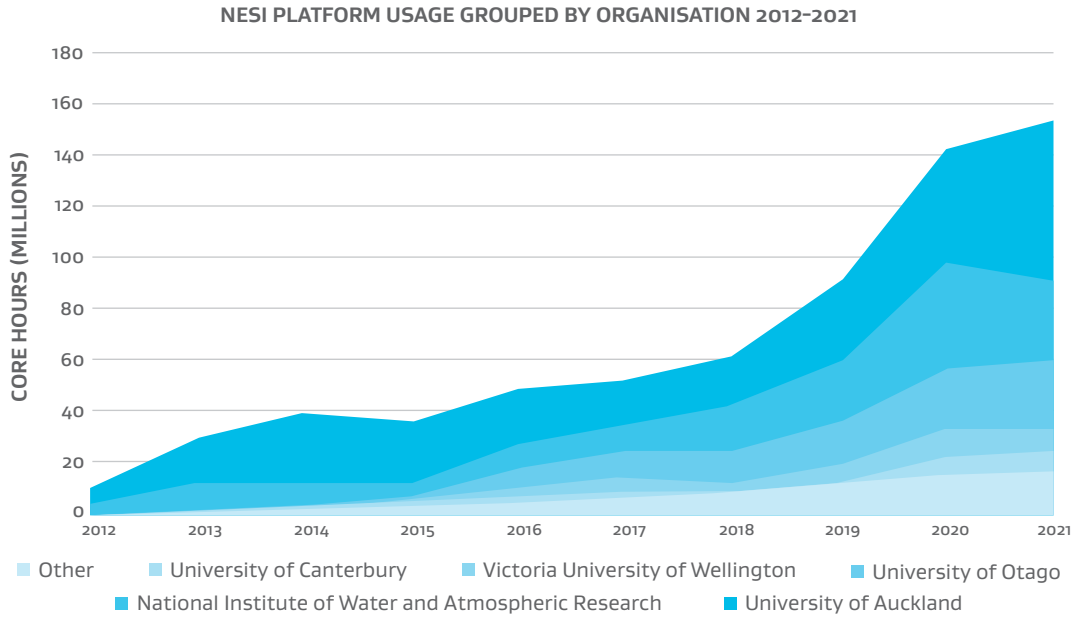


Figure 1: Three Core Collaborators Dominate the Use of NeSI's HPC Capacity

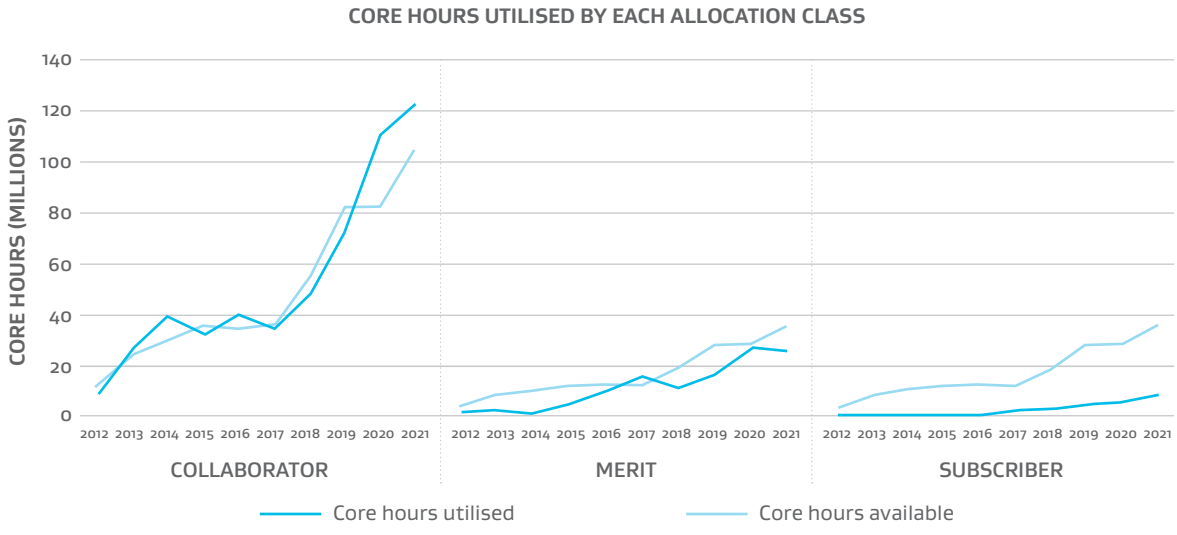


Figure 2: Usage Compared to Theoretical Capacity of NeSI's Compute Platforms by Allocation Class

60. NeSI also provides well-appreciated skills development and technical support programmes for researchers. Apart from their own service support team (9 FTEs), NeSI draws on expert professionals employed by collaborating institutions (especially the University of Auckland, the University of Otago and NIWA), who cover a wide range of topics and levels of sophistication.

61. These skills development and technical support programmes are an integral part of the outreach that NeSI has embarked on to encourage greater use of their HPC facilities.

62. NeSI also provides support for:

- institutions on their research data management challenges (Genomics Aotearoa, in particular)
- AgResearch on its flexi HPC environment (see Box 2)
- research software, its application and its engineering.

63. At this stage, these are relatively small initiatives and they are lightly resourced. However, they indicate the work ahead for the system if research data management is to be brought closer to modern standards, and the potential advantages of a more joined-up approach to HPC provision.

64. Auckland University, as host of NeSI, puts significant effort into enabling the NeSI model. It is directly funded to the extent of $2 million per year for this purpose. This funding covers the salary of a full-time Director, and other resources such as project management, technical, accounting, human resources, communications and legal.

65. NIWA is a key collaborator and influence in the NeSI field of interest. As host of the NeSI HPC capacity, as well as being a very high user of both its own and NeSI's capacity, NIWA inevitably has a major interest in the current usage of the NeSI facilities and future investment intentions. By virtue of the skills and experience within its own personnel, NIWA is a significant provider of skilled support for NeSI and an important influence in the wider national eResearch landscape.

BOX 2: AGRESEARCH'S PARTNERSHIP WITH NESI

In 2020/2021, NeSI entered an agreement with AgResearch to host and manage AgResearch-owned infrastructure in the Tamaki Data Centre as part AgResearch's eResearch infrastructure refresh. Prior to this, AgResearch's eResearch platforms were fragmented and had reached capacity due to the exponential growth in data production, with forecast data storage needs projected to grow at 30 percent per year in the medium term. Single points of failure for key personnel supporting the infrastructure and scientists undertaking eResearch also existed.

By entering an arrangement with NeSI, AgResearch gained:

- access to systems expertise across different areas of infrastructure, security and virtual environments
- an increased level of capability and service to its user base
- alignment in service delivery between AgResearch's infrastructure and NeSI's national clusters, making the transition from local scale to national scale almost seamless
- reduced training needs for onboarding, accessing software, tools and knowledge
- a team of established engineers, which reduces key personnel risks.

The arrangement may also provide a model for other research institutions to leverage capabilities, standardise environments and obtain access to skills that are difficult to procure and retain at an institutional level.

Current Funding Models

66. The Terms of Reference for this report stated that it should not make recommendations on funding levels for our eResearch Infrastructure. However, they also stated that it could identify any implications for funding mechanisms to support the system.

REANNZ

The REANNZ funding model involves a balancing act between a small number of members and significant fixed operating costs

67. REANNZ operates a form of club membership model – that is, members pay a fee to join the REANNZ 'club' and are then able to access the facilities provided by the club. However, this funding model poses a significant risk to REANNZ, as there is limited scope to increase the number of members and member fees operate on a one-size-fits-all model. Members contributed, on average, around 40.7 per cent of REANNZ revenue over the last four financial years (see Table 1), while the Crown has contributed $3 million annually via the Strategic Science Investment Fund (SSIF).

68. There has been little change in these fees over the years. Financial pressures, the availability of cheaper alternatives (albeit with lower capacity and lower service levels) and other issues have led some to question the value of the service (see Figure 3). Ultimately, three universities left the group, rejoining in 2017/2018 after successfully negotiating lower fees. With reduced revenue, REANNZ has had to run deficits in order to maintain its services.

Table 1: REANNZ Financials from 2018 to 2021

	2018[1] ($M)	2019[2] ($M)	2020[3] ($M)	2021[4] ($M)
Revenue				
Crown	3.0	3.0	3.0	3.0
Hawaiki Cable	5.3	3.0	–	–
Members	8.0	6.1	6.1	6.2
Other*	5.2	5.2	4.6	4.5
Total	**21.5**	**17.3**	**13.7**	**13.7**
% user funding	**37.2%**	**35.2%**	**45.2%**	**45.2%**
Network expenditure				
Depreciation and amortisation	2.0	2.2	2.2	2.1
Employment expenses	2.3	2.4	2.0	1.9
Network operating expenses	8.1	11.3	7.8	8.2
Total	**12.4**	**15.9**	**12.0**	**12.2**
Operating expenditure				
Depreciation and amortisation	0.2	0.2	0.1	0.1
Employment expenses	2.2	2.1	1.6	2.0
Other operating expenses	1.5	1.7	1.6	1.1
Total	**3.9**	**4.0**	**3.3**	**3.2**
Total expenditure	**16.3**	**19.9**	**15.3**	**15.4**

* Other includes managed services, internet and interest revenue.

[1] REANNZ (2018), REANNZ Annual Report 2018, Year ended 30 June 2018.

[2] REANNZ (2019), REANNZ Annual Report 2019, Year ended 30 June 2019.

[3] REANNZ (2020), REANNZ Annual Report 2020, Year ended 30 June 2020.

[4] REANNZ (2021), REANNZ Annual Report 2021, Year ended 30 June 2021.

69. A 2018 Sapere report[5] on the REANNZ situation concluded that the organisation was facing a challenging financial future. The main sources of challenge included high costs for service provision, particularly over long-haul international routes, low numbers and density of researchers across the local network, under-utilised capacity and increasingly competitive service offerings from commercial internet service providers. That analysis remains valid now.

REANNZ required significant Crown funding to support its 2023/24 network refresh

70. As a Schedule 4A Crown Company and as set out in its constitution, REANNZ is required to build up cash reserves to become self-sustaining and to fund the network upgrades, with the next upgrade due in 2023/24. However, REANNZ's cash reserves were forecast to be exhausted by 2023/24. In Budget '21, REANNZ successfully bid for an additional $23.95 million in funding. This included $10 million in capital expenditure for the network upgrade and a gradual increase in Crown funding to $6.5 million a year in 2023/24.

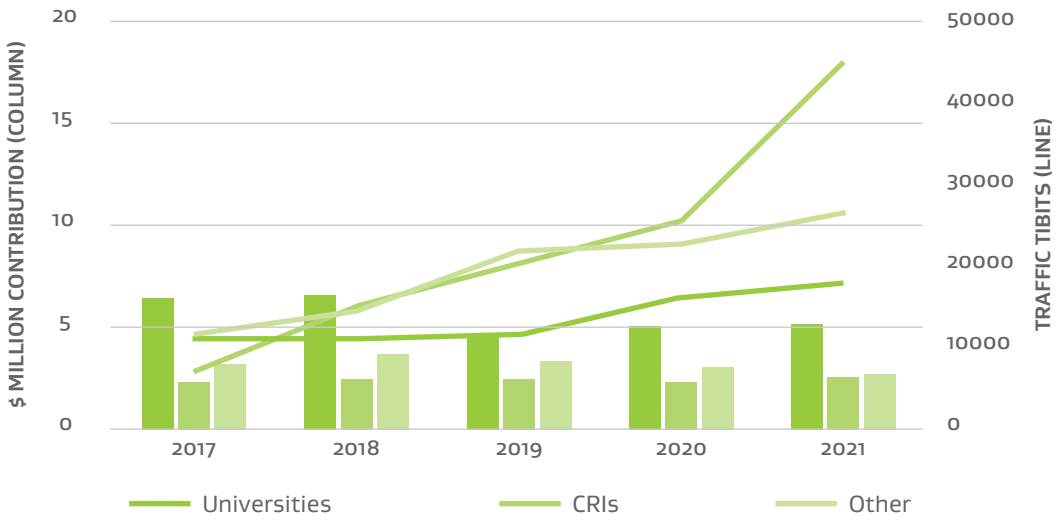


Figure 3: REANNZ Funding Contribution and Traffic, by Sector

[5] Moore, D., Tran, L., Uddstrom, M., & Yarrall, D. (2018). New Zealand's high speed research network: At a critical juncture. Sapere.

NESI

NeSI also operates a club funding model, with significant contribution from collaborators and the Crown

71. NeSI also runs on a club funding model. However, their institutional and funding structures are significantly more complex than those of REANNZ. Access to the HPC can be granted via three different pathways: the collaborator, the subscriber and merit pathways (see Figure 2).

72. The collaborators are the founding members of the club alongside MBIE, with each paying for the purchase and operations of the HPC, user support and training, and capability building, as well as reinvesting to fund future upgrades as required. Over the last four financial years, the collaborators have contributed an average of around 43.5 per cent of NeSI's total revenue (see Table 2).

73. Non-collaborators (other research and educational institutions who are beyond the primary collaborators) may access NeSI's services via the fee for service subscription (typically year-long contracts) or the fully subsided merit pathways.

Table 2: NeSI Financials 2018 to 2021[6]

	2018 ($M)	2019 ($M)	2020 ($M)	2021 ($M)
Revenue				
Crown contribution	6.4	7.9	7.5	7.4
Collaborator contribution	4.0	5.4	4.9	5.0
Collaborator reinvestment	3.7			
Total contribution to NeSI	**14.1**	**13.3**	**12.4**	**12.4**
Cash funding from the Crown	*7.2*	*7.2*	*7.2*	*7.2*
Crown surplus/(deficit), closing balance	*3.4*	*2.6*	*2.3*	*2.1*
% user funding	**54.6%**	**40.6%**	**38.7%**	**40.3%**
Expenditure				
People	5.8	7.2	6.9	7.8
Other operating expenditure	0.8	0.7	0.4	0.5
Platform operating expenditure	2.1	2.8	2.3	2.5
Platform depriciation	1.7	2.6	2.7	1.6
wCapital expendiure	3.7			
Total expenditure	**14.1**	**13.3**	**12.4**	**12.4**
Sector				
Revenue	0.2	0.4	0.6	0.5
Expenditure (Excl. commitments)		0.2	0.2	0.6
Sector fund, net for the period	**0.2**	**0.1**	**0.4**	**(0.1)**
Sector fund, closing balance, running	**0.3**	**0.4**	**0.8**	**0.7**

[6] Data provided by NeSI.

74. NeSI revenue from the Crown comes via the SSIF, making up an average of 54 per cent of NeSI's yearly revenue. Under NeSI's current funding agreement with MBIE, $2.0 million of this is designated as the University of Auckland's hosting fee, which funds the NeSI Directorate, being the national head office, for salaries and overhead costs to support the national leadership and management of NeSI alongside science sector engagement. The remaining $5.2 million of Government funding is used to support the NeSI collaboration via co-funding at a 1:1 rate with the collaborating organisations (i.e. could be used for the purchase of software or hardware, salaries and overheads across all other teams, and depreciation expenses).

75. MBIE recently agreed to increase NeSI's baseline SSIF funding by $0.64 million in 2022/23, to address an expected deficit caused by increased activity levels, staffing and planned recruitment over the two-year extension of the NeSI contract.

NeSI's funding model generates tension for the collaborating institutions

76. This funding model has generated significant tension within the club of investing organisations, with organisations both inside and outside the club perceiving differing future HPC needs and facing differing effective costs. Reconciling those differences within the club and across the interests of the wider eResearch sector is an ongoing and substantial challenge, as well as a governance distraction.

International Trends

ERESEARCH TRENDS

Rapid technology development is transforming research

77. The continued development and improvement of compute technologies, such as graphics processing units (GPU), has resulted in advanced techniques that are transforming both research and industry, including AI, ML and deep learning (DL). These technologies are allowing research to collect, process and analyse large quantities of data at unprecedented rates and have already had significant impacts in the fields of healthcare, neuroscience and agriculture.

78. Due to the power of AI, ML and DL, and the range of fields in which they can be applied, their popularity is growing, with uptake of these technologies in non-traditional fields such as education and the humanities/social sciences.

79. Further advancements in compute infrastructure, including quantum, exascale and edge computing, will continue to increase the power and speed of AI, ML and DL, allowing researchers to address increasingly complex questions that are currently beyond reach.

80. In this section, we consider eResearch infrastructure developments in countries comparable to New Zealand where useful lessons and comparisons may be found.

Researcher needs are used to inform investment

81. Advancements in relevant technologies have given rise to several common themes in terms of the evolving needs of eResearchers for infrastructure across various jurisdictions. These were highlighted in the Canadian Digital Research Alliance Needs Assessment[7] (see Box 3).

JURISDICTIONAL TRENDS

There has been a shift towards a more integrated eResearch approach with a longer-term view

82. Internationally, many countries have moved to an integrative approach to service provision at both the capability and institution levels. This includes taking a roadmap-based approach to investments, typically covering three to five years. In most cases, this shift has been associated with increased funding levels allowing highly subsidised eResearch infrastructure and services, including the provision of centralised commercial Cloud offerings. There has also been a convergence with respect to addressing data sovereignty issues.

83. Differences still arise in the ways countries integrate their service provision, each with varying levels of success. Relevant examples of these jurisdictional trends include Australia, Canada, Singapore and Sweden.

Australia has invested significantly in research data management

84. Australia's two 'Tier 1 compute facilities' are the National Computational Infrastructure and the Pawsey Supercomputing Research Centre. Both focus on compute, data visualisation and storage. While Pawsey hosts HPC, it also manages Nimbus, a national research-specific Cloud service.

85. The Australian Research Data Commons (ARDC) provides access to nationally significant, data-intensive digital research infrastructure and platforms (i.e. the National eResearch Collaboration Tools and Resources [Nectar] research Cloud), and has a sustained focus on improving data management, along with an integration of wider capabilities (see Box 4).

[7] Digital Research Alliance of Canada. (2021c). Researcher Needs Assessment: summary of what we heard. Government of Canada.

BOX 3: DIGITAL RESEARCH ALLIANCE OF CANADA (DRAC): RESEARCHER NEEDS ASSESSMENT

In 2021, the newly formed DRAC engaged in extensive consultation with the eResearch sector to better understand the diversity of needs and challenges in digital research infrastructure. Their findings were categorised into four areas:

Awareness and Accessibility

There is a need for a national catalogue of digital research infrastructure with the funding, administrative and language barriers reduced to increase the accessibility of services.

Governance and Policy

There is a need to promote collaboration with indigenous peoples, industry and the private sector to drive innovation, while maintaining data sovereignty, security and intellectual property. Stronger policies that facilitate the discoverability and interoperability of national data sets were identified as ways to improve collaboration.

Operations

There is a need to improve access to remote-computing resources, low-cost data storage, curation, analysis and burst computing via either commercial or national platforms.

New national standards, data repositories, funding streams and support staff are required to ensure long-term storage and management of data, to maximise its value. Coupled with this is the need to improve cybersecurity to allow secure storage, transfer and access to sensitive data.

Support

The education and retention of support staff (data curators, analysts, archivists, metadata experts, software developers and system engineers) at the national and institutional levels is essential to the long-term success of the eResearch strategy.

86. The Australian Government sets the funding and direction for infrastructure investments in five-year cycles via the roadmap for infrastructure investment generated by the National Collaborative Research Infrastructure Strategy (NCRIS) programme. The Research Infrastructure Investment Plan sets the two-year funding cycles that align to the NCRIS Roadmap.

87. In 2018, Australia's national science agency, the Commonwealth Scientific and Industrial Research Organisation (CSIRO), launched a 'Cloud Right' program that provides researchers with managed accounts for Google, Amazon Web Services and Microsoft Azure. The programme has a skills and investment partnership with the commercial providers and a sophisticated automatic-billing platform to support access.

88. Australia's Academic and Research Network (AARNet), which is a not-for-profit company owned by 38 Australian universities and CSIRO, is Australia's NREN. It provides Australian research and education communities with the required network backbone and federated identity services, as well as an expanding range of cybersecurity, data storage (i.e. CloudStor) and sharing platforms.

89. This network is primarily funded by membership subscriptions for access to the network and additional services. Currently, it has a wide customer base, including the shareholding organisations, most of the publicly funded research agencies (e.g. Australian Nuclear Science and Technology Organisation), several state government agencies, hospitals, schools, TAFEs (Technical and Further Education) and GLAMs (galleries, libraries, archives and museums).

BOX 4: AUSTRALIAN RESEARCH DATA COMMONS (ARDC)

Australia's eResearch infrastructure is more mature than New Zealand's, following an investment of around $4 billion over 12 years (2017/18–2028/29) to support nationally significant research infrastructure. A total of $911 million has been invested in digital data and eResearch platforms, including $160 million for the ARDC,[8] which was established in July 2018 as a result of recommendations made under the Australian Government's NCRIS 2016 Roadmap.[9]

The formation of the ARDC involved bringing together the Australian National Data Service, National eResearch Collaboration Tools and Resources (Nectar) and Research Data Services into a company limited by guarantee.

The ARDC provides access to nationally significant, data-intensive digital research infrastructure, platforms, skills and collections of high-quality data. As a national research infrastructure provider, the ARDC also facilitates partnerships to develop a coherent research environment that enables researchers to find, access, contribute to and effectively use services to maximise research quality and impact.

Canada has a strong focus on national integration and coordination

90. From 2012 until 2022, Canada's eResearch ecosystem was led by Compute Canada, a national federation of 38 universities and four regional HPC-centric facilities. It provided governance and management of the advanced computing platforms to ensure the delivery of a world-class service. However, its effectiveness was constrained by conflicts between its institutional partners and government funders.

91. To address these issues, Canada created a new organisation, the Digital Research Alliance of Canada (DRAC), which officially took over the management of HPC facilities on 1 April 2022. During the transition period, Canada put into place policies to gradually consolidate HPC facilities and coordinate the funding of eResearch infrastructure. The DRAC also identified that less than 40 percent of HPC demand was being satisfied through national offerings and it is currently looking to develop a national strategy for the use of public and commercial Cloud computing.[10]

92. The Canadian Network for the Advancement of Research, Industry and Education (CANARIE) operates Canada's NREN with support from 13 provincial and territorial partners. The resulting network connects over 750 Canadian universities, colleges, school boards, research hospitals, government agencies, research institutes, not-for-profit organisations, private sector organisation, business incubators and accelerators. It hosts further capability and service provision in cybersecurity, identity and access management, research software, research data management, and the Digital Accelerator for Innovation and Research Cloud program that connects Canadian start-ups and small to medium enterprises to each other and to free Cloud platforms.

93. This network can provide these services to the research and education sectors at heavily subsided rates, as most of its funding is provided by the Canadian Government. In 2020–2021, $30.2 million of its $31 million funding was provided by Innovation, Science and Economic Development Canada.

[8] Australian Department of Education, Skills and Employment (2018). National Collaborative Research Infrastructure Strategy: 2018 Guidelines, Australian Government, pg. 27.

[9] Australian Department of Education, Skills and Employment (2016). National Research Infrastructure Roadmap Research Infrastructure Investment Plan. Australian Government, pg. 22.

[10] Digital Research Alliance of Canada. (2021). Strategic Plan 2022–2025. Government of Canada.

Singapore's approach has involved partnering closely with the commercial sector

94. The National Supercomputing Centre (NSCC), Singapore's primary eInfrastructure service provider for researchers and industry, is closely aligned with that Government's innovation agenda and funding. NSCC has mainly focused on creating competitive compute and storage infrastructure, as well as creating robust international connections through its NREN, Singapore Advanced Research and Education Network (SingAREN). Recently, NSCC has begun to focus on the provision of a national research Cloud platform, leveraging the aggregated demand across NSCC's stakeholders and partner organisations to create an affordable Cloud-based resource.

95. Commercial Cloud offerings are already available to Singaporean researchers, start-ups and industry in AI/ML fields via AI Singapore's partnership with Google Cloud. By partnering with the commercial sector, Singapore's AI/ML researchers and AI apprenticeship programme can utilise cutting-edge technologies to build capability and tools to address real-world problems.

96. The Government-funded Agency for Science, Technology and Research also has capability in eResearch via the Institute of High Performance Computing, which acts as a focal point for meeting the system's computational modelling, simulation and AI needs.

97. As noted above, SingAREN is Singapore's NREN provider, with further capability and service offerings in federated identity management, eduGAIN and eduroam, and database mirroring. These services are provided to its institutional members, including universities, polytechnics, schools, the Government, institutes and industries, for an annual membership fee.

98. However, SingAREN's network upgrades are funded by the Singapore Government via the National Research Foundation of Singapore National Research Infrastructure framework. In 2013, SingAREN received five years' funding for the Lightwave Internet exchange network. Subsequently, it received a similar funding package in 2020 to upgrade the Lightwave Internet exchange network to boost Singapore's network speed and quality.

Sweden's integration efforts have occurred at three levels: national, Nordic and European Union

99. Sweden benefits from access to European Union-funded capabilities, including PRACE (Partnership for Advanced Computing in Europe), EGI Federation and EUDAT (European Data Infrastructure), EOSC (European Open Science Cloud) and EuroHPC (the European High Performance Computing Joint Undertaking), simplifying service delivery as many requirements are met by the EU capabilities.

100. Sweden connects to these international resources via its NREN, the Swedish University Computer Network at the national level. This network is primarily financed by member organisations and grant funding from the Swedish Ministry of Education via the Swedish Research Council. It also provides a range of add-on services, which include cybersecurity, identification and Cloud-based services for data storage and management.

101. The network is also a member of NORDUnet, an international collaboration between the NREN in the Nordic countries and GÉANT, a pan-European network of 39 NRENs. This provides the Swedish education and research sectors with greater access to network infrastructure, services and support.

102. At the national level, the Swedish National Infrastructure for Computing coordinates and allocates access to six HPC centres across the country. While the six members receive Government funding to maintain compute requirements, the investment process is not coordinated. The eResearch community has started to discuss strengthening the organisation to encourage complementarity in investment and to promote consolidation.

103. This organisation is also a member of the Nordic eInfrastructure Collaboration (NeIC), a federation of national eResearch organisations from the Nordic countries. Each organisation within it operates independently, with members choosing whether to participate in joint initiatives, and with successful initiatives driven locally and subsequently opened up to the federation.

104. The Swedish eScience Research Centre was established to develop and provide state-of-the-art eResearch tools to improve research outcomes at the ecosystem level. However, it currently acts as a networking and brokering resource.

Key Challenges

105. This report is being undertaken as the Crown considers options for future investment in New Zealand. Several reports have been produced over recent years, exploring issues and opportunities for eResearch in New Zealand, including in the form of officials' advice to Ministers. This attention has generally been sparked by funding shortfalls, particularly with respect to the REANNZ connectivity services, as well as by instability in membership, particularly of the NeSI collaboration.

106. Underneath those immediate sources of pressure is a range of other factors that can undermine the stability of the national eResearch platform. This section explores those factors in depth.

SCALE

New Zealand faces an inherent challenge because of economies of scale, exacerbated by fragmentation in eResearch infrastructure arrangements

107. Scale matters. Compared with other countries against which we benchmark ourselves, our economy is small, our per capita incomes are small, our science and research community is small, and we commit a relatively small share of our incomes, in both the public and private sectors, to science and research. We have a small number of research and educational institutions across which to distribute the cost of eResearch infrastructure. Therefore, the costs per institution are comparatively high. That price hurdle is accentuated by the smaller proportion of costs that are covered directly by the Crown, in comparison with other developed economies.

108. The impacts of these scale issues emerge in numerous ways. Compared to their counterparts in larger research systems, our researchers have access to a narrower range of specialised computers, specialist software or specialist technical support personnel to assist them. Moreover, our researchers are less likely to be interacting with a broad range of fellow researchers in ways that stimulate new ideas and new solutions to their problems.

109. This is, of course, the story of New Zealand. There are strategies that can reduce the impact of small scale and isolation. Avoiding fragmentation in our domestic infrastructure arrangements is a good place to start. Active collaborations with others, domestically and internationally, can lessen the disadvantages of small scale. Beyond that, smart strategic choices and agility in execution are always important.

INSTITUTIONAL ARRANGEMENTS

Current institutional forms create conflict and limit the accessibility of services

110. REANNZ is established as a Schedule 4A Crown Company under the Public Finance Act 1989. It is also subject to the Companies Act 1993 and certain sections of the Crown Entities Act 2014, including those relating to financial obligations and reporting. The Minister of Research, Science and Innovation appoints the Board of Directors, the members of which carry all the usual governance roles and responsibilities of company directors.

111. The institutional and governance arrangements in REANNZ appear to be sound and functional. However, as discussed below, funding arrangements are problematic with respect to generating the revenue required to ensure the company has long-term viability. This is exacerbated by its long-term commitment to fund a capacity contract with the Hawaiki cable provider. A 2018 Sapere[11] report concluded that REANNZ faced a structural deficit with very limited opportunity to boost membership income.

[11] Moore, D., Tran, L., Uddstrom, M., & Yarrall, D. (2018). New Zealand's high speed research network: At a critical juncture. Sapere.

112. NeSI is, in a literal sense, a non-entity. It has no formal corporate form but exists as a contract-based collaboration between the host, University of Auckland, and the investing institutions Manaaki Whenua, NIWA and the University of Otago. While NeSI is overseen by a Board of Directors, this is advisory only. The investing members hold decision-making rights over future investments. Personnel working within NeSI are employees of investing members.

113. In many respects, NeSI's successes have come despite its institutional form – good people with a positive attitude can make almost any structure work. But the structure is a long way from ideal for an organisation that needs to have a strong focus on strategically important outcomes and agility in execution. This is especially problematic given the environment of rapid technology developments and the imminent arrival in New Zealand of large-scale commercial Cloud operators. These will likely open up a range of new options by which to meet data management, storage, analysis and transfer needs – for both NeSI and its current partners and users.

CLUB MEMBERSHIP MODEL

Club models work well for like-minded members but fray as the members' needs diverge

114. The club membership model employed by REANNZ and NeSI can be a cost-effective and efficient way for members to gain access to facilities than none individually could afford to purchase.

115. The secret of functional clubs is that members are like-minded – that is, they are seeking similar benefits, have similar ability to pay and may find advantages in the collaboration with other club members beyond simply having access to common facilities or services.

116. However, the research community is characterised by a wide diversity of interests, needs and ability to pay. Accordingly, there is a wide disparity in the real and perceived 'value for money' in club membership, as highlighted by members leaving the arrangements. The clubs are not stable.

117. As noted earlier, only a small number of researchers are undertaking the type of research that requires large HPC and connectivity of the sort provided by REANNZ and NeSI. For them (and their funders), access to facilities of this type is essential. The large number of users that have smaller needs can have their needs met via access to commercial solutions for compute and connectivity that are both functional, at least at first glance, more cost effective and convenient for researchers. This drives the tendency to fragmentation of the clubs.

118. While the club membership is dominated by the CRIs and universities, these two groups are very different in character. That suggests a need for more 'segmentation' of system clients with more closely targeted solutions and pricing packages.

FUNDING AND PRICING

High levels of member funding affect perceptions value

119. Current annual expenditure for REANNZ and NeSI totals around $28.4 million (2021 data). Of that, $10.4 million is funded via MBIE SSIF contracts, the remainder being contributions from participating institutions, split into club fees and user charges.

120. However, the diverse interests and needs in eResearch infrastructure affect the users' perceptions of the value of membership. Those differences emerge most clearly when institutions choose to invest in compute and connectivity facilities outside the club. Both REANNZ and NeSI have taken steps to tailor their offerings to members to better meet their individual needs and to soften their value-for-money concerns. But membership, especially for NeSI, remains unstable.

NeSI is seeking additional funding to increase compute capacity, while REANNZ received additional funding via Budget 2021 to meet its current obligations and fund the network upgrade. Further funding pressures lie ahead. Commercial offerings will provide added pressure as their prices decline

121. It is likely that commercial services will put added pressure on the clubs, with growing options and better pricing emerging on both compute and connectivity facilities. The niche that REANNZ and NeSI service will likely continue to be squeezed over time. That squeeze comes from two directions.

122. Diminishing costs and improved capability of commercially offered services for connectivity and HPC will constantly nibble at the appeal of the REANNZ and NeSI offerings, especially for smaller-scale eResearcher tasks.

123. For larger-scale projects, advances in Cloud computing, together with bundled software and support offerings, make that option increasingly attractive. Cloud accessibility is likely to take a further step with the imminent arrival of several local Cloud facilities.

Transparency of full cost is needed

124. To navigate this emerging landscape of viable alternative services, transparency of full costs and value delivered will be even more important. Efficient decision-making on all fronts is challenging enough, without adding opacity in underlying costs and benefits to the mix.

125. There is a belief within parts of the research community that access to eResearch infrastructure should be provided without cost to researchers. This is the free, 'all-you-can-eat buffet' model that is thought to encourage greater participation as well as reduce complexity and other obstacles in engaging in eResearch activities.

126. A commonly suggested funding model involves simply 'top-slicing' the Crown's overall Research, Science and Innovation budget allocation and removing all cost-recovery elements relating to use of the infrastructure.

127. It is clear, but hardly surprising, that researchers respond to incentives as much as any other group. The merit access route to HPC facilities offered by NeSI free of charge, including to non-members, has certainly boosted the number of eResearchers using those facilities (see Figure 3). But as the system approaches capacity, allocation decisions must be made. Opacity of costs is unlikely to make those allocation decisions easier or more efficient.

128. Across the system, there are choices to be made by participating institutions about which projects are best suited to the NeSI and REANNZ specialist, high-performance but expensive facilities, and which are best done within their in-house facilities. Pricing mechanisms can play a critical role in aligning those decisions with what is efficient in terms of getting the requisite volume of high-quality eResearch undertaken for the lowest infrastructure costs.

129. The optimal mix of core Crown funding and cost recovery will be shaped by the nature and extent of quality and price advantage gained through central negotiation of contracts for compute and connectivity capacity, compared against the results individual collaborators could gain by negotiating their own deals. The bigger any advantage through central negotiation and provision, the more it is to the advantage of both the Crown and collaborators to incentivise researchers into using the central facilities.

DATA AND STORAGE

New Zealand's research data management maturity is at an early stage

130. Data is not just the raw material of eResearch; it is increasingly being seen as a valuable asset to be carefully husbanded for wider and future benefit. Numerous reports have argued that New Zealand is well short of best practice in its management of research data. During our engagement, we found a very mixed pattern – some pockets of excellence in data curation, but other areas of deep frustration that research data is not being adequately managed. In too many instances, it seems, data is neglected once research projects have run their course and funding stops.

131. In 2016, REANNZ, NeSI and New Zealand Genomics Limited jointly published a report under the eResearch 2020 banner entitled 'National Research Data Programme: The Case for Research

[12] REANNZ, NeSI, & NZGL. (2016). eResearch 2020: National Research Data Programme: The Case for Research Data.

Data',[12] which explored the New Zealand research data field and proposed the formation of a National Research Data Programme involving an investment of $50 million to $70 million over a 5-year period. The aim of the programme was to "transform New Zealand researchers' abilities to work with and create value from data".

132. The proposed major programme outcomes were:

- a national metadata catalogue for research data
- a comprehensive professional development programme
- operational support for data management
- development of active data bridges
- aligned national and institutional policies.

133. Professional, effective data curation is an essential element of any eResearch capability programme. Jones and Dietrich explored this subject in 2018, noting that the field "sits at a relatively exploratory stage of evolution, with new tools, approaches and norms being developed all the time".[13] Other countries have been investing significantly in this field, few with more effort and persistence than Australia with its ARDC programme of recent years (described earlier in Box 4).

Bridging the gap needs to involve working closely with Māori

134. New Zealand has a significant and potentially costly gap arising from the lack of national leadership in establishing principles, protocols and practices in data curation across the national research system. There is clearly a case for building a community of practice; that is, national leadership working with local researchers and data owners, aimed at making progress towards the types of outcomes proposed in the eResearch 2020 report by REANNZ, NeSI and New Zealand Genomics.

BOX 5: CARE DATA PRINCIPLES

New Zealand Māori academics have worked with the Global Indigenous Data Alliance to develop the following CARE data principles:

Collective benefit: Data ecosystems shall be designed and function in ways that enable Indigenous Peoples to derive benefit from the data.

Authority to control: Indigenous Peoples' rights and interests in Indigenous data must be recognised and their authority to control such data be empowered. Indigenous data governance enables Indigenous Peoples and governing bodies to determine how Indigenous Peoples, lands, territories, resources, knowledges, and geographical indicators, are represented and identified within data.

Responsibility: Those working with Indigenous data have a responsibility to share how those data are used to support Indigenous Peoples' self-determination and collective benefit. Accountability requires meaningful and openly available evidence of these efforts and the benefits accruing to Indigenous Peoples.

Ethics: Indigenous Peoples' rights and wellbeing should be the primary concern at all stages of the data life cycle and across the data ecosystem.

These principles aim to ensure the current movement toward open data and science fully engages with Indigenous Peoples' rights and interests, while complementing the existing data-centric approach represented by the FAIR Principles (Findable, Accessible, Interoperable, Reusable).

[13] Dietrich, M. & Jones, N. (2018). Understanding the eResearch Ecosystem in New Zealand. NeSI pg. 16.

BOX 6: RAKEIORA PATHFINDER PROGRAMME: MĀORI CO-GOVERNANCE OF DATA

The field of human genomics is undergoing rapid global growth, but to maximise its benefits for New Zealand, we need to be mindful of our unique cultural and genetic composition, significant health inequities and historical misuses of genomics.

Therefore, we need coherent national processes to bring together genomic data with multiple other forms of health-related information from multiple areas of research and practice, to improve healthcare research and support precision medicine. This requires co-developing a foundation for innovative and scalable national genomics research infrastructure with Māori to manage and govern data, its ownership and guardianship, including New Zealand-specific genomic databases.

Rakeiora, a 'Pathfinder' programme, is being developed to test options to acquire, protect, use and store genomic datasets for this purpose, with a particular focus on addressing the country's health inequities by developing genomic tools that put the needs and priorities of Māori at the centre.

The researchers behind Rakeiora have therefore co-developed and co-governed the programme with Māori. The programme's success has relied upon forming the relationship with Māori prior to the inception of the programme, working with an 'equal voice' in decision-making, developing a mutual understanding of each other's perspectives and approach, and working at each other's pace.

The successful establishment of Rakeiora may help in developing a framework more generally for researchers to work with Māori in a way that best supports Māori interests in wellbeing and the protection of data.

135. An important reason for a national leadership role in data curation is the pressing need to progress issues relating to Māori data sovereignty. The CARE data principles[14] to support indigenous rights and interests (see Box 5) provide a foundation for addressing this issue. As well, there is scope to draw on previous work conducted in this area, including that of Te Mana Raraunga (Māori Data Sovereignty Network), the Pan-CRI Māori Data Sovereignty working group and the Rakeiora Pathfinder Programme model of developing co-governance over Māori data (see Box 6).

Institutions have signalled the need for data storage at a national level

136. The New Zealand research system operates on a full-cost funding basis, whereby the cost of data storage and management is included in research funding proposals. Institutions are expected to procure and manage their own data storage and management infrastructures. However, research proposals receive funding for only a limited period, leading to 'funding cliffs' for maintaining and storing data.

137. Many institutions have signalled that they are experiencing data storage constraints due to the exponential growth in the production of data. However, the cost of data storage is not reducing at a comparative rate. This is compounded by the lack of long-term funding for data storage.

138. Currently, a variety of publicly and privately funded collections (systematically collected physical specimens) and databases (repositories of subject-specific information) are held by research institutions such as museums, universities, and local and central Government. While many of these are physical collections, the database components fall into the eResearch sphere of interest. Examples of national-scale databases include the following:

[14] Carroll, S.R., Garba, I., Figueroa-Rodríguez, O.L., Holbrook, J., Lovett, R., Materechera, S., Parsons, M., Raseroka, K., Rodriguez-Lonebear, D., Rowe, R., Sara, R., Walker, J.D., Anderson, J. and Hudson, M., (2020). The CARE Principles for Indigenous Data Governance. Data Science Journal, 19(1), pg.43.

- The database components of the Nationally Significant Collections and Databases SSIF platform: These lack national coordination on data management and storage, leading to inconsistent practices and hindering their accessibility. Static funding has left the hosts struggling to maintain the databases adequately.
- NeSI's persistent and nobackup/scratch data storage service (hosted at Greta Point): This is in high demand, with both aspects reaching capacity despite measures to regularly audit and clean the data and manage the amount.
- The Aotearoa Genomic Data Repository: This is an online resource, funded by Genomics Aotearoa, for the storage and sharing of genomic data. It was jointly developed with NeSI to provide a secure place to store data within a Māori values context.

139. Additional funding for national data repositories such as the above could alleviate the data burden, encourage data standards to be met and protect ongoing access to valuable data sets. However, this is an expensive undertaking and calls for carefully structured frameworks, principles and criteria to guide decisions on what is stored, by whom and for how long.

140. In Australia, the ARDC has invested heavily in developing frameworks for managing data retention policies. These identify the criteria by which decisions can be made on who (i.e. institutions vs state/national governments) should have responsibility and what data (i.e. working data, scale and complexity of data sets, local or national significance) should be held, with what access conditions. New Zealand could readily gain from the Australian work and there would be considerable benefit in seeking trans-Tasman compatibility in metadata standards and beyond.

SKILLS AND TRAINING

The uptake of eResearch is largely a skills gaps challenge

141. Leading-edge research draws from researchers' creativity and skills. Increasingly, the core skills required involve eResearch capabilities. Science, Technology, Engineering and Mathematics (STEM) disciplines have taken up eResearch techniques fairly readily. Generally, the researchers in those fields start with some familiarity and confidence in the statistical, mathematical, modelling and compute skills required. This is less likely in other disciplines, such as the humanities and some social sciences in which eResearch techniques have not been commonly applied. That is changing.

142. As appreciation of the potential for eResearch techniques to enhance research impact grows, so the need for skills development grows. Leading researchers who have deep subject-matter expertise cannot be assumed to have the knowledge and experience to identify and exploit eResearch opportunities in their fields. Recognising this gap, both REANNZ and NeSI have offered programmes to build researcher skills and provide technical support in the use of their hardware. Universities are also increasing their offerings in eResearch-related subjects, to build student skill levels in these fields. Moreover, commercial Cloud providers offer training courses, generally online, for those wishing to build capability in, and experience with, the latest techniques and technology.

eResearch training needs to be tailored and offered according to the level of expertise

143. There is a good deal of segmentation of learners in these fields. At the top end of big users of HPC and connectivity technology, research is conducted in large teams. Those are likely to include personnel with either high-level eResearch skills or the ability to recruit those skills as necessary. They are unlikely to be heavy users of the course offerings from NeSI; indeed, they are more likely to provide experts to assist in teaching NeSI programmes.

144. At the lower end, undergraduates and perhaps Master's-level students are significant users of the carpentries and other offerings from NeSI. These are largely at the level of basic principles and practical techniques – necessary starting points for a researcher – but not delivering the high-level skills required for New Zealand researchers to engage in very large-scale data science projects, either here or in international collaborations.

145. These lower-grade NeSI programmes are necessary stepping stones for researchers making their way into higher-level eResearch projects. They are also at levels where we should expect our universities to be delivering more in their standard courses, at least at the generic level.

RAPID TECHNOLOGICAL ADVANCEMENT

Commercial operators have rapidly increased their New Zealand footprint

146. The future delivery of eResearch infrastructure will need to be adaptable in the face of technological advances that have the potential to squeeze the specialist niches occupied by offerings from REANNZ and NeSI. Commercial Cloud operators are investing massively in new compute technology. Several of the major firms will soon have large-scale data storage and compute facilities operating in New Zealand. The local presence of large-scale Cloud operations may well be accompanied by further investment in undersea cables to support their operations, with the potential for further competition with the REANNZ offering.

Determining the right levels of infrastructure and support for diverse users will be a balancing act

147. At the lower end of the research spectrum, university students are, in aggregate, significant users of NeSI and REANNZ facilities. In a funding-constrained environment, there may be strong incentives to undertake this work on low-cost, off-the-shelf desktop computers and commercially provided networks.

148. For mid-sized projects run by CRIs and some university-based teams, specialist infrastructure may be required and central provision is likely to be more efficient than standing up multiple HPC hubs across our research institutions. At this scale, university researchers who could make efficient use of the central infrastructure may prefer to maintain close control of their compute work and by acquiring their own dedicated machine. A combination of funding/cost-recovery strategies and managerial policies will be required to shape decisions on what compute work goes where.

149. Finding an efficient mix of central versus local infrastructure is not a straightforward process and the efficient balance is never going to be static. For this reason, transparency of costs and pricing (i.e. being clear about the full cost of different options, regardless of who is paying) is essential, to enable the making of well-informed investment decisions over time.

Commercial Cloud offerings allow access to cutting-edge technology

150. As with any field in which technology is advancing rapidly, investments in eResearch infrastructure face very rapid depreciation rates, as performance advances can render equipment either obsolete or heavily devalued within a year or two. Managing investments and balance sheets in such an environment is always going to be difficult. One option is to leave that task to others and simply rent the capability and capacities required – the Cloud option.

151. At present, Cloud offerings seem to be relatively expensive. Accurate comparisons need to be made on a like-for-like basis, including elements such as cybersecurity protections, provision of high-resilience facilities, access to back-up and support, and the flexibility to quickly scale capacity up or down. On some occasions, the marginal costs associated with adding capacity to an existing local facility will be an appropriate basis for comparison. Our sense is that too often, those opting for local compute capacity are undervaluing those associated services and protections provided in Cloud offerings and under-estimating the full costs of their local solutions.

152. As has been seen in other countries grappling with these issues, Cloud offerings are likely to form a part (probably growing) of the future HPC offering for eResearch purposes. That option will require careful exploration and negotiation, drawing on a thorough understanding of researchers' needs, provider offerings and viable alternatives, including cross-border collaborations.

INTERNATIONAL COLLABORATION

Being an effective international collaborator should be a factor in New Zealand's eResearch investment decisions

153. Internationally, there is a shift towards open science strategies, including global open science Clouds and Commons. For example, the European Open Science Cloud was established as a federated environment for hosting and processing research data at a pan-European scale.

154. Wider international collaborations are a continuing theme as the cost of cutting-edge technologies puts them beyond the reach of individual countries. Therefore, there is an increased driver for international collaborations to allow access to physical and Cloud-based infrastructure. Some examples are the following:

- Australia's National Computational Infrastructure and Singapore's NSCC have signed a memorandum of understanding that includes exploring collaboration in HPC infrastructure and capability development (e.g. exascale systems, green data centre technologies, greater research network connectivity and secure data transfer).
- In the first arrangement of its kind outside Japan, Singapore's NSCC has gained regular access to Japan's Research Organization for Information Science and Technology Fugaku system, the world's most powerful supercomputer.
- In 2020 40 European NRENs (represented by GÉANT) conducted a pan-European tender for commercial Cloud services, resulting in range of Infrastructure as a Service (IaaS) offerings from multiple commercial Cloud providers that are tailor-made and priced for academia.

155. New Zealand could leverage existing agreements/arrangements to access technologies that are out of scope and scale for us to invest in on our own. Current international agreements with a research infrastructure component include:

- 2019 New Zealand-Singapore Arrangement relating to Science Technology and Innovation Cooperation[15]
- 2017 Australia–New Zealand Science, Research and Innovation Cooperation Agreement (ANZSRICA)[16]
- 2010 Agreement between the Government of New Zealand and the Government of the United States of America on Science and Technology Cooperation Contributing to Domestic and External Security Capabilities.[17]

156. In essence, New Zealand has laid significant groundwork for international collaboration, but needs to take more steps to realise the full benefits.

[15] Arrangement relating to Science Technology and Innovation Cooperation, New Zealand-Singapore, (2019), Paragraph 5 Statute 6-8, pg. 6-8.

[16] New Zealand Science, Research and Innovation Cooperation Agreement, Australia-New Zealand (2017), Article 5 (4).

[17] Agreement between the Government of New Zealand and the Government of the United States of America on Science and Technology Cooperation Contributing to Domestic and External Security Capabilities, United States-New Zealand (2010), Article 2.

Future State

The case for Crown intervention needs to consider the rapidly improving commercial offerings

157. The rationale for central provision of eResearch facilities is to support and grow a strong national capability in eResearch, in its many forms, for the benefit of all our people. The benefits gained from this central provision must be greater than those we could reasonably expect without centralised facilities. That is not an easy test and passing it becomes harder as commercial offerings of both compute and connectivity capability become cheaper, more capable and more convenient, for both large- and small-scale projects.

158. Being clear about the existing gaps that could most usefully be filled by the provision of national facilities is a key strategic task for the research community and the Crown's funding mechanisms.

159. There are complex dynamics at play in this field. Technology, both hardware and software, is advancing rapidly with major impacts on capacity, performance and unit price. Data flows are growing exponentially. It is entirely possible that expanded national compute capacity could remain fully utilized even as the overall system fragments with individual research entities adding to their local facilities. The critical resource that will be hardest to acquire is high quality skills and experience.

Emphasis should be placed on growing the eResearch community

160. The general case for growing the New Zealand cohort of highly skilled eResearchers is reasonably clear-cut. eResearch capability and techniques can be applied across disciplines, in new ways, to provide new and valuable insights in complex ecosystems. It is a core capability in modern research and an essential element if New Zealand is to hold a respected place near the top tables of international research. International collaborations will increasingly be formed around teams with strong eResearch capabilities.

161. With that in mind, it is important to attract talent into eResearch fields and provide pathways for skill development and the effective application of those skills in high-impact research. The availability of eResearch hardware, easy access to it and strong technical support and advisory services are all critical components.

162. The case for high-capacity connectivity facilities, NREN and associated functionality is comparatively clear-cut. The strategic issues on that front relate to how best to provide access to the necessary capacity in a manner that meets the diverse needs of researchers, is readily scalable as required, and is cost efficient for users and the Crown.

CHARACTERISTICS OF A DESIRABLE FUTURE STATE FOR NEW ZEALAND'S ERESEARCH

Ambition is required to grow eResearch as a core capability of New Zealand's Research, Science and Innovation sector

163. A refreshed and relaunched eResearch platform should come with high ambition for New Zealand's eResearch future. It should encompass growing numbers of graduates from our universities with eResearch skills for application in both public and private sector employment, both research oriented and commercial. We should also seek greater engagement in international collaborations on major research projects. If they are to function at the highest international levels on leading-edge projects, our eResearchers, across multiple disciplines, need opportunities to develop high-level skills and experience, both in their particular fields and in relevant data sciences. Above all, our eResearch system should aim for research excellence and high impact.

164. Ambition in this field has a price tag. Inevitably, what we can aim for in eResearch will be strongly correlated with available funding. It is currently clear that our national eResearch capability and maturity are at uncomfortably low levels, with only a few institutional exceptions. Without commitment, strategy and investment, that will remain the case.

eResearch platforms are researcher-centric, easy to access and use

165. Our eResearch functions should be convenient for researchers (and their employing institutions) to engage with, making it easy for them to utilise the eResearch infrastructure to undertake their work, and to do so efficiently with respect to their own time and usage of the hardware.

166. Clarity around the criteria applied for prioritisation and scheduling of access to the infrastructure is clearly important, as is ready access to quality technical support.

167. Ease of access to the eResearch infrastructure must also extend to Māori researchers, with diligent efforts to uncover and remove barriers to entry. Many of the barriers will not be at the point of access to the eResearch infrastructure, but much further away in the development of the underlying skills required to engage in eResearch activities. That suggests the required strategies will involve working with the universities and other tertiary education providers to support teaching programmes aimed at ensuring students, including Māori, are equipped to engage in eResearch programmes as they attain higher levels of education. NeSI has an initiative in this area underway. This will require ongoing support.

Enduring institutional governance provides a focal point for the sector and tailors offerings to user needs

168. The future institutional structure requires a solid base to attract and hold Crown and research sector investment. It will need to facilitate both effective governance and tight stakeholder engagement, including partnership with, and inclusiveness of, our Māori researchers and wider communities.

169. This is a complex field, given widely diverse stakeholder interests, directions and maturity levels, rapidly evolving technology, and rapidly evolving service offerings from commercial providers. Those elements of complexity are compounded by the usual challenges of negotiating Government fiscal priorities.

170. While there are diverse interests across stakeholders, what is important is how effectively those diverse interests can be catered for in terms of smartly tailored service and cost-recovery structures that deliver value for money for members in their various circumstances and needs. Without success on that front, the eResearch system will inevitably face disintegration as members opt to go it alone to meet their compute and connectivity needs. There may be choices in organisational form to assist with bringing members to the table and encouraging them to remain as active participants.

171. Given these various challenges, there is a premium on strong governance with tight strategic focus, energetic pursuit of strategic aims, and careful establishment and monitoring of performance metrics that measure progress towards meeting these strategic objectives.

The financial model must be stable, plan long term, contain the right incentives and encourage efficient resource use

172. An ambitious development strategy for the eResearch sector requires a long-term, sustainable financial model. The combination and structure of Crown and user-pays revenue should be designed to optimise the use of available facilities, drawing in larger and more sophisticated eResearch projects.

173. Given the diversity of needs and interests across the system, a more prominent place, at the institutional level, for fees related to system usage is likely to be more effective and sustainable than the current club subscription model. Clearly, the degree of Crown subsidy in such arrangements is highly influential in shaping demand for central services.

Costs are transparent to stakeholders making investment choices

174. While many within the system would prefer a free, 'all-you-can-eat buffet' method for accessing eResearch infrastructure, that is unlikely to serve the eResearch community well in the long term. If, as expected, we face a future with greater availability of Cloud-based services, we expect to see more user-pays practices across the system.

175. Rationing of access to centrally provided infrastructure is as inevitable as the rationing of funding for eResearch. Obscurity or invisibility of the real costs of providing services can only impede our ability to make good choices in the allocation of access to infrastructure.

176. Deciding exactly where the costs should fall (i.e. who pays) is not the same as ensuring those costs are transparent. However, those who need to make key choices on investments in either hardware or research funding should be clear about the relative resource costs of the options before them.

The middle tier of researchers is supported to grow

177. The case for the central provision of eResearch infrastructure rests substantially on meeting needs that cannot be met (or can only be met at high expense) commercially. That points towards prime users being those engaged in larger and more sophisticated projects. Only a very small group of New Zealand eResearch teams fit that category. They are already heavy users of the existing infrastructure.

178. Across our CRIs and universities sits a much larger group of 'medium-scale' researchers/users. Generally, the eResearch maturity level of these institutions is quite low. To maximise our future eResearch impact, this group is probably the most fertile territory for growth through further support and development – those who could be most productively pulled into the fold through access to high-impact training courses, technical assistance and increased familiarity with the use of the available infrastructure, including research software.

Research data management practices are supported by a central function

179. Given the very uneven standards of research data curation across our research community, there is a strong case for providing resource at the centre, leading a drive towards improved performance in this field. This could build on existing pockets of good practice already present in the system, as well as drawing on international models, to establish core principles, protocols and practices across the New Zealand research community, including with Māori. While the research data remains with (and the responsibility of) its originators and the relevant domains, it is likely that resourcing from the centre will be required to make progress towards the relevant CARE and FAIR concepts.

Central eResearch service provision is future focused and responsive to evolving commercial offerings

180. Technology and research needs can evolve rapidly. An effective eResearch platform will need to be agile in responding to both.

181. It is likely that agility, in this environment, will entail central provision of some form of 'portal' through which eResearchers can access a menu of compute options, from current-level Central Processing Unit (CPU) and GPU capacity through to a range of centrally negotiated Cloud offerings. Usage will migrate to whichever mode proves most efficient and cost effective over time and investment in, or contracting of, compute capacity should follow that evolving usage pattern.

182. Currently, the 'connect' component of the eResearch ecosystem is mainly addressed through a mix of leased and owned local networks and the subsea Hawaiki cable. Some gaps remain in the local network, with some customers, current and potential, underserved in terms of capacity. Issues for the future relate to the potential impact of remote sensing technologies and uptake of the Internet of Things. Both suggest a demand for connect options that go beyond the existing network.

183. Negotiating future investment decisions in this environment requires clarity of strategy, hard-nosed financial and business case analysis, and equally hard-nosed governance calls.

184. A considerable skill and experience base lies within the existing structures and could support pursuing future ambitions. Protecting these existing strengths will be an important consideration in the planning and rollout of future institutional reforms.

Future Options

185. The pathway ahead has options that largely hinge on how much ambition (and funding) we wish to inject into the eResearch platform.

ASSUMPTIONS AND JUDGEMENTS

186. The key assumptions and judgements underlying the options outlined in the next section are as follows:

- There is a valuable place for nationally provided eResearch infrastructure of the type currently delivered by REANNZ and NeSI. The alternative would be to leave research entities to develop their own arrangements according to need – either going their own way or collaborating with others in the system, where they see mutual benefit in doing so.
- The go-it-alone option would likely result in higher costs, duplication of facilities and support services, and unhelpful fragmentation of our small eResearch capacity.
- To be successful, centralised HPC capacity must include very easy accessibility, quality technical support and high agility, to meet the specific needs of the research community. It must be a more attractive proposition, on performance and price, than in-house options.
- Given the centrality of eResearch in modern research capabilities, the cross-cutting nature of the associated capabilities and capacities, and the relative immaturity of eResearch capabilities within our research system, there is a place for a platform within the research system that aims to foster the development of a vibrant eResearch community. This platform would sponsor and facilitate the development of relevant skills and experience, sponsor and support domestic and international collaborations, and act as a conduit to a suite of connectivity and compute options to support the eResearch community.
- While the institutional form and governance arrangements for REANNZ are fit for purpose, those for NeSI fall well short of the task. NeSI functions despite its formal structure.
- Leading and supporting the development of a strong community of practice around research data curation, including a framework for longer-term research data storage, is a function that could fit readily alongside the provision of other eResearch infrastructure-related functions. It is also a missing link in the current architecture.
- Research software and associated engineering skills are an essential component of a high-functioning eResearch platform, integral to supporting researcher use of the platform.
- New technologies are allowing massive growth in remote sensing and other real-time data collections that provide exciting opportunities for the new and expanded application of eResearch techniques, as well as added challenges in data curation, storage and connectivity services.
- Neither REANNZ nor NeSI currently have a viable funding base, meaning the status quo is not sustainable.
- A future financial model is likely to involve both increased Crown core funding and some form of user charges, probably levied at the institution level.
- The user-charges and Crown-funding regime should be designed to provide effective incentives that steer (without compelling) researchers towards the use of the central facilities. Such an incentive structure could also be designed to encourage (but not compel) participating institutions to pool their HPC investments for wider system availability and more efficient overall utilisation of capital and support skills.

- There is likely to be benefit in centrally negotiated contracts providing access to Cloud computing, storage and other necessary infrastructure, as opposed to leaving each research institution to negotiate its own arrangements.
- We are likely headed for a future in which the central compute infrastructure takes the form of a single portal through which researchers will be able to access a menu of compute options – CPU, GPU, Cloud – from which to choose the option that best meets their particular needs.
- Future commercial Cloud computing options will be based on user-pays models and will shift the balance of funding needs towards more operating expenditure and less capital expenditure.
- Cloud computing offerings may come with bundled software and technical assistance options.
- The reform choices available will be directly related to the scale of Crown funding available, although the quality of the options offered should influence the Crown's willingness to provide funds.

187. As we develop and consider options for the delivery of future eResearch services, the key criteria are as below.

188. Does the option:

- encourage science excellence
- increase the uptake of eResearch and enhance its maturity/sophistication
- encourage science collaboration across our domestic science community
- encourage international research collaboration
- achieve available economies of scale and scope, increasing the efficiency of our capital commitment in science and research
- increase engagement between our science establishment and the private sector
- avoid displacement of the private sector as service provider.

THE OPTIONS

Option 1: Status quo+

- This offers the least disruption, least ambition, least additional spend and least additional impact.
- REANNZ is functional and remains as it is. Hard decisions remain on how to extract the highest value from its existing and future contracted capacity. More work should be done on the local network, including its reach and performance. There will likely be increased commercial competition from new international cables.
- NeSI is folded into a new organisational and governance structure – probably a new not-for-profit independent Crown entity. Options for organisational form that provide for direct shareholdings by research institutions could also be explored. A renewed governance board is charged with reviewing current HPC assets, pricing and service mix. Membership of the governance board has significant overlaps with the REANNZ board.
- The existing funding is concentrated in the new organisation, enabling it to purchase or contract services, including those currently provided in kind where required and appropriate.

- This option aims to maximise value in existing assets while maximising impact on eResearch activity. The flexi HPC model pioneered with AgResearch is developed further, and the 'portal' model for adding access to the Cloud is explored.
- The approach to training and technical assistance is reviewed, including outreach to universities on eResearch teaching and the nature of engagement with existing training providers from NIWA and the Universities of Auckland and Otago.

Option 2: Incremental expansion

- This involves some added cost for added services and some extra ambition.
- REANNZ remains as it is, but with future funding streams agreed.
- NeSI is remodelled as in Option 1, with additional resource and expectations.
- Enhanced funding is provided for skills development and encouraging a greater flow of researchers, including Māori, into the eResearch fold, working with universities and compute providers who have outreach capability.
- As in Option 1, the flexi HPC model is developed further and pathways to Cloud options are opened up.
- NeSI's successor is funded to encourage enhanced capacity across the eResearch system in research data curation and storage. This is largely a community-of-interest model, with capacity at the centre to assist, encourage and facilitate efforts to build better practices across the research community. Some funding support is available for members working to upgrade their research data curation practices.

Option 3: Integrated eResearch platform

- This option aims to substantially accelerate the volume and levels of maturity in eResearch capability across the education and research communities.
- A scaled-up Data and Digital Research Institute (DDRI) is created as a new research sector platform, probably on a not-for-profit Crown entity model (also consider part-stakeholder shareholding options) with a suitably skilled and experienced governance board.
- REANNZ and NeSI are folded in together.
- Up-graded skills development and training capacity are made available, working in conjunction with universities.
- Research data curation and storage are added, as in Option 2.
- The provision of research software and the development of associated software engineering skills are specifically recognised and funded.
- Expansion in compute capacity is provided for, through both a flexi HPC model and a 'portal' model that provides access to existing and expanded HPC capacity as usage requires, plus the development of new Cloud options as they become efficient.
- Existing funding is concentrated in the new organisation, enabling it to purchase or contract services, including those that are currently provided in kind where required and appropriate.
- Funding is increased to support ambition, including in enhanced compute options and capacity, the growth of eResearch skills and experience, and opening new collaboration options and relationships with corresponding agencies abroad, especially Australia, Singapore and the United States of America.
- Sufficient 'mass' is created to build energy, momentum and innovation across the system.
- The DDRI becomes the national champion of eResearch and its possibilities, working closely with the broader education and research community to scale up the national eResearch capability and capacity, as well as its effective contribution.
- The aim of the DDRI is to build capacity and capability within participating organisations, not to crowd out those developments from the centre.

CONSIDERATIONS TO FORM AND FUNCTION

189. Through our consultations on the draft report, a number of issues were raised with respect to the options presented and considerations relevant to choosing a preferred path. Many, understandably, seek more detail, greater elaboration of intent and rationale. This report will not present a detailed design for future structure. That will need to emerge at a later stage when the Crown and the key stakeholders have reached broad agreement on the strength of their ambitions for New Zealand's future eResearch capability and the extent to which they are prepared to support that ambition with funding commitments.

190. Key questions emerging from our consultations are discussed below:

Central vs distributed capability

191. The DDRI concept hinges on the capacity of a central, national capability to generate meaningful benefits for the collective compared to those available to individual institutions. The possible pathways to achieving such benefits are through:

- negotiating better/more efficient access to compute and connectivity services
- more effective utilization of hardware capacity than can be achieved in a more fragmented arrangement
- collective scale enabling greater capacity for provision of specialized hardware, software and technical advisory and support skills for eResearchers
- maintaining and reinforcing functionality where a single national provider is the norm (e.g. NREN)
- active leadership roles driving greater consistency in data curation, particularly through establishing frameworks at the level of metadata standards and broad principles and protocols for research data management and storage, enabling improved research value from local data
- fostering and facilitating stronger linkages for eResearch collaborations nationally
- fostering and facilitating stronger linkages for international collaborations via relationships with similar bodies abroad, including negotiating access to international eResearch facilities for New Zealand researchers
- building an informed and coherent overview of national and institutional eResearch activity and capability from which to support a national strategic dialogue on how New Zealand can best progress its eResearch capability and impact.

192. A judgement on whether, how and to what extent these possible benefits can be achieved will underpin the choices made. We don't believe that there is a viable option in which the Crown walks away from an on-going strategic and financial commitment to some form of central eResearch infrastructure, leaving future developments substantially in the hands of education and research institutions. That would simply reinforce the existing state of fragmented investment, immaturity and patchy capability across a field that is so central to future research advances.

Crown-owned, stakeholder-owned, or mixed ownership?

193. The proposed DDRI proposal is open as between an independent Crown-owned not-for-profit entity, or other ownership forms that could include whole or part shareholdings from user research and education organisations. Should a DDRI or similar become the preferred option, this is obviously a key choice to be resolved.

194. The wholly Crown-owned option is clearly simpler to put in place and operate. Given the scale of Crown funding associated with the activities envisaged to be within the DDRI scope, the Crown will have a very strong interest in how that funding is being deployed, how the governance function is performed and how the overall function supports its wider research and education strategic directions and performance.

195. Introducing part-ownership by interested stakeholders across the education and research field could help build user engagement and commitment to the facilities. However, as this report has already discussed, the existing club approach has well illustrated the diversity of interests, needs and ability to pay across the user organisations. That provides a warning for the potential challenge of drawing together a wider user-group and the potential for future instability in the shareholding group. It would undoubtedly complicate the Crown's role as strategy leader for overall education and research priorities and performance, dominant funder of activities, and underwriter of the eResearch system.

196. Proceeding down a mixed shareholder path would ideally be based on engagement from a broad range of our research and education institutions, including endorsement of the strategic significance of upgraded national eResearch capability and a commitment to support and use the facilities and collaborate in developing capability.

197. Should that prove to be a step too far for too many, a pragmatic alternative would be to base the part ownership model on a coalition of the willing – those universities and CRI's that are deeply committed to REANNZ and NeSI and prepared to collaborate on a transition to a stronger and more holistic approach to building a fit-for-purpose institutional platform for our eResearch future. This could usefully include a commitment to an active outreach programme aimed at drawing other institutions and their researchers into the fold over time.

198. Some have suggested a further option for institutional form modelled on the NZ Synchrotron Group Ltd (NSGL), a company with 12 members (10 funding members) drawn from across the New Zealand university and CRI sector. Secretarial and administrative services are provided by the Royal Society of New Zealand. NSGL is the New Zealand investor partner in the Australian Synchrotron facility and administers New Zealand use-rights to local researchers. Funding is substantially provided by the Crown.

199. While this is may be a functional model for the synchrotron function, the eResearch challenge seems to be more complex, albeit with a smaller number of large collaborating partners. Critically, member interests within NSGL appear to be much more closely aligned than we see in the eResearch field.

200. Choosing between these various institutional forms – or, indeed, other options – is a task in its own right and can only be done with a firm understanding of how the Crown and the potential partner institutions view their longer-term strategic interests in the eResearch field and what appetite exists for future funding.

REANNZ in or out?

201. As the report makes clear, REANNZ is well established in form and function, but not financially sustainable. Would it be best left outside of a DDRI platform with attention only to providing a lasting solution to its funding needs?

202. Arguments advanced for leaving it outside DDRI – at least in an initial phase - include:

- REANNZ is functional and a merger would be disruptive, risk dislodging valuable personnel and derailing current progress
- A solution to the NeSI structural issues should be addressed before consideration of a possible merger with REANNZ
- In a merged organization, management and governance attention would be distracted from the particular needs, issues and complexities of the connectivity work stream
- The HPC and connectivity investment cycles are on different time scales (one short, one long) which complicates strategic planning and decision making
- There is a distinction in customer orientation in the compute (B2C) and connectivity (B2B) activities that limits potential gains from integration.

203. The DDRI concept is premised on the prospect of achieving benefits from an integrated, researcher-centric service delivery model, allied with a shift to a stronger emphasis on system-wide capability building across the wider eResearch spectrum, enhancing skills and experience, active collaborations nationally and internationally, and catalyzing improved frameworks and progress on data curation and storage.

204. Any organizational reorganization is potentially disruptive and risky. Careful management of change, including timing, sequencing and communication is essential. A suggestion that integration of REANNZ might best come as a secondary phase of change is something that could be considered, but it would risk extending uncertainty and also lessening the REANNZ influence in shaping the future organization.

205. Any issues associated different investment cycles in the compute and connect activities seem not to be especially unusual or intractable for a future DDRI. Nor do differences in customer orientation appear to be problematic. Both REANNZ and NeSI have elements of B2B and B2C in their current models and that is likely to be the case in the future. There is a high degree of overlap of customers served by both organisations. Presenting a unified package of services should be helpful in strengthening the customer focus and simplifying engagement.

Owning or contracting hardware?

206. Future strategies should simply follow the best value options – either owning hardware or contracting use rights as seems cost and performance effective and generally advantageous in meeting researcher needs. As discussed earlier in this report, judgements about best value rest on careful analysis on a like-for-like basis, including factors such as cybersecurity risk, access to support services, ability to readily scale activity up and down, and backup services.

207. This report envisages a future shift to the use of Cloud computing capacity. But that is a choice that would be driven from an assessment of researcher needs, best available options for meeting those needs and relative costs. There is no particular reason to believe that existing NeSI owned CPU & GPU facilities will cease to be an efficient solution in the near term. Indeed, there may be another generation or more to be had from similar, owned or contracted HPC capacity.

208. We envisage a future central compute "portal" through which researchers can gain access to a broader range of compute options, including potentially a range of centrally-negotiated Cloud facilities and, ideally, facilities operated by international research partners. Such an arrangement offers the prospect of access to more specialized or tailored compute capacity and more powerful facilities than are likely to be viable with only locally ownership.

A private industry pathway?

209. NeSI currently offers a private industry pathway to its compute facilities in addition to its core research and education users. Some limited usage by the private sector has occurred. The door is open and should remain so.

210. Encouraging closer engagement between our research community and private sector researchers and innovators should be a strong feature in future strategic thinking for our eResearch infrastructure providers. Such public/private research engagement has been notably limited in New Zealand.

211. REANNZ is constrained by its contract for the Hawaiki cable which stipulates those activities should be principally for research and education. The Hawaiki contract does, however, permit use by the GLAM sector (galleries, libraries, archives and museums) and time limited innovation activities. Given available capacity, and subject to pricing and service levels compared to private network providers, REANNZ should continue to pursue opportunities in those additional market segments.

RECOMMENDATIONS

212. Given the rapidly evolving and cross-disciplinary nature of eResearch in modern science, Aotearoa New Zealand needs to increase its commitment to this field. The current arrangements leave New Zealand lagging in terms of developments in this core national research capability. Current arrangements are also resting on unstable foundations, both financial and, in some part, institutional. Change and additional investment is ahead.

213. The Report Panel recommends the following:

I. The Government pursues Option 3, the Integrated eResearch Platform (or a variant on that concept), with investments intended to create a Data and Digital Research Institute (DDRI).

II. The DDRI should be established as a not-for-profit Crown entity (or similar form, possibly including stakeholder shareholdings).

III. The DDRI should be funded sufficiently to deal with the existing contractual commitments that are currently challenging REANNZ and NeSI.

IV. The DDRI should be funded to enable it to unwind existing in-kind and other sector contributions, with the intention that the DDRI will purchase or contract services from collaborating institutions (or others) as required to conduct its business efficiently. That will likely involve some reassignment of funding that is currently going directly to collaborating partners.

V. The DDRI should be funded to support targeted skills development programmes aimed at building capability in eResearch across the education and research establishment. This should be done in collaboration with universities as they increase their commitment to the field.

VI. Technical support and advice for eResearchers is integral to the eResearch platform. The capacity to provide such support, including on research software and the associated engineering skills, should be part of the DDRI offering.

VII. The DDRI should be funded to provide support across the research community for enhanced research data curation and storage. This would help to develop national standards and support research institutions to adopt and apply improved practice in data curation.

VIII. User charges should apply (at least at the institution level) across the connect and HPC usage, aligned with use but at a lower proportion of total costs than is currently applied. There should be no user charges for skills development or research data curation functions.


illumina
NeSI
eResearch NZ 202
biojulia

Appendix 1: Infrastructure and Related Services to Support New Zealand's eResearch Future: Terms of Reference

INTRODUCTION

1. The Ministry of Business, Innovation and employment (MBIE) periodically considers the direction for investments it administers to ensure settings support government's objectives for the science system. Government contributes funding toward nationally accessible eResearch infrastructure and related services through the Strategic Science Investment Fund (SSIF) mechanisms. This Terms of Reference sets out the scope and process for a report that considers the direction for eResearch infrastructure and services in New Zealand.

CONTEXT

eResearch is crucial for the Research Science and Innovation system

2. Data intensive research (eResearch) has increasingly become a fundamental enabler for most forms of research and is key to promoting Government's goals for the Research, Science and Innovation (RSI) system.

3. Advanced digital tools enable greater global collaboration and ambitious scientific research at a larger scale. The vast amount of data and compute power made readily available through New Zealand's advanced research networks and High Performance Computing capabilities have enabled our scientists to lead developments in areas such as the modelling of climate change and its impact, as well as the sequencing of genomes and our understanding of the COVID-19 virus and its risks.

Researchers' uptake of eResearch requires a co-ordinated effort across institutions

4. Growing eResearch in New Zealand relies upon informed, supported researchers with access to up-to-date tools and methods.

5. Supporting eResearch involves multi-layered capabilities (see Annex Two), underpinned by connect, compute, and data infrastructure pillars. Multiple institutions are involved in delivering these capabilities, including universities and Crown Research Institutes. The major investments in this area are the advanced research network provided by Research and Education Advanced Network New Zealand (REANNZ) and High-Performance Compute (HPC)[1] Services provided by National eScience Infrastructure (NeSI).

There are challenges in ensuring New Zealand's limited resources are best invested in a highly dynamic area

6. To be competitive internationally, New Zealand needs to consider which aspects of the eResearch capability frontier it should pursue and how much. There are a number of challenges to consider, for example:

- Technology advances (and in turn, user needs) in this space are rapid, and models for service provision are constantly evolving. For example, there are trends towards more generic compute infrastructure that serve a larger group of researchers (commoditisation), growing availability of commercial cloud compute and storing data close to compute infrastructure.
- Collaborating on a strategic approach to deliver eResearch capabilities across New Zealand institutions can leverage limited resources[2], however:
    - co-ordination involves a common understanding of roles, a shared sense of direction, and a framework to support complementary individual investment decisions – it's not clear these settings are in place

[1] HPC are state-of-the-art specialised supercomputers and larger clusters of commodity-based processors that allow scientists to scale up their research.

[2] For example, combining to scale up computational capabilities of our RSI system

- building cases to obtain multi-party consensus and often government approval for joint investment can sometimes be at odds with agility needed to chase the frontier
- collaborating institutions highlight difficulties in sharing costs in a way that is both equitable and create the right incentives

7. This work is intended to inform Crown and wider institutions' co-ordinated investment in eResearch. Its findings will also need to weigh the importance of digital infrastructure in the broader economic context. Commercial needs and private research and innovation will influence changes in the market for HPC and related services. Government investment in eResearch capabilities needs to be aligned with Government's strategic goals for digital innovation within industry more generally.

In scope

8. The purpose of the report is:

   What are New Zealand's future needs of infrastructure and related services to support eResearch, and how can NeSI and REANNZ best position their role and capabilities within the system to support these?

9. The report will consider the following areas:

- current state of play for eResearch activities
- strategic context: including key future trends, international comparisons, New Zealand's eResearch problems and opportunities, New Zealand's future eResearch needs
- future state: How New Zealand should position its High-Performance Computing capabilities and related services, Data Management and Storage into the future
- the role of government and the role REANNZ and NESI should play relative to other institutions in the eResearch system
- any implications for funding mechanisms to support the eResearch system.

10. The report questions in Annex One provide further detail to these areas.

11. Given the range of capabilities across institutions needed to support eResearch are broad, this report needs to manage its scope to focus on areas of best impact, particularly where government support may be needed. Annex Two sets out the capabilities that this report will focus on (noting that this focus can be adjusted, if needed, during the course of the report).

How the report product will be used

12. Being clear about how the report product will be used helps inform its shape. The report is intended to (at a high level):

- support collaboration in developing eResearch capabilities across institutions by offering a widely engaged upon future state which is specific enough to offer a framework for decision taking
- support institutions' investment planning to allow a more coordinated approach
- support further policy work on options around role implications and related funding mechanisms for NESI and REANNZ
- support decisions on future Crown investment in eResearch.

Out of scope

13. The report is not intended to be a 'report card' on performance – it will only surface these issues in order to clarify the extent of the shift needed for actors to position their capabilities for the future.

14. The report will not cover the REANNZ connect network in detail – this will be considered as part of REANNZ's network upgrade work.

15. The report will not make recommendations on funding levels, although its findings may have implications for this area.

INTERDEPENDENCIES

16. The following workstreams have links to this report:

- Future Pathways is exploring how to create a connected, adaptable and resilient science system (see: https://www.mbie.govt.nz/have-your-say/future-pathways/). The Future Pathways discussion paper has identified the lack of mechanisms to identify where focused investment in research infrastructure would deliver more value for New Zealand. The discussion paper is currently consulting on design principles for research infrastructure, along with future funding, ownership and operational models for research infrastructure which is relevant to this report. This report is on a faster track than the Future Pathways programme, however report findings and decisions will need to be tested for alignment with direction of travel for the larger Future Pathways programme.
- NESI contract extension and business case: MBIE contributes $7.2 million annually towards NESI. MBIE has agreed a contract extension with NESI until 2023 while MBIE considers policy and investment settings. NESI anticipates the next compute upgrade will be 2023/24.
- REANNZ National Network Strategic Review: in parallel to this work, REANNZ is undertaking preparatory consultation as it plans for its 2023 network upgrade.
- Universities are currently running a joint process to consider future research infrastructure investment. In parallel, Universities are discussing a number of pressing research data management issues, including open data, data sovereignty, and Protective Security Requirements.

PRINCIPLES

17. The report process will be:

- future focused – on future New Zealand needs, science trends, and the eResearch system's adaptability to future shifts in these areas
- facilitative – ensuring participation of expert voices across participating institutions (whilst balancing this with the potential for engagement fatigue)
- focused on a system level view of eResearch needs and capabilities
- open and transparent – ensuring that there are "no surprises" for either system actors or MBIE
- inclusive of Māori, particularly around ensuring their needs around Māori data sovereignty are identified and considered
- efficient – ensuring compliance costs for gathering information is minimised
- sensitive to the need to ensure appropriate protection of information.

KEY STAKEHOLDERS

18. Key stakeholders include:

- NESI
- REANNZ
- NESI host: University of Auckland
- Other NESI Collaborators: NIWA, Manaaki Whenua – Landcare Research, University of Otago

19. Other significant stakeholders include wider CRIs and Universities.

20. Wider stakeholders include Independent Research Organisations, Departments and Corporates with significant research capabilities.

PROCESS FOR THE REPORT

Report Owner

21. Oversight of this process is through MBIE's Infrastructure Steering Group. The report owner is the General Manager, Science Innovation and International.

Report Panel

22. An external Report Panel (Report Lead and Panel Member) will be appointed that has the following strengths: credibility at a governance level, strong facilitation skills, an understanding of shared service and funding models, and an understanding of the eResearch space.

Expert Reference Group

23. An Expert reference Group of three to five people (including international members) will be identified by MBIE (in discussion with key stakeholders) to support the Report Panel throughout the process with views on:

- expert user perspectives
- eResearch trends
- comparisons with overseas jurisdictions
- the Report Panel's draft direction of travel.

Report process

24. The expected report phases are:

    1. Terms of Reference finalised in consultation with key stakeholders
    2. Background material developed with input from NESI, REANNZ and other key stakeholders
    3. Report Panel recruitment, commissioning meeting with MBIE and review of background material
    4. Identify and on-board International Reference Group
    5. Approach for engaging with Māori on data needs and sovereignty issues identified
    6. Initial Panel meeting with NESI and REANNZ Boards: Boards own reflections on the landscape and where they see the issues and opportunities
    7. Initial meetings with other key stakeholders
    8. Panel initial meeting with Minister
    9. Interviews with wider stakeholders
    10. Mid-way Check in with NESI and REANNZ Boards
    11. Expert Reference Group check in
    12. Future state workshops with key stakeholders and development of options
    13. Expert Reference Group check in
    14. Any follow up engagement necessary
    15. Draft Report discussion with MBIE, Key Stakeholders, and Expert Reference Group
    16. Report finalisation, including Key Stakeholders' response
    17. Panel meeting with Minister to discuss report
    18. Release.

25. These phases may be refined based on discussions between the Report Panel, NESI and REANNZ. The Panel will have regular 'check ins' with NESI and REANNZ, along with representatives from key Stakeholders throughout the process to test thinking and ensure no surprises. The Panel will also check in with MBIE on progress.

26. The process is expected to commence January 2022, with a final report and response targeted for late May 2022.

ADMINISTRATION

27. MBIE's Entity Performance and Investment Team (MBIE) will recruit and provide support for the Report Panel throughout the process. The key contact for the process is Alan Vandermolen, Director, Entity Performance and Investment. MBIE will work with NESI and REANNZ to develop background material for the Report Panel.

28. The cost of the report will be met by MBIE.

29. The report will be subject to requests under the Official Information Act (1982). The Panel's final report will be released publicly, with any commercial-in-confidence material withheld.

Annex One: questions to consider in developing report product

For eResearch infrastructure[3] and related services:

Current arrangements

a. Who are the key users in New Zealand?

b. What is the current demand in New Zealand?

c. What are the current roles (and intersects) of institutions in New Zealand and existing arrangements?

d. What capability exists through New Zealand (both public and privately owned)?

Strategic context:

e. What does the international landscape for provision of services look like - including the range of:

- Future trends
- capabilities available
- models for provision of services

f. What are New Zealand's most pressing problems, gaps, and opportunities?

g. Considering technology trends, what is the likely future demand from existing and future users in New Zealand?

Future State

h. How far off are we from the frontier in our capabilities, and where do we want to be?

i. How should New Zealand be positioning our capabilities into the future, specifically taking into account:

- How to grow demand and researcher capability?
- public versus private provision?
- Opportunities for international collaboration (or with adjacent New Zealand activities)
- potential resilience and data sovereignty (including Māori data sovereignty) issues

Role and funding implications

For organisational roles and funding of eResearch Infrastructure and related services:

j. What are the trends in overseas jurisdictional arrangements?

k. What are New Zealand's most pressing problems and opportunities?

l. What is government's role? What is NESI and REANNZ's role relative to other institutions (and the private sector), and how could this best be positioned?

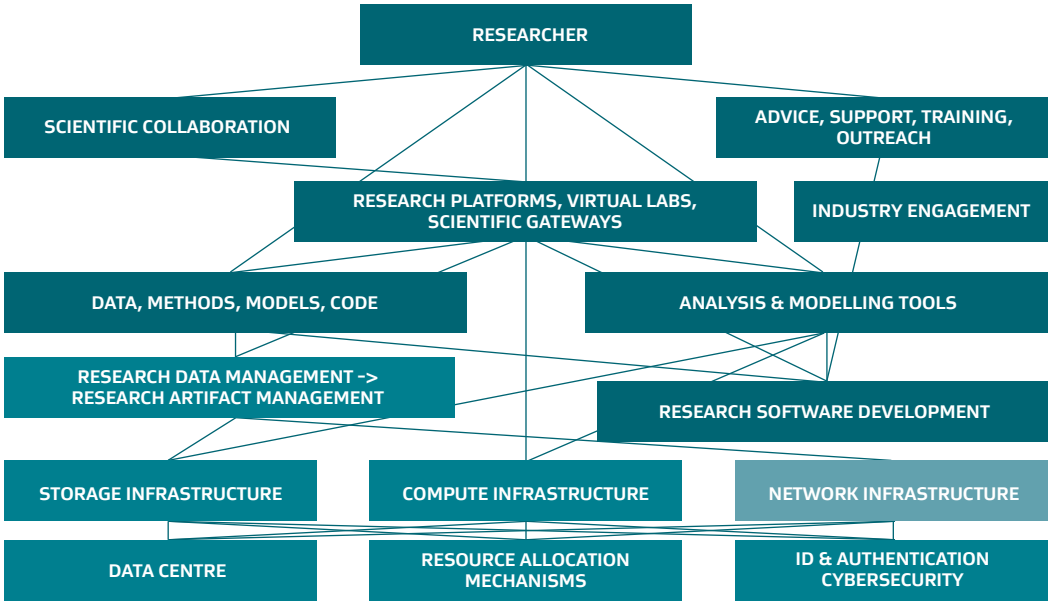
m. How could funding mechanisms best support eResearch optimisation and efficiency while being reasonably equitable across collaborators?

[3] HPC , Research Data Management and Data Storage (including ID and authentication, and cybersecurity) – see Annex 2.

Annex Two: more detail on report scope

The diagram below[4] sets out:

- The main capabilities needed within the eResearch system to enable Researchers, which are delivered by a range of institutions and private providers.
- The indicative level of focus that will be applied to these areas (in order to manage the scope of the report, and put more effort into areas of greatest concern/ opportunity) – these scope judgements will be tested in early exploratory work with Key Stakeholders.



Summary coverage:
main focus will be on clarifying roles and setting context for other areas (while strategically important, no significant service provision co-ordination issues have been highlighted to date in scoping the report).

Out of scope:
to be covered as part of REANNZ network refresh work, although funding aspect models are in scope.

More focus:
Remaining areas are main focus of this report (although Research Data Management is potentially a large area and won't be covered in-depth).

[4] Refer: Mark Dietrich, Delivering Effective eResearch Services: Better Practices Drawn from Benchmarking Eight Jurisdictions Around the World, Final Report of International Benchmarking Study for New Zealand eScience Infrastructure (NeSI), 2019


Carl
Olsen

Appendix 2: Report Panel Profiles

MURRAY SHERWIN CNZM
Chair

Murray is an economist with over 40 years' experience across a wide range of public policy roles. His previous roles include Chair of the New Zealand Productivity Commission, CEO/Director General of the Ministry of Agriculture and Forestry, Deputy Governor of the Reserve Bank of New Zealand and Chief Manager of the Reserve Bank's Banking Supervision, Financial Markets and International Departments. He has experience as an economist at the OECD, member of the Prime Minister's Policy Advisory Group, member of the Board of Executive Directors at the World Bank, and member of the Advisory Board of the New Zealand Debt Management Office.

He was Chair of the Innovation Partnership in 2012 to 2016, working with partner organisations, including Google, Chorus and Internet New Zealand, to promote smart use of internet technologies. He currently Chairs the Plant Germplasm Import Council, working with New Zealand's plant-based industries and the Ministry for Primary Industries to facilitate efficient and biosecure access to new plant genetic material to boost innovation, productivity and resilience in our plant industries.

Murray is an accredited member of the New Zealand Institute of Directors, member of the New Zealand Association of Economists, New Zealand Institute of Forestry and New Zealand Institute of Agricultural and Horticultural Science. He is a graduate of the University of Waikato (MSocSci, Hons.).

JOE MARKS
Panel Member

Joe Marks is Weta Digital's Chief Technology Officer, overseeing the company's technology initiatives across visual effects and animation.

Joe brings 35 years' experience as a technology executive. He was most recently Executive Director of the Center for Machine Learning and Health at Carnegie Mellon University, where he worked at the forefront of innovation in digital healthcare. Previously, he was Vice President and Research Fellow at Disney Research, leading research and development across labs in Pittsburgh, Zurich, Boston, San Francisco and Los Angeles, and Research Director at Mitsubishi Electric Research Labs.

Joe's research interests include computer graphics, video processing, media distribution, robotics, human–computer interaction, mobile computing, computer vision, artificial intelligence, sensors, embedded systems and behavioural economics. He has been active in the SIGGRAPH community, serving as its Papers Chair in 2004 and Conference Chair in 2007. He also has a broad portfolio of peer-reviewed publications in applied computing. Joe has managed significant corporate and academic research and development, in addition to co-founding two e-commerce start-ups. He holds a Bachelor of Arts in Applied Mathematics and a Ph.D. in Computer Science, both from Harvard University.

Appendix 3: Expert Reference Group Members

ASSOCIATE PROFESSOR JOSEPH LANE
University of Waikato
Associate Professor, Physical & Theoretical Chemistry

PROFESSOR KAREN WILCOX
University of Austin Texas
Director, Oden Institute for Computational Engineering and Sciences

PROFESSOR MICHAEL WITBROCK
The University of Auckland
Professor of Computer Scientist in the field of artificial intelligence

ROSIE HICKS
Australia Research Data Commons
Chief Executive of the Australia Research Data Commons

Appendix 4: Consultation List

The Report Panel is grateful to the following organisations who contributed insights to this report:

- AgResearch (AgR)
- Amazon Web Services
- Associate Minister of Research, Science and Innovation
- Auckland University of Technology (AUT)
- Callaghan Innovation
- Department of Internal Affairs
- Genomics Aotearoa
- Google
- Government Communications Security Bureau (GSCB)
- Hawaiki Cable
- Institute of Environmental Science and Research (ESR)
- Institute of Geological and Nuclear Research (GNS)
- Kordia
- Lincoln University
- Manaaki Whenua Landcare Research (Manaaki Whenua)
- Massey University
- Microsoft New Zealand
- Ministry of Business, Innovation and Employment
- Ministry of Education
- Ministry of Primary Industries
- New Zealand eScience Infrastructure (NeSI)
- NeSI Reference Group
- New Zealand Forest Research Institute Limited (Scion)
- New Zealand Tech Alliance (NZ Tech)
- National Institute of Water and Atmospheric Research (NIWA)
- Plant and Food Research
- PlantTech
- Prime Minister's Chief Science Advisor
- Research and Education Advanced Network New Zealand (REANNZ)
- Science New Zealand
- The Science for Technological Innovation Challenge (a national science challenge)
- The Treasury
- The University of Auckland
- Universities New Zealand
- University of Canterbury
- University of Otago
- University of Waikato
- Victoria University of Wellington
- Xerra

Appendix 5: Key Documents Reviewed

Agreement between the Government of New Zealand and the Government of the United States of America on Science and Technology Cooperation Contributing to Domestic and External Security Capabilities (2010), United States-New Zealand, Article 2.

AgResearch. (2021). Board Papers: Procurement of eResearch Infrastructure.

AgResearch. (2021). Digital Blueprint Rautaki Matihiko FY22 - FY25.

Allen + Clark. (2017). Evaluation of the New Zealand eScience Infrastructure (NeSI). New Zealand Government.

Amazon Web Services. (2022). Cloud Powering Faster, Greener, and More Collaborative Research.

Arrangement relating to Science Technology and Innovation Cooperation (2019), New Zealand-Singapore, Paragraph 5 Statute 6-8, pg. 6-8.

Australian Department of Education, Skills and Employment. (2016). National Research Infrastructure Roadmap Research Infrastructure Investment Plan. Australian Government.

Australian Department of Education, Skills and Employment. (2018). National Collaborative Research Infrastructure Strategy: 2018 Guidelines.

Australian Department of Education, Skills and Employment. (2020). National Research Infrastructure Roadmap Research Infrastructure Investment Plan.

Australian Department of Education, Skills and Employment. (2021). 2021 National Research Infrastructure Roadmap. Australian Government.

Australian Research Data Commons, & Institute Methods for Innovation. (2019). Investigating the Link Between Research Data and Impact. Australian Government.

Australian Research Data Commons. (2021). Australian Research Data Commons Strategy; Leading Australia to data-driven research impact. Australian Government.

Carroll, S.R., Garba, I., Figueroa-Rodríguez, O.L., Holbrook, J., Lovett, R., Materechera, S., Parsons, M., Raseroka, K., Rodriguez-Lonebear, D., Rowe, R., Sara, R., Walker, J.D., Anderson, J. and Hudson, M., (2020). The CARE Principles for Indigenous Data Governance. Data Science Journal, 19(1), pg. 43.

Compute Canada. (2016). Compute Canada Report: Increasing Canadian Research Impact.

Constitution of Research and Education Advanced Network New Zealand Limited (REANNZ) (2007).

Department of Internal Affairs. (2014). Cloud Computing Information Security and Privacy Considerations.

Dietrich, M. & Jones, N. (2018). Understanding the eResearch Ecosystem in New Zealand. NeSI.

Dietrich M & Jones N. (2019). Delivering Effective eResearch Services: Better Practices Drawn from Benchmarking Eight Jurisdictions Around the World. NeSI.

Digital Research Alliance of Canada. (2021a). Digital Research Alliance Canada Corporate Plan, 2021-2022. Government of Canada.

Digital Research Alliance of Canada. (2021b). Renewed Policy Guidance on Integrated Hardware (Contributed Systems). Government of Canada.

Digital Research Alliance of Canada. (2021c). Researcher Needs Assessment: summary of what we heard. Government of Canada.

Digital Research Alliance of Canada. (2021d). Strategic Plan 2022 - 2025. Government of Canada.

Ministry of Business, Innovation and Employment, The Treasury, & Ministry of Education. (2020). Reviewing feasibility for closer collaboration between REANNZ and the Network for Learning Ltd.

Ministry of Business, Innovation and Employment. (2014). Review of the New Zealand e-Science Infrastructure (NeSI).

Ministry of Business, Innovation and Employment. (2015). Briefing 1426 15-16 Strategic Refresh of NZ's eResearch Infrastructure Investment: Draft Report.

Ministry of Business, Innovation and Employment. (2015). Strategic Refresh of eResearch Infrastructure Investment.

Ministry of Business, Innovation and Employment. (2016). Future Directions for Research Infrastructure in New Zealand.

Ministry of Business, Innovation and Employment. (2019). Briefing 0108 19-20 - A new approach to eResearch infrastructure in New Zealand.

Ministry of Business, Innovation and Employment. (2021). Briefing 2021-1766 eResearch issues and opportunities.

Ministry of Business, Innovation and Employment. (2021). Briefing 2021-2926 eResearch work programme.

Ministry of Business, Innovation and Employment. (2021). Briefing 2122-2010 eResearch work programme update: Report on Infrastructure and Related Services to Support New Zealand's eResearch Future.

Ministry of Business, Innovation and Employment. (2021). Te Ara Paerangi Future Pathways Green Paper 2021.

Ministry of Business, Innovation and Employment. (2022). Digital Technologies: Draft Industry Transformation Plan 2022-2032.

Moore, D., Tran, L., Uddstrom, M., & Yarrall, D. (2018). New Zealand's high speed research network: At a critical juncture. Sapere.

National Collaborative Research Infrastructure (NCRIS), & Lateral Economics. (2021). National Collaborative Research Infrastructure (NCRIS) spending and economic growth.

National Science Foundation. (2019). Transforming Science Through Cyberinfrastructure: NSF's Blueprint for a National Cyberinfrastructure Ecosystem for Science and Engineering in the 21st Century.

New Zealand Digital Skills Forum. (2021). Digital Skills Aotearoa Report: An updated analysis of the Digital Skills landscape of New Zealand.

New Zealand Government. (2021). Te koke ki tētahi Rautaki Matihiko mō Aotearoa: Towards a Digital Strategy for Aotearoa.

NeSI. (2018). Growing researchers' computational skills to meet future needs; Review and Analysis of NeSI's Training Strategy 2015-2017.

NeSI. (2019). NeSI 3 Investment Case 2019-2026.

New Zealand Science, Research and Innovation Cooperation Agreement (2017), Australia-New Zealand, Article 5 (4).

OECD Global Science Forum. (2020). Building digital workforce capacity and skills for data-intensive science.

OECD. (2022). OECD Economic Surveys: New Zealand 2022.

REANNZ (2018), REANNZ Annual Report 2018, Year ended 30 June 2018.

REANNZ (2019), REANNZ Annual Report 2019, Year ended 30 June 2019.

REANNZ (2020), REANNZ Annual Report 2020, Year ended 30 June 2020.

REANNZ (2021), REANNZ Annual Report 2021, Year ended 30 June 2021.

REANNZ & NeSI. (2013). Options Paper: Towards a Smarter eResearch Infrastructure Sector.

REANNZ, NeSI, & NZGL. (2016). eResearch 2020: National Research Data Programme: The Case for Research Data.

TDB Advisory. (2014). Review of National eScience Infrastructure. Ministry of Business, Innovation and Employment.

Wilkinson MD, Dumontier M, Aalbersberg IJ, et al. (2016) The FAIR Guiding Principles for scientific data management and stewardship. Sci Data. 15;3:160018.