NCI AUSTRALIA

ANNUAL REPORT

2021-2022

SUPERCOMPUTING AND BIG DATA LEADERSHIP FOR AUSTRALIAN INNOVATION

This report and additional data are available from the NCI website at nci.org.au/about-us/annual-reports

Requests for authorisation and enquiries concerning the contents of the report should be directed to enquiries@nci.org.au

Produced by
National Computational Infrastructure
143 Ward Road, The Australian National University, Acton, ACT 2601

Designed by Result Design.

Cover Image:
Detail from the Nirin artwork by Anthony Best (Canberra-based Indigenous artist). Commissioned for the introduction of NCI's latest cloud computing system, this artwork shows lines of communication, meeting at gathering places represented by the circular features. The arch shapes represent scientific communities researching in the Nirin cloud surrounded by data.

We acknowledge the Traditional Custodians of the ACT, the Ngunnawal people. We acknowledge and respect their continuing culture and the contribution they make to the life of this city and this region.

CONTENTS

INTRODUCTION

About NCI

Australia's National Computational Infrastructure, NCI, is a unique supercomputing and big data facility supporting a vast network of scientists, research infrastructure providers and national science agencies. Comprising one of the country's most powerful supercomputers alongside extensive data repositories, NCI provides the data management, infrastructure and services to meet Australia's high-performance data, storage and computing needs.

NCI's service offerings are critical to the nation. They support high-impact research shaping policy development and drive the innovation needed to address contemporary challenges and realise new economic opportunities. This includes providing essential resources for the development of new materials, quantum technologies and health advances such as personalised medicine.

The Australian Government and the Australian research sector come together through NCI in a broad collaboration involving the largest national science agencies (CSIRO, the Bureau of Meteorology, Geoscience Australia), the National Collaborative Research Infrastructure Strategy (NCRIS) network of research infrastructure capabilities, universities, the Australian Research Council and industry partners. This collaboration underpins NCI's offering for the research community: a powerful, shared resource that no single organisation could build on its own. The collective leveraging of investments from across the country has created a highly successful enabler of innovative research. NCI empowers government agencies, universities and industry across nationally significant research domains.

NCI's provision of tightly integrated supercomputing hardware, data and computational expertise to researchers enables advanced scientific outcomes. Robust computational results derived by Australian researchers provide insights that inform and benefit public policy and accelerate new scientific and technological advances. Whether it is computational simulations for drug and materials development, high-resolution modelling for extreme weather prediction or code modernisation using the latest methods in artificial intelligence and machine learning, NCI makes possible the impossible, facilitating high-impact research that could not otherwise be undertaken in Australia.

Our Mission and Strategic Vision

NCI's mission is to radically enhance the high-performance computational methods and capabilities available to Australian researchers. By doing this, we add value to Australian society, fulfil the needs of our collaborators and enable transformative ideas. NCI underpins a robust research sector which delivers benefits to Australia's industry, environment, policy and society.

We expand the range, scope and ambition of Australian research. Our world-class computing and data systems, services and practices enable thousands of computational and data scientists each year to produce breathtaking science that benefits and impacts the country. As data-intensive methods become ever more central to Australian science, NCI's reliable and innovative high-performance platform is essential. NCI is pushing the boundaries of what such a facility can offer: tightly integrated high-performance computing (HPC), high-throughput computing (HTC) and high-performance data (HPD) open up previously unimagined possibilities for Australian research.

NCI enables Australia to have a globally competitive HPC, HTC and HPD capability. Through us, Australian researchers continue to experience a significant shift in what they can do computationally. As a result of our services and investment in the capability of our users, the number of researchers using high-end computing in Australia is growing, laying the foundations for Australia to produce even more high-impact discoveries and innovations.

Chair's Report

High-performance computing, including big data (HPCD), has become an indispensable tool for advanced analysis and simulations, across science and industry. This is why countries around the world are investing in increasingly powerful computing and data capability: this is the driving force behind scientific innovations and the development of future technologies. HPCD provides a much sought-after competitive edge on the global scale.

The National Computational Infrastructure (NCI) plays a central role in ensuring Australia remains at the forefront of HPCD. Supported by the Department of Education's National Collaborative Research Infrastructure Strategy and our Collaboration partners, NCI is continually investing in both computing capability and, most importantly, people. It is this combination that is key to the world-leading services NCI offers. We couple continual enhancements to our supercomputing, data and cloud infrastructures with a collaborative and tailored approach to working with the research community.

A commitment to effectively meeting the needs of our users is evident in many of NCI's key achievements in 2021–2022, with initiatives focused on ensuring Australian researchers can access, and are best-equipped to use, HPCD resources.

NCI's new Adapter Allocation Scheme, launched in December 2021, was developed in response to demand from our current and emerging user communities for a scheme that allows for flexible access to HPC and cloud. After a successful pilot in early 2022, Adapter is up and running, supporting meritorious research projects that need flexible compute access over a short period. Adapter is

not only increasing the number of researchers able to make use of NCI's facilities, but is also encouraging innovative approaches to undertaking computational and data-intensive research.

Similarly guided by user feedback, NCI has made a significant investment in our training offering since 2021. As a result, today NCI's training program is a complete package, with diverse courses ranging from foundational HPCD skills to artificial intelligence and machine learning applications. The many workshops, sessions and courses NCI offers on a regular basis benefit thousands of researchers each year, equipping our users to make the most of NCI's computing capability and pursue new fields of enquiry.

Strong partnerships have underpinned NCI's ability to effectively support the Australian research ecosystem. This continues to be the case as NCI deepens its contribution to new areas, including bioinformatics and genomics research. As part of PrOSPeCT (the Precision Oncology Screening Platform enabling Clinical Trials), NCI is working with diverse public and private partners to deliver a game-changing genomic medicine initiative. Through this project, new treatment paths will open up for thousands of Australians diagnosed with rare and 'untreatable' cancers.

The recent successful launch of the Australian Earth System Simulator (ACCESS-NRI) also highlights how NCI is helping other national research infrastructure facilities to achieve their missions. NCI supercomputing has been essential to the development of the ACCESS suite of climate and weather models to date, and NCI looks forward to providing the computing support required as ACCESS-NRI works to amplify the performance of this new critical infrastructure. As the 2021 National Research Infrastructure Roadmap highlights, this type of close integration is key to realising the greatest impact from Australia's research investment. NCI welcomes the release of the Roadmap and looks forward to continuing to work collaboratively with all stakeholders to progress the important issues it outlines as part of the Research Infrastructure Investment Plan process.

This year, NCI welcomed the University of New South Wales as a new Major Collaborator, showing the value that the research community sees in NCI. I am particularly pleased that the Board was further strengthened by the appointment of 3 new independent members who commenced in December 2021. NCI recognises the value of bringing new and diverse partners, researchers and end-users into the organisation. This broadens the expertise and experience that can be drawn on, enhancing the capabilities available to Australian researchers through increased connections across research, industry and government. Critically, it also means that more end users than would otherwise be possible can benefit from the world-class scientific computational capabilities that NCI offers.

As I write this report, NCI is undertaking the strategic planning needed to prepare for major upcoming capability expansions in 2022 and 2023. This bolstered capability will allow NCI to further broaden and deepen the support it offers the research community, helping to solve scientific computational problems in the national interest for an ever more diverse group of end users and beneficiaries.

I am very pleased to present to you NCI's 2021–2022 Annual Report and excited to see what can be achieved in the coming year.

Dr Greg Ayers NCI Advisory Board Chair

NCI Governance

THE NCI ADVISORY BOARD

NCI is governed by The Australian National University on the advice of the NCI Advisory Board, which comprises:

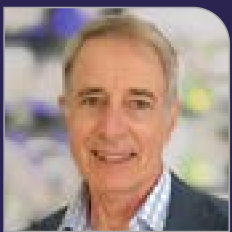
> an independent chair appointed by the Advisory Board,

> the Director of NCI,

> one member appointed by each of the Major Collaborators, and

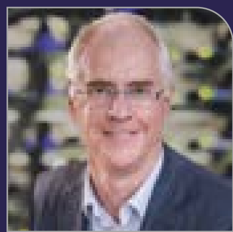
> additional independent board members appointed by the NCI Advisory Board on the basis of their expertise.

ADVISORY BOARD MEMBERS



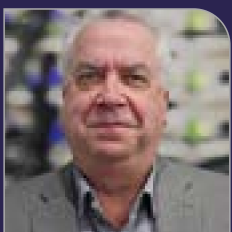
Dr Greg Ayers
Chair

Professor Sean Smith
FAAAS
NCI, Director

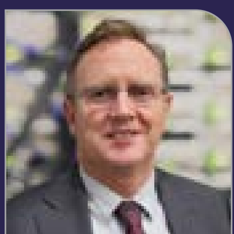


Professor Keith Nugent
FAA
The Australian National University, Deputy Vice-Chancellor (Research and Innovation)

Professor Elanor Huntington
CSIRO, Executive Director Digital, National Facilities and Collections



Dr Gilbert Brunet
FCMOS
Bureau of Meteorology, Chief Scientist and Group Executive (Science and Innovation)



Dr James Johnson
Geoscience Australia, Chief Executive Officer

Professor Grainne Moran
UNSW Sydney Pro-Vice Chancellor Research Infrastructure

Dr Simone Richter
ANSTO, Group Executive (Nuclear Science & Technology and Landmark Infrastructure)

Professor Melodie McGeoch
La Trobe University, School of Life Sciences

Dr Rosemarie Sadsad
Evidentli, Head of Solution Delivery

Ms Clare McLaughlin
National Health and Medical Research Council, General Manager

Accurate as at 30 June 2022.

NCI'S ORGANISATIONAL STRUCTURE

In the second half of 2021, NCI revised its structure to ensure the organisation is best placed to deliver the next generation of HPCD in Australia. The restructure allows NCI to build increased capacity in the organisation and better serve users, including by consolidating communications, training, user support and administrative service roles within a newly created Business Development and User Engagement team.

Under the new arrangements, 3 divisions, each led by a Deputy Director, report to the NCI Director. In addition to Business Development and User Engagement, there are the Innovative Compute Environments team (supporting infrastructure, development and security for HPC, cloud and storage services) and the HPC and Data Innovation team (supporting the growth and development of NCI's data collections, data science, modernisation of HPC and data codes, and visualisation).

In addition, a small directorate comprised of administrative, finance and Executive Officer staff provides support to the NCI Director and broader organisation.

Director’s Report

Dear colleagues and friends,

Reflecting on yet another extraordinary and at times challenging year, I am very proud of the exemplary services NCI provides to the Australian research community and how we are finding new opportunities to offer even greater benefits from diversified utilisation of NCI’s high-performance computing and data capabilities. This is no mean feat, and relies on both the strategic vision of our Board, ably led by Dr Greg Ayers, and the collective efforts of NCI staff, who consistently bring their commitment and expertise to make this vision a reality.
I sincerely thank everyone for their ongoing contributions.

NCI has long recognised that to unlock the full potential of our world-class computational infrastructure, including the Gadi supercomputer, an entire ecosystem of resources and support is needed. We have recently focused on pursuing initiatives that further enhance this ecosystem: making our services even more accessible, supporting our users to build computational science skills, and expanding our involvement in new and emerging areas of HPCD service.

I am glad to say these efforts are bearing fruit, with a number of major achievements and milestones reached in 2021–22. More users and research communities are now able to access, and benefit from, NCI services, and we are establishing new capabilities to support a greater variety of computational and data-analysis projects.

- In April, NCI officially launched our Nirin cloud computing platform, enabling interactive workflows, data analysis capabilities and other tailored tools for a variety of research communities. The eye-catching artwork commissioned to mark the event, created by Indigenous artist Anthony Best, is featured on the cover of this report.
- The Adapter Allocation Scheme, launched in late 2021, is NCI’s newest merit scheme and caters to researchers needing smaller high-performance computing and cloud resources over a short period. The flexible quarterly application schedule is bringing new users to NCI and encouraging innovative compute approaches.
- We have significantly expanded the range and number of training opportunities we provide. Over 3,600 people attended courses and events we hosted in 2021–22, gaining the skills needed to make the most of computational and data science methods. NCI will continue to work with the research community to ensure our training program keeps meeting users’ needs.
- As a foundation partner in the PrOSPeCT project, NCI is helping to deliver an initiative that will open up new treatment paths for people across Australia with difficult-to-treat cancers. The project reflects NCI’s increased focus on developing tailored services and technologies to help enable the enormous potential of bioinformatics and genomic medicine to be realised.

NCI’s partners and Collaborators are key to our work. We continue to strengthen our existing relationships with Australia’s national research infrastructure facilities, while also building new connections domestically and abroad. Of particular note in 2021–22 was the official launch of ACCESS-NRI as a valuable new infrastructure facility within the NCRIS network. NCI has been closely involved in the development of the ACCESS model suite and looks forward to supporting the next generation of climate and weather models.

On the international front, in March this year NCI and the National Supercomputing Centre Singapore signed a Memorandum of Understanding, building on our already close partnership. This provides a foundation for further joint activities, including software development and data sharing. This type of regional cooperation delivers major benefits for supercomputing, big data and high-throughput computing users throughout the Asia-Pacific region. I look forward to further deepening regional ties, including through jointly organising events such as the Supercomputing Asia conference in 2023 and 2024 with colleagues in Singapore, Thailand, Japan and South Korea.

Professor Sean Smith, FAAS
Director, NCI

Highlights of the Year

Delivered 16 Skills Sharpening training courses on fundamental programming skills in partnership with Intersect Australia.

Announcement of the PrOSPeCT project. NCI is a foundation partner in this industry-research collaboration to develop a cancer screening platform.

Launch of NCI's research cloud platform, Nirin, enabling more interactive workflows, data analysis capabilities and support for the research community.

N D J F M A M J

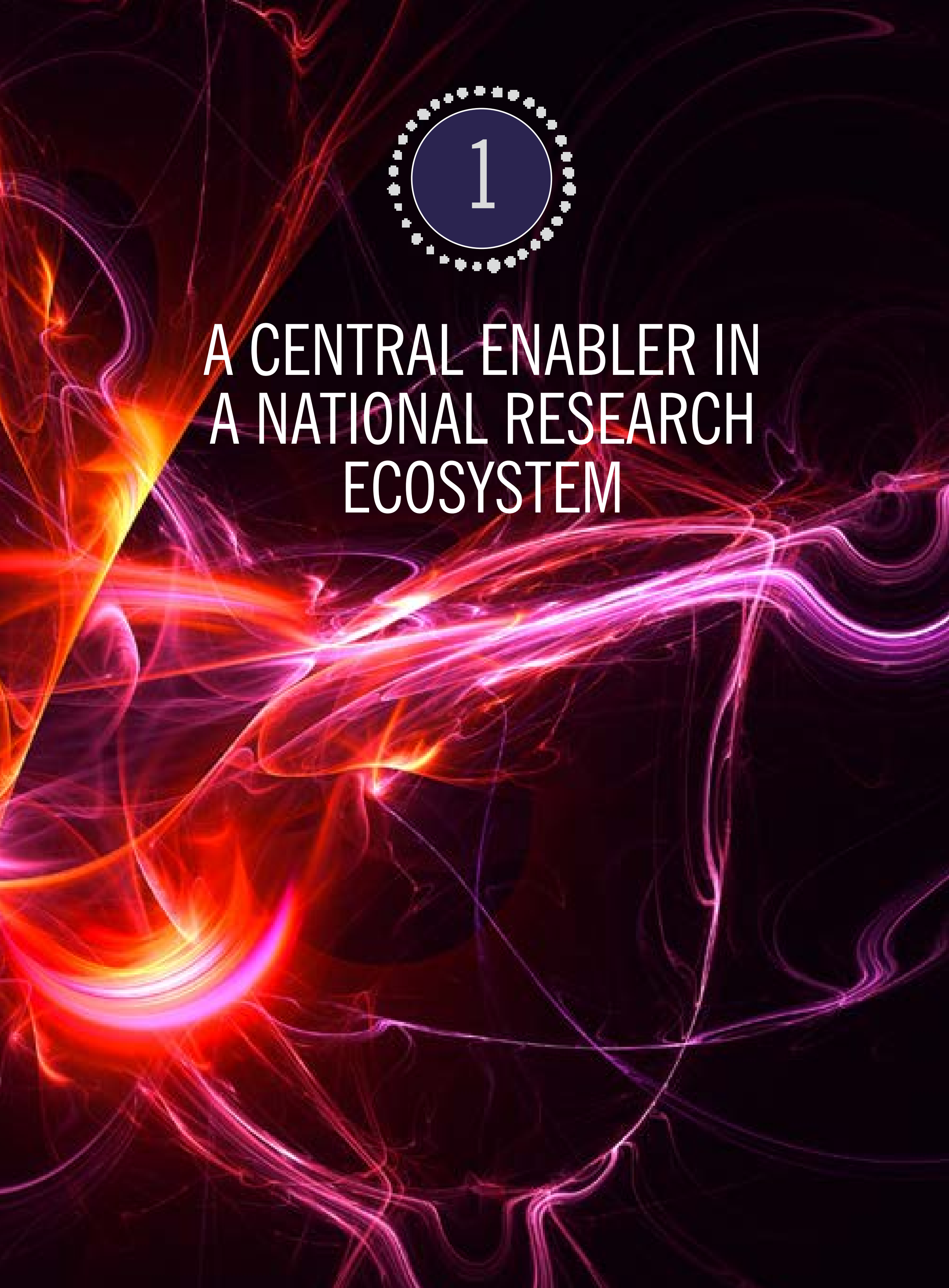
2021 2022

NCI launches the Adapter Allocation Scheme, a new merit scheme providing flexible access to high-performance computing and cloud.

NCI and the National Supercomputing Centre Singapore sign a Memorandum of Understanding, strengthening existing collaboration.

NCI announces grants to Australia's leading computational researchers totalling 140 million units of computing time.

ALCG



1

A CENTRAL ENABLER IN A NATIONAL RESEARCH ECOSYSTEM

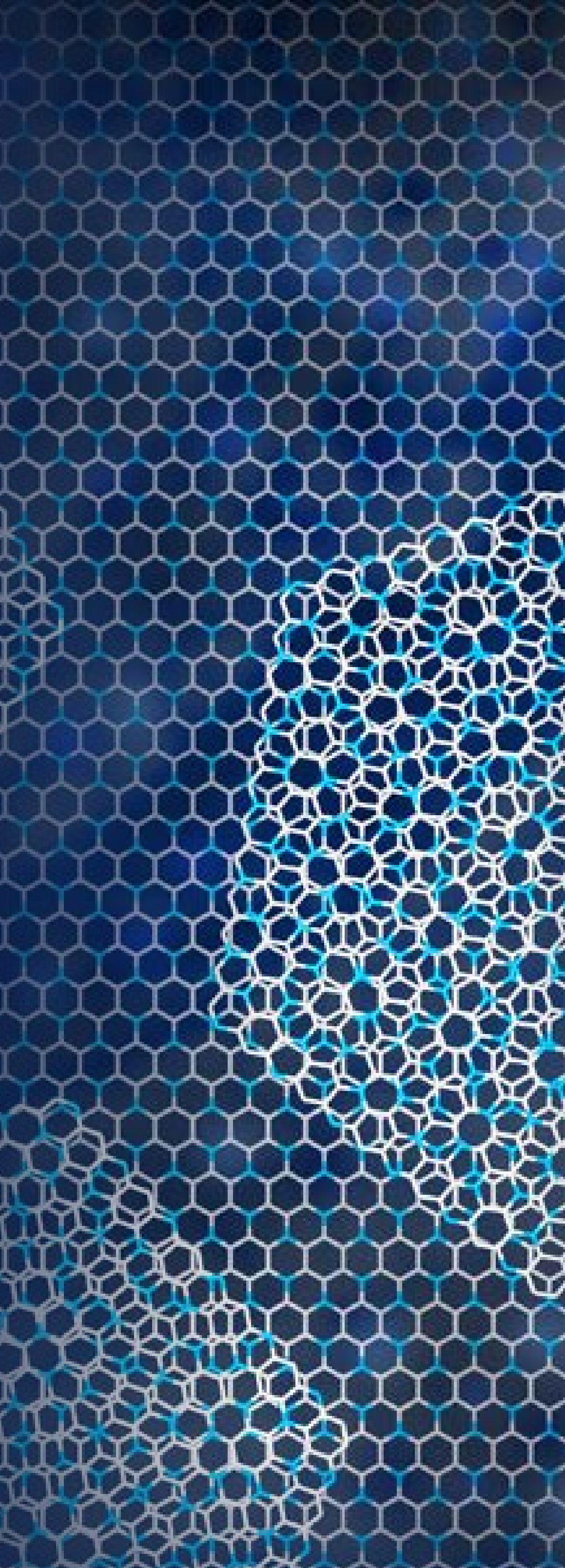
Snapshot – Developing the Materials and Technologies of the Future

Computational science undertaken at NCI is playing a key role in the design of diverse new materials and technologies. These innovations will underpin efforts to address the world's most pressing challenges, improve the quality of products available to consumers, and facilitate the next breakthroughs in science. In some instances, this research is occurring in close collaboration with industry, with patents lodged and partnerships established to oversee the long process of taking new products to market. In other cases, researchers are laying down foundational knowledge, enabling future opportunities to be explored and unforeseen challenges addressed.

Developing and applying new scientific knowledge to help us live more sustainably is a central theme of many researchers' work. Professor Katya Pas from Monash University, recipient of a 2021 Australasian Leadership Computing Grant, uses large-scale quantum chemical calculations to run simulations of up to 500,000 atoms in a system. This includes investigating the properties of phase change materials (PCMs) that can be used to efficiently heat and cool buildings. Better understanding intermolecular interactions like hydrogen bonding will help unravel the origin of the bulk properties of these complex materials.

Professor Pas and Dr Guiseppe Barca (from the ANU), her collaborator on the grant, have already uncovered how the interplay in intermolecular properties results in an increased latent heat at lower temperature of PCMs based on organic ionic liquids. This discovery, made possible by the new GPU-based code called Exess, has already demonstrated a design of better PCMs that absorb heat at much lower temperatures than existing organic ionic salts. The outcomes of Professor Pas' research will help to further unlock the true potential of PCMs to reduce the world's energy consumption.

Professor Debra Bernhardt from The University of Queensland is also undertaking research to better inform the design of products and help with the transition to a clean energy economy. She uses Gadi to run computational simulations that inform the design of a variety of new materials used in devices including supercapacitors, batteries and micro-electronics.



Visualisations of bilayers of two-dimensional C3N with different twist angles from Professor Debra Bernhardt's research. Twisting of one layer with respect to the other results in different electronic properties of the bilayers.

“Supercapacitors and rechargeable batteries are a vital energy storage technology required to make the availability of renewable energy more reliable. This is necessary to address the challenges of climate change,” she says. “Our work involves developing and applying novel molecular modelling methods to improve energy storage.” These new methods make it possible for out-of-equilibrium systems, such as those that are being charged or discharged, to be more accurately represented in simulations. The research has provided numerous insights, including how pore design impacts the charging and discharging mechanism in carbon-based electrodes. These can be harnessed by industry in the design of more effective and efficient electrodes.

Separate studies pursued by Professor Bernhardt’s team focused on overcoming the technical challenges involved in using carbon-based materials such as graphene to develop electronic devices. While there is significant excitement around the potential use of these materials in electronics, applications have to date been constrained by our limited ability to ‘tune’ their electronic properties: that is, to control whether materials are more or less conductive. The study found that the conductivity of bilayer C3N, a carbon-based 2-dimensional material, could be altered solely by changing the orientation of stacked 2-dimensional layers of the material. As the bilayer is very thin, it is transparent and flexible, so this finding could lead to applications where transparent circuitry is required.

While many researchers are focused on either optimising or finding new applications for existing technologies, others are focused on the development of entirely new innovations, such as quantum computing. Further development of quantum computers relies on the ability to understand and closely control materials at the nanoscale. While critical, experiments are difficult to undertake and the data produced can be complicated to interpret. For these reasons, simulations play an important role in modelling the quantum computing systems being built, helping to guide experimentation and make sense of measurements.

Associate Professor Rajib Rahman from UNSW Sydney is undertaking the detailed simulations needed to support the further development of silicon-based quantum computing in Australia. “We rely on NCI’s Gadi supercomputer to run detailed simulations, atom by atom,” he says. “Our results feed back directly into experiments, informing for example where atoms are placed in the silicon matrix and how voltages are applied. This is key to the development of prototypes with more qubits or quantum bits, particularly systems where these can interact and work together.”

Professor Rahman and his team are excited to see how the technology develops in coming years. While the building of a large-scale quantum processor is a key goal and would likely revolutionise many areas of science and industry, there may equally be unpredicted spin-off technologies that arise from the ground-breaking physics being undertaken. A member of Professor Rahman’s team, Dr Edyta Osika, notes, “At the time classical computing emerged, those developing the technology had no idea of the transformational effect it would have on almost every aspect of our lives. Similarly, it will be exciting to see how quantum computing develops and the unanticipated ways it may be used to improve our society.”

Researchers from across the country are undertaking world-leading materials science supported by NCI’s robust and powerful computing infrastructure. This diverse work has implications across sectors and society, driving the innovations that will support a more sustainable, healthier and prosperous future for Australia and the world.

National Benefits

NCI enables computational and data-intensive science that helps ensure Australia’s continued prosperity and the welfare of the Australian people. The development of new materials and technologies leads to the emergence of new industries, creating more skilled jobs for Australians. Advances in weather modelling provide more accurate forecasting of extreme weather events, enabling better mitigation and planning efforts. Machine learning techniques facilitate the development of autonomous robots and better monitoring of the environment, helping to maintain Australia’s competitive advantage in agricultural industries. The innovations resulting from the diverse research undertaken at NCI drive positive change across sectors, from public health and medicine to environmental protection, transport and manufacturing.

In 2021–22, the Australasian Leadership Computing Grants, designed to use the full scale and capability of the Gadi supercomputer, provided an opportunity for Australian researchers to pursue ambitious computational research. Across diverse fields, these projects highlighted what is possible when world-class infrastructure is combined with the capabilities of Australia’s foremost researchers. The incredibly competitive process led to 5 successful projects sharing in a total allocation of 130 million units of computing time on the nation-leading supercomputer Gadi.

Led by some of the most experienced computational scientists in Australia, these projects were large and ambitious, enabling our researchers to compete at an international scale and research topics of national significance. Professor Ben Corry from The Australian National University investigated how COVID-19 is recognised by humans to initiate an immune response. He and his team showed how mutations in the specialised receptor proteins that recognise viral RNA can either inhibit or activate the receptor. The significant computing allocation meant that protein-RNA binding could be estimated with accuracy, providing valuable information that could inform future therapeutic approaches, including mRNA vaccine design.

In the past year, NCI continued to expand its offering of computational and data services to the Australian research community, building user capabilities and ensuring the greatest possible benefit is derived from Australia’s research funding investments. This includes an enhanced training offering, with over 106 sessions conducted over 2021–22 collectively involving 3,600 attendees from more than 100 organisations (see *Building User Capabilities* on page 40 for more information). NCI also continues to respond to the specific needs of various research communities, developing and making available tailored data analysis environments, improving the user experience and, ultimately, research outcomes.

Closely integrated supercomputing and big data systems, powerful analysis environments, expert support and a focus on capability development: NCI delivers the key components that drive discovery and high-impact Australian science efficiently and effectively.

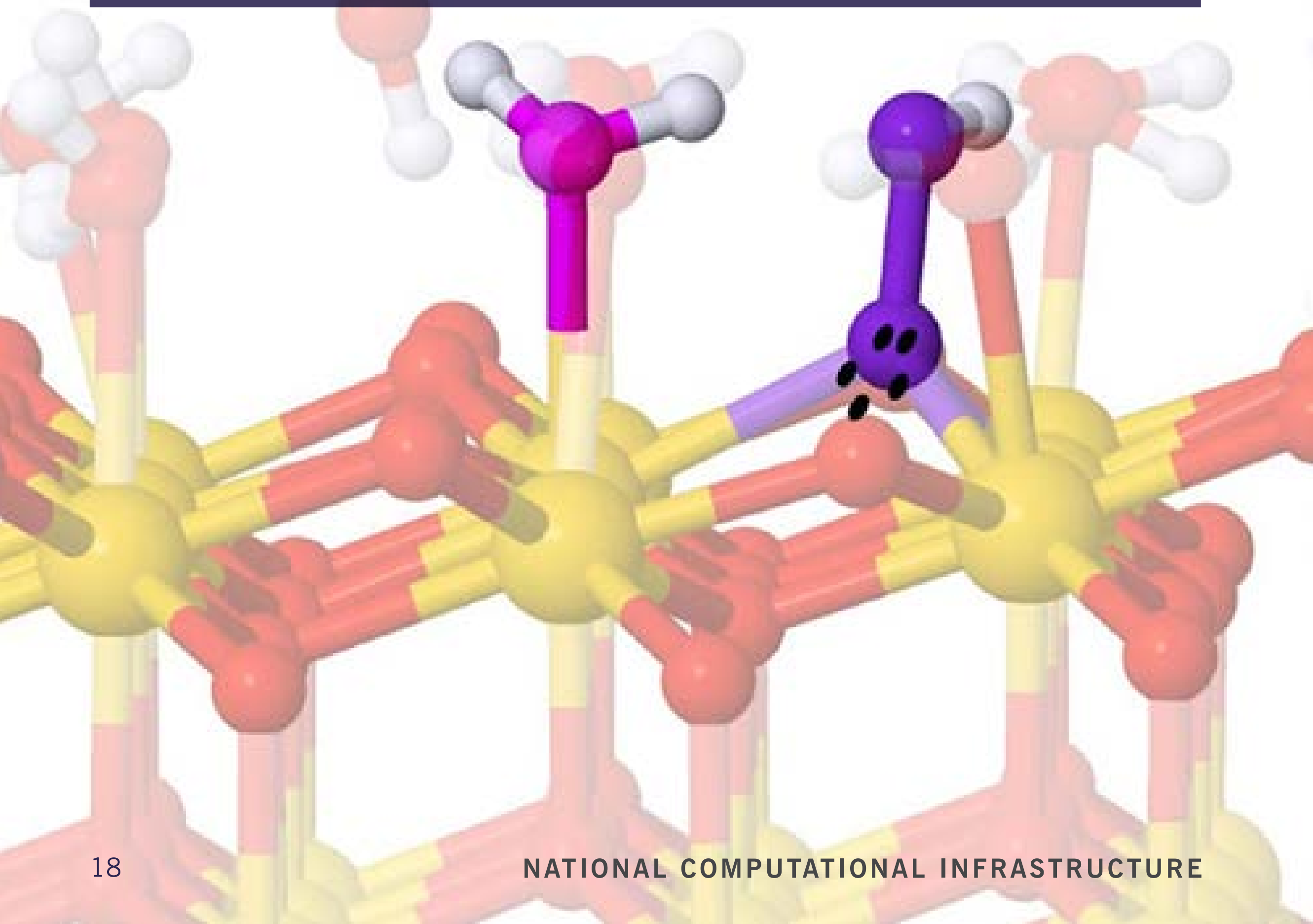
UNDERSTANDING ATOMIC INTERACTIONS TO INFORM CANCER THERAPY DEVELOPMENT

Many promising nanomaterials appear to have qualities that could make them well-suited for use in anti-cancer therapies. The successful development of a candidate therapy can often be hampered, however, by a lack of understanding about its mechanism of action at the atomic level, which is very difficult to determine experimentally. Without a detailed appreciation of how a material may be acting to achieve a certain outcome, it is difficult to harness or further manipulate it for therapeutic purposes.

It is exactly this problem that Dr Judy Hart from UNSW Sydney sought to overcome as part of her 2021 National Computational Merit Allocation Scheme grant. Dr Hart used computing time on NCI's Gadi supercomputer to simulate surface interactions and better understand the action of cerium oxide.

Cerium oxide has been used for decades in a range of industrial processes, including as an effective catalyst to increase the speed of critical chemical reactions. More recently, the potential use of cerium oxide nanoparticles in medicine is being explored, with the exciting possibility that they may be used to selectively target cancer cells.

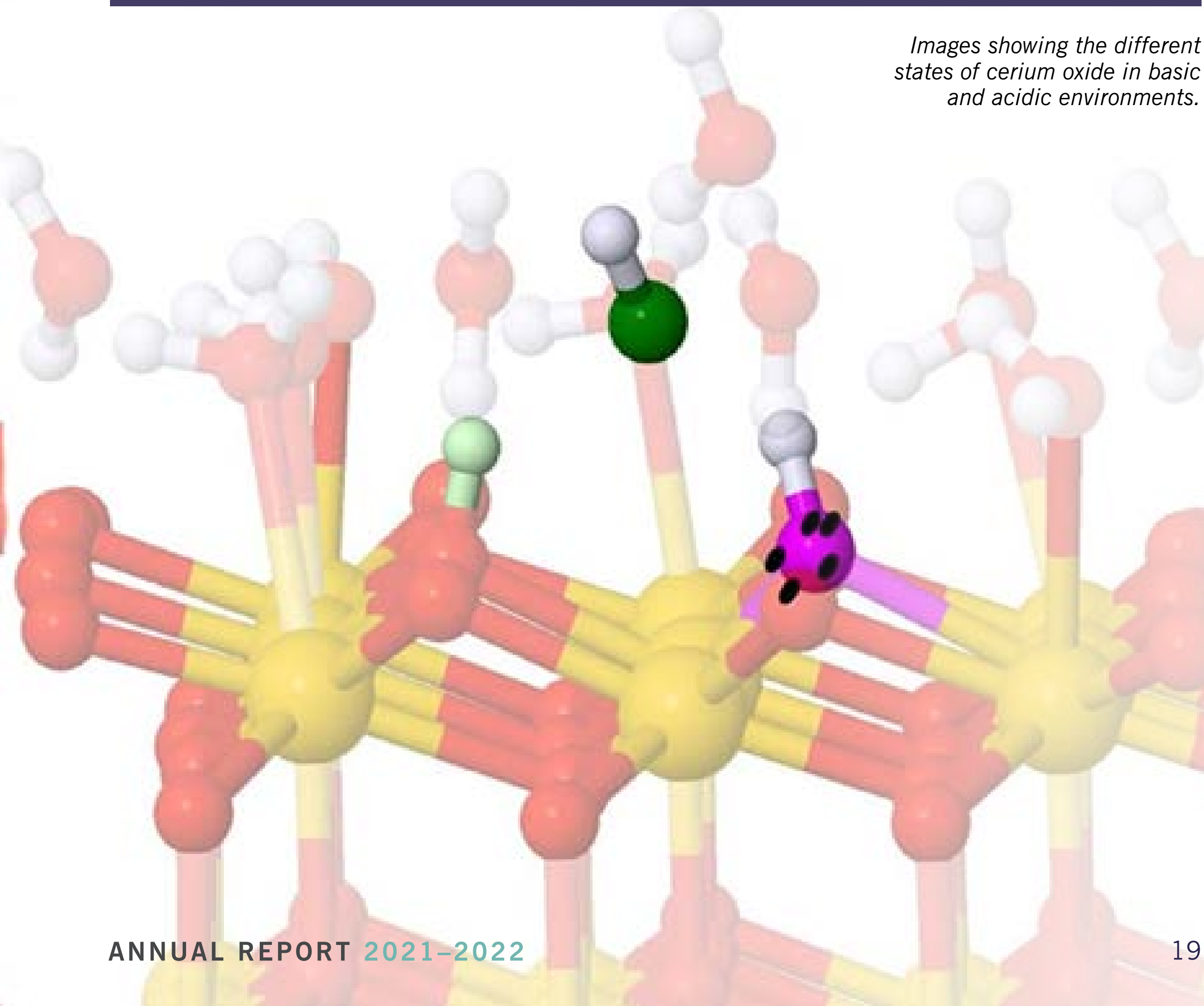
Critically, cerium oxide has been observed to behave differently in particular cellular contexts. "In an alkali or basic environment, typical in healthy cells, cerium oxide acts beneficially, converting potentially damaging reactive oxygen species (ROS) to harmless water and oxygen," Dr Hart explains. "However, in more acidic environments, typical in cancer cells, cerium oxide appears to act in the opposite way, actually driving the production of harmful ROS that could bring about cell death."



The possible therapeutic applications of such pH-dependent behaviour are clear. However, the lack of understanding of the atomic-scale mechanism at play has, to date, hampered exploration of its potential biomedical use.

To gain insights into how cerium oxide achieves its effects, Dr Hart's research team used density functional theory calculations to study the interactions between the surface of cerium oxide and various molecules. Dr Hongyang (Jason) Ma, who worked as a PhD student on the project, says, "These computational approaches allow you to understand what is happening at the atomic scale, to 'see' the interactions between atoms and also control the system, modifying features to investigate how this changes surface chemistry and reactivity." The team's research findings provided evidence for the specific mechanism behind cerium oxide's effects, an important step towards better control of the material's pH-dependent behaviour and realisation of its cancer-treatment potential.

The future promises even greater outcomes from computational chemistry approaches. As the power and speed of supercomputers continue to be enhanced, the ability to model complex systems, more closely representative of reality, also increases. Dr Hart says, "The theory is actually ahead of the computational power available. Today we are able to model 100 atoms using fairly long-standing theory. While incredible insights are possible, this limits the scope of what you can do. In 10 years, with more powerful computing capacity and newer theories that have already been developed, we may be able to model 1000 atoms, and more effectively predict the real-world behaviour of materials."



Images showing the different states of cerium oxide in basic and acidic environments.

Impactful Science

Some of the most rapidly evolving elements of Australian society rely on knowledge generated through computational and data-intensive science at NCI: battery and hydrogen technologies to support a decarbonised electricity grid and future industries, satellite imagery to help farmers and infrastructure managers, and human genomics to support more efficient disease diagnosis and personalised treatment options.

Australia's leading supercomputer and research data repositories at NCI are central to scientific research into weather forecasting, climate change, new materials, health, resources and emerging technologies. With the research tools that NCI makes available, scientists are producing results that create an impact on our everyday lives, enabling new opportunities to be explored and helping us to respond to key challenges.

In recent years, nothing has had a bigger impact on Australians' lives than the coronavirus pandemic. During this time, computational chemists, biologists and public health professionals have been using NCI to investigate every aspect of the COVID-19 virus, including its features and transmission. Through the provision of dedicated resources, as well as ongoing support for impactful projects from around the country, NCI enabled critically important national research. Scientists applied their decades of expertise and all the computational and data analysis tools at their disposal to assist with vaccine development, medical treatment discovery and future pandemic preparations.

Alongside the pandemic, many Australians have experienced, or have been recovering from, extreme weather events that have caused unprecedented human loss and damage in many communities. These events underline the challenges associated with a changing climate, and the need to better understand these changes and how they drive more extreme weather. Models are the main tool for better understanding current climate and weather, as well as more accurately projecting future events. The high-performance computing infrastructure and data services available at NCI underpin the work of Australia's foremost climate and weather science researchers and modellers, including those from the Bureau of Meteorology, CSIRO and the ARC Centre of Excellence for Climate Extremes (see Case Study *Understanding why it rains: From the local to the global* on page 22).

The science taking place at NCI benefits Australians now and into the future. The impact of the discoveries computational scientists are making will continue to grow in coming years. Some of the central technologies and advances of the future, whether they be in weather science, medicine, energy storage, land management or advanced manufacturing, are making their start in the nation's leading supercomputing and big data facility.

UNDERSTANDING WHY IT RAINS: FROM THE LOCAL TO THE GLOBAL

In recent years, Australians are all too familiar with the significant impacts and damage that extreme rainfall can cause. Some regions have experienced devastating floods on more than one occasion, with the human and economic cost of these events challenging the resilience of affected communities, and forcing difficult discussions around how best to manage ongoing risks.

Better understanding the processes underlying these types of events is one of the priorities of the ARC Centre for Excellence for Climate Extremes. The Centre comprises many of Australia's leading climate scientists across 5 universities, including researchers from The University of Melbourne's School of Geography, Earth and Atmospheric Sciences. These researchers are drawing on NCI's computing power and data infrastructure to analyse how the various drivers of rainfall interact, from local factors to global climate and weather processes.

"Rainfall is complex, and understanding it is a multiscale problem," says Dr Claire Vincent, Senior Lecturer at The University of Melbourne and a Chief Investigator at the Centre. "It can be a very local event, so it is necessary to drill down to those small scales. At the same time though, rainfall is strongly influenced by very large scale variability, such as La Niña and El Niño, that is controlled by sea surface temperature in the Pacific Ocean. It is in the space of these interactions that extremes sometimes occur."

This presents a challenge for researchers, with high-resolution models needed over large areas and timeframes, linking up very different scales of variability to see how they control the rainfall climate in various locations. One project, led by Dr Vincent, focused on the occurrence of thunderstorms in the tropics. While these are known to be driven by heating of the earth's surface, which in turn initiates clouds and storms, the project investigated how to better take into account how cloud formation is also affected by planetary-scale waves propagating around the equator.

This is computationally intensive work and reliant on access to large datasets made available at NCI. This infrastructure, together with support offered by the Centre's computational modelling support team, greatly benefits researchers. "It allows you to get on with your research. The computing environment is set up in an optimal way, you do not have to stress about making a compiler work or finding a computer that can manage to run your model. Rather, you can focus on the science, on trying to better understand processes to more meaningfully interpret the data available."

NCI is supporting the foundational science that is improving Australia's ability to predict climate extremes and helping us better prepare for the future.

Merit-based Access to NCI

NCI provides computing resources to the most impressive Australian researchers. Across the merit-based portion of NCI's allocations in 2021–22, NCI allocated more than 580 million units of computing time to researchers from across the scientific spectrum. Peer-reviewed and allocated through a stringent allocation process, the successful proposals are richly deserving for both their scientific excellence and national significance.

In 2021–22, NCI made allocations of computing time under 4 merit allocation streams, supported by investments from the Commonwealth Department of Education's National Collaborative Research Infrastructure Strategy (NCRIS), the Australian Research Council (ARC) and the NCI Collaboration.

The Australasian Leadership Computing Grants (ALCG) cater to the biggest, most ambitious computational science in the country, with 5 projects being awarded 130 million units of computing time over the 2021–22 financial year. Professor Alan E Mark from The University of Queensland's School of Chemistry and Molecular Biosciences, a 2021 ALCG recipient, used atom-based simulation methods to help design more effective organic semiconductors for use in a variety of products, including solar cells and lighting (see Case Study *Insights to improve the manufacture of new materials* on page 26). The 2022–23 ALCG, announced in May 2022, will support 3 projects across diverse science domains including a search for new kinds of insecticides, fundamental investigations of electron-molecule collisions and next-generation low emissions technology.

The National Computational Merit Allocation Scheme (NCMAS) funds the majority of merit-based access to supercomputing resources in Australia, with around 200 recipients each year across the major Australian supercomputing facilities. In 2022, computing time was awarded across

3 facilities: NCI, the Pawsey Supercomputing Research Centre and MASSIVE. Around 300 million units of computing time were awarded to 142 projects on Gadi, covering all domains of science, from turbulence simulations to materials chemistry, genomics, astrophysics, molecular interactions and extreme weather modelling. This represented more than half of the total NCMAS computing allocation on offer.

The Flagship Scheme provides computational resources to certain ARC Centres of Excellence and research institutes, supporting their nationally significant missions. The current Flagship Scheme recipients specialise in astrophysics, climate extremes, materials science, fundamental physics and plant production. Further support goes to research infrastructures in NCRIS and from across the scientific community. These highly meritorious centres and research organisations share in 150 million units of computing time each year with the support of both NCRIS and the ARC.

To complement these 3 existing schemes, and in response to demand from current and emerging NCI users, in late 2021 NCI launched Adapter, a new flexible merit scheme. Adapter is designed to identify meritorious research projects that need flexible compute access over a short period, catering to users who need smaller compute and data resources in short bursts. Following a successful pilot in early 2022, the recipients of the first quarterly application process were announced in June 2022, with research to begin in July. This initial round provides 39 researchers from 16 universities access to around 10 million units of computing time on NCI.

Together, these schemes provide tailored and flexible support to some of the most experienced and important Australian research teams looking into battery technology, climate variability, high-efficiency turbines and much more.

INSIGHTS TO IMPROVE THE MANUFACTURE OF NEW MATERIALS

Organic semiconductors are new materials generating significant excitement among researchers and industry. They are low-cost, lightweight and often mechanically flexible, and have already become key components in lighting, solar cell and micro-circuitry products. However, to fully realise the potential of these materials, it will be necessary to improve the efficiency of their manufacture and develop a better understanding for how arrangements at the nanoscale affect their performance.

As part of his 2021 Australasian Leadership Computing Grant, Professor Alan E. Mark from The University of Queensland ran atomic simulations at NCI to model the manufacture and performance of organic semiconductor materials. The work he and his team have undertaken has helped clarify the cause and mechanism underlying commonly observed issues, including how the performance of organic light emitting diodes (OLEDs) are impacted by the presence of water or oxygen.

OLEDs consist of many layers, including a cathode, an anode and an 'emission layer', the layer that emits light. It has long been observed that even a tiny amount of water can significantly impact the performance of an OLED, with some scientists speculating this may be due to effects on either the cathode or anode layers. By running computationally intensive models simulating a drop of water

placed on an OLED, Professor Mark and team have helped clarify matters, demonstrating that water molecules are actually directly interacting with the OLED's emission layer.

"What we see is that while the OLED layers appear impervious, if a drop of water is placed on top, water molecules will gradually work their way in," says Professor Mark. "We used much of our time on the Gadi supercomputer to investigate how the water diffuses through the various layers. While there have been suggestions regarding how water may affect these materials, we showed for the first time how water manages to get in and move through what is effectively a solid material."

In the simulations, the water 'likes' to be within the emission layer, spontaneously binding to the emitter molecules. This changes the electronic structure and prevents these molecules from emitting light. This knowledge is needed to develop engineering approaches to minimise these effects. "People had been proposing that you may need to make the outer layers of the OLED more waterproof. Our research shows that an alternative may be to chemically modify the emitter itself so that water does not bind to the emitter."

The models developed by Professor Mark and his team are delivering insights to industry that would not be possible through physical experiments alone. These computational approaches, enabled by the analytical power of NCI's Gadi supercomputer, are revealing how the nanoscale structure of organic semiconductor materials affects their properties and manufacture, and laying the foundation for the next advances in manufacturing.

2

A HUB FOR THE RESEARCH COMMUNITY

Snapshot – Helping to Realise the Potential of Genomics

Genomics researchers are unlocking the rich information contained in genetic material, using insights gained to develop new applications across healthcare, agriculture and environmental management fields. Supercomputing and high-performance data services are key to enabling this ground-breaking research, providing the infrastructure and specialised tools needed to effectively manage and analyse the complex and incredibly large datasets generated. NCI is proud to provide this support, collaborating with partner institutions to realise the potential of this emerging and exciting scientific area.

In recent years, NCI has deepened its involvement in human genomics and health research, becoming a trusted partner to Australia's leading researchers. This includes a close partnership with the National Centre for Indigenous Genomics (NCIG), a one-of-a-kind Indigenous-led research organisation. NCI works with NCIG, helping to effectively manage and analyse valuable genomic data and bring donor control and sovereignty over this important resource. In the past year, NCI has further built upon its support for NCIG's critical mission, committing, with The Australian National University, to cover the costs of 10 million units of computing time per year and 500 terabytes of data storage.

NCI pursues collaborations with varied partners to help develop and advance new genomics and bioinformatics frameworks and approaches. This includes hosting and maintaining a node of the Galaxy Australia web platform – through Bioplatforms Australia's BioCommons capability – that provides tools and reference datasets to researchers. NCI also collaborates with organisations, including the Sydney Informatics Hub, to develop and optimise scalable workflows to assist with processing whole genome sequence data. These types of projects are ensuring Australia's medical researchers and clinicians have the resources they require to work efficiently and 'focus on the science' (see Case Study *Helping provide the best tools for children's cancer researchers* on page 56).

Building on this experience, NCI is a foundation partner in the recently announced Precision Oncology Screening Platform Enabling Clinical Trials (PrOSPeCT). NCI looks forward to applying our sovereign data management capabilities, technical excellence and scale to support this potentially game-changing cancer genomics project. PrOSPeCT is being delivered by a consortium of public and private organisations led by the Australian Genomic Cancer Medicine Centre (Omico), and will allow fast screening of cancers to support clinical trials, drug research and development.

> "Omico is delighted to work with Australia's National Computational Infrastructure on the largest cancer precision oncology program, PrOSPeCT. This partnership will change treatment options for more than 20,000 Australians battling incurable cancers."
>
> – Professor David Thomas, head of Genomic Cancer Medicine at the Garvan Institute and CEO of Omico

NCI will play a key role in the project, working with partners to help derive the greatest value possible from the enormous amounts of clinical and genomic patient data collected. By successfully ingesting patient data and transforming it into a format that can be readily used for clinical and drug development applications, we will help create a digital asset that will ultimately deliver benefits for individuals who currently have limited treatment options. This will be achieved through the creation of a genomics database, with applications that can reach into this data, and leveraging the ability to scale-out analytics on NCI's Gadi supercomputer and Nirin cloud. The project will also draw upon the specialised data analytics environments NCI has developed for our users, including web-based notebooks and machine learning and AI libraries.

"It is exciting to play a key role in building a whole new ecosystem that has the potential to deliver so many benefits," says Dr Warren Kaplan, NCI Science Lead (Genomics and Biomedical Data). "It will support the development of treatments for patients who have run out of options, and enable opportunities to build and commercialise more effective drugs, as well as utilise existing drugs in new and innovative ways."

Supporting National Research Infrastructure

NCI is a research infrastructure platform, an investment for the community to underpin nationally significant science. Researchers from across the scientific disciplines access NCI for the analysis-ready datasets, data storage, computational performance, scientific visualisation and expert support we provide. By leveraging the investment of our Collaborators from across the spectrum of the national science agencies, research universities and medical institutes, along with the Australian Government, we support a larger and more capable set of computing and data services than any single institution could build individually.

The Australian Government, through the Department of Education's National Collaborative Research Infrastructure Strategy (NCRIS) program, funds many other research infrastructure facilities, collectively making up the NCRIS network. Through strategic collaborations, NCI compute and data services underpin many of the research tools and data portals that other NCRIS facilities provide their users. By supporting the rest of the national research infrastructure network, we help researchers get easy access to the data collections and robust analysis tools they need.

NCI's variety of services covers the requirements of different scientific disciplines, workflows, code bases and data sources. From the data processing of optical astronomy to the virtual analysis of environmental sensors, NCI helps other capabilities provide their expert scientific platforms in the ways that work best for their users. As digital skills and infrastructure become increasingly important to the work of all NCRIS facilities, NCI's Tier-1 HPCD infrastructure and services are a cost-effective way for the entire research community to access the resources they need.

Whether supporting biodiversity researchers, urban planners, geophysicists or ocean modellers, NCI is the hub for the research of some of Australia's major scientific communities. The Gadi supercomputer's finely tuned performance, the capacity and accessibility of its filesystems, and the availability of virtual analysis environments make NCI a key part of the research community's toolset for Australia's most intensive computational science.

National Research Infrastructure Vignettes

ACCESS NRI:

As a new National Research Infrastructure formally launched in June 2022, the Australian Earth System Simulator (ACCESS-NRI) will further enhance Australia's research modelling capabilities in weather, climate and Earth systems, and facilitate inter-disciplinary collaboration across the Australian climate research community and globally.
NCI supercomputing power has been pivotal in the development of the ACCESS suite of models since their inception and is an integral partner to ACCESS-NRI, providing the high-performance data and computing infrastructure required to deliver cutting-edge simulations and models.

TERRESTRIAL ECOSYSTEM RESEARCH NETWORK:

The Terrestrial Ecosystem Research Network (TERN)'s central goal, to measure and share time series multi-scale observation and model-ready data of key terrestrial ecosystem attributes, requires robust data access and discovery capability. NCI helps to bring the data from hundreds of locations around Australia to the research community through computational and data storage infrastructure, including NCI's FAIR data repository, which underpins TERN's Data Discovery Portal.

AUSTRALIAN RESEARCH DATA COMMONS:

NCI and the Australian Research Data Commons (ARDC) work together to support Australian eResearch activities from across the scientific disciplines. Together, NCI and the ARDC partner to provide enhanced capabilities and future-focused technological developments to research communities, including for climate prediction, weather forecasting and geophysics. Enhanced capabilities include the creation of dedicated Data-enhanced Virtual Laboratories, and the rescue, management and sharing of valuable geophysical data collections.

AUSCOPE:

AuScope provides research tools, data, analytics and support to Australia's geoscience community. AuScope data includes geophysical data that is generated from geophysical sensors located across the Australian continent and enables researchers to understand and predict important natural and human-induced environmental changes and geohazards over time. As part of the 2030 Geophysics Collections Project, a partnership between NCI, the ARDC, AuScope, and TERN, rawer high-resolution versions of these key datasets are being made available for high-performance computing. AuScope is also making NCI's compute and storage resources accessible to the broader geoscience research community, enabling, for example, sophisticated simulations of the Earth's continental processes over time.

ASTRONOMY AUSTRALIA LIMITED:

NCI is supporting Astronomy Australia's Optical Data Centre project as part of a collaboration with Macquarie University and The Australian National University. The Optical Data Centre, a service for holding and serving multiwavelength astronomy data, also supports research teams to carry out their research through simplified data management, team management, user-focused customisable interface and visualisation tools, and data service provisioning. NCI provides robust data storage and management for this new and exciting astronomy research space.

BIOPLATFORMS AUSTRALIA:

Bioplatforms Australia (BPA) provides a wide range of biology and health-related research services. Through BPA's Australian BioCommons capability, NCI hosts and maintains a node of the Galaxy Australia web platform. Galaxy Australia lets researchers conduct accessible, reproducible and transparent computational biology research with more than 1,500 tools, 220 reference datasets and extensive online training materials. To enable the generation of valuable data assets such as reference genome assemblies, the Australian BioCommons and NCI have also collaborated to offer an Australian BioCommons Leadership Share (ABLeS). This provides life science communities with a tailored mix of computational resources and specialist expertise as well as support for a shared repository of tools and software.

AARNET:

AARNet network services connect research facilities across the country to the computing power and data capabilities available at NCI, enabling life-changing research projects. A recent initiative involved AARNet partnering with the University of New South Wales IT team and the Garvan Institute of Medical Research (Garvan) to design and deploy the network infrastructure required for Garvan's human genome sequencing centre, including a high-speed AARNet link connecting Garvan to NCI. The link is enabling faster access to data than ever before for researchers aiming to find cures and improve treatments for COVID-19, cancers, and other diseases.

ATLAS OF LIVING AUSTRALIA:

NCI hosts and makes available high-quality, analysis-ready datasets for the research community, including the Atlas of Living Australia (ALA), Australia's national biodiversity infrastructure. ALA users can browse, analyse and download biodiversity data, including over 100 million species occurrence records, eDNA records, and species information, as well as spatial and environmental layers. Under the surface, NCI provides storage and support for the data portals to run smoothly and efficiently.

Image of a Southern Brown Tree Frog from the Atlas of Living Australia. Credit: Third Silence Nature Photography.

NCI AND PAWSEY SUPERCOMPUTING RESEARCH CENTRE (PAWSEY):

NCI and Pawsey are Australia's 2 Tier-1 supercomputing facilities. The facilities work closely together to enable ground-breaking science across the nation, supporting the computing needs of a large and growing shared user community. This includes jointly administering the National Computational Merit Allocation Scheme (NCMAS), Australia's premier grant scheme for access to high-performance computing resources. Collaboration extends to many other areas, including engaging with the international supercomputing community to continually enhance Australia's capabilities, and co-hosting training events for Australian researchers.

Supporting the NCI Collaboration

Central to NCI's operations is the NCI Collaboration, a grouping of Australia's most prominent national science agencies, research universities and medical research institutes, which provides the funding and leadership that underpin NCI's work and supports that of scientists around the country. The collaborative model allows each Collaborator to contribute to and share in NCI's systems and services, thereby benefiting from a more capable computing and data ecosystem than any individual facility could have built on their own.

The NCI Collaboration consists of the 5 Major Collaborators, a growing group of Minor Collaborators and a selected set of Merit Flagships. The Major Collaborators are CSIRO, the Bureau of Meteorology, Geoscience Australia, The Australian National University and, as of November 2021, UNSW Sydney. NCI Minor Collaborators include Deakin University, the Garvan Institute of Medical

Research, RMIT University, University of Technology Sydney, The University of Wollongong, Intersect, Flinders University and The University of Sydney. Supported by the Department of Education's NCRIS program and the Australian Research Council (ARC), NCI also provides no-cost allocations of computing time to select ARC Centres of Excellence with large computational requirements. A complete list of NCI's Collaborators is on page 82.

The NCI Collaborators conduct some of Australia's most important science. From the Bureau of Meteorology's regular improvements to regional weather forecasts to CSIRO's critical climate modelling activities and Geoscience Australia's mapping of Australia's underground resources, the Collaborators produce high-impact, otherwise impossible science at NCI. Making use of the Gadi supercomputer, high-performance data storage systems, virtual analysis environments and expert support, the NCI Collaboration has access to Australia's most integrated and capable computational science platform.

NCI Collaboration Vignettes

BUREAU OF METEOROLOGY:

The Bureau of Meteorology invests in NCI as part of its research and development process. Gadi and the significant data collections and services at NCI are an ideal environment with which to develop and test improvements to the national weather forecasting software suite, ACCESS. Recent achievements include significant efficiency gains enabling better resolution and accuracy further into the future, and major improvements to both capital city forecasts and tropical weather forecasts in particular. The Bureau is also a key contributor to the Australian Earth System Simulator (ACCESS-NRI) with which NCI is closely engaged.

THE AUSTRALIAN NATIONAL UNIVERSITY:

ANU researchers have access to Gadi and all of NCI's data services for their high-impact science. These capabilities are supporting research across diverse fields, from the earth sciences and evolutionary biology, to immunology and astrophysics. This includes initiatives like the SkyMapper project that provides an invaluable service to the astronomy research community. Led by ANU, the project's data collection is a unique set of images of the Southern Sky that provide rich insight into some of the fundamental questions about the universe. Innovations like Skymapper are enabling fundamental science questions to be answered, as well as ensuring we can respond to societal challenges. ANU researchers are driving advances in areas such as natural language processing, Indigenous genomics and renewable energy that are directly enhancing Australia's ability to achieve its public policy objectives.

CSIRO:

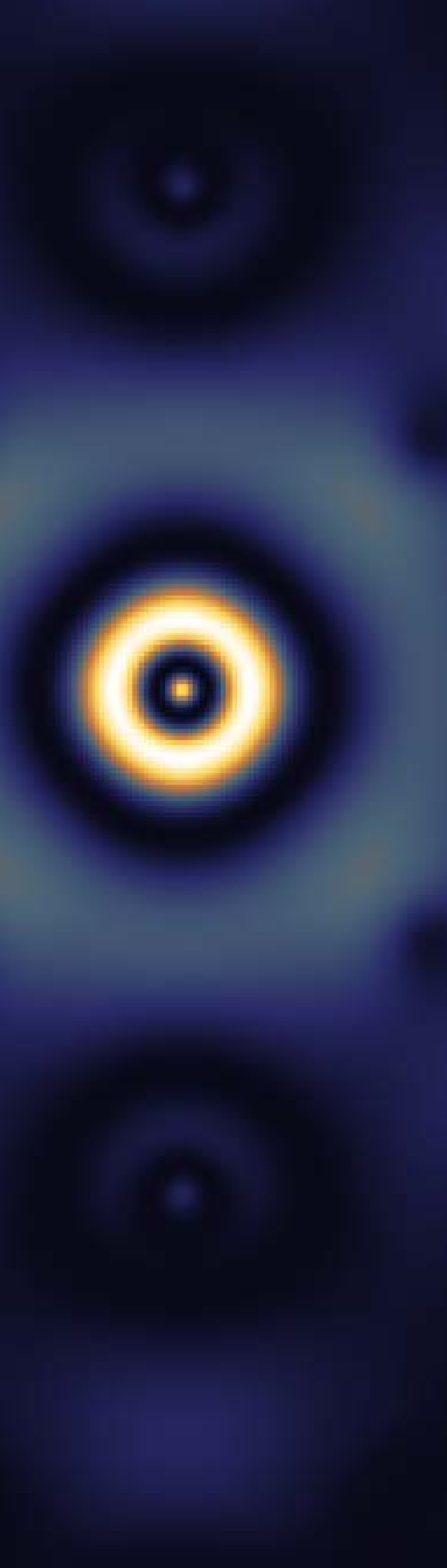
Australia's national science agency, CSIRO, performs climate and weather modelling over many timescales and resolutions. Recently, scientists Richard Matear, Simon Marsland and their research teams have been conducting climate studies at both long-term and decadal timescales. Incredibly computationally intensive and producing huge amounts of data, high-resolution simulations such as these rely on access to a high-performance supercomputer and filesystem, an ultra-fast network and optimised software environment, all of which Gadi provides researchers. Some of this work was delivered through the National Environmental Science Program, delivering global simulations that explore future possible climates. This was part of a much larger international project coordinated by the World Climate Research Program. These model simulations contribute to our understanding of possible future climates, and were recently assessed as part of the Intergovernmental Panel on Climate Change (IPCC) Sixth Assessment Report.

GEOSCIENCE AUSTRALIA:

Learning about Australia's deserts, forests, oceans, rivers, cities and underground resources requires large datasets and rigorous processing. Geoscience Australia (GA) uses NCI as a trusted repository and analysis platform for many of its research data collections, comprising daily updated collections of images of the Australian surface taken from a range of satellites in space, derived data products, direct measurements of ocean temperature and colour and much more. The Earth observation data collections that GA uses are stored, quality-controlled and made easily available at NCI. NCI has developed tools such as the GSKY and On Demand data services to enable easier access to these valuable data collections.

UNSW SYDNEY:

Demonstrating its commitment to computing and data pathways for Australian research, UNSW joined NCI's Collaboration Agreement as a Major Collaborator in late 2021. This enables UNSW to contribute to the strategic leadership of NCI, including representation on the NCI Advisory Board. UNSW researchers undertake a range of computational science using NCI's Gadi supercomputer and data infrastructure, including bushfire modelling, computational chemistry, quantum computing and more. Over the coming years, NCI looks forward to strengthening its connection with UNSW and pursuing further collaborations that harness UNSW's relationship with other NCRIS-supported facilities.



Associate Professor Rajib Rahman from UNSW Sydney is undertaking detailed simulations to support the further development of silicon-based quantum computing in Australia. This visualisation shows the probability density of an electron bound to a single phosphorus donor impurity in silicon crystal.

Building User Capabilities

NCI has made a significant investment in its training offering over the last 18 months, guided by feedback from our stakeholders and users showing there was strong demand for greater skills and capability development opportunities. In doing so, NCI is playing a key role in equipping more Australian researchers to use computational approaches in their analyses, as well as deepening the capabilities of existing computational scientists to pursue new methods of analysis in their work.

Today, NCI's training program is a comprehensive offering tailored to both new and existing researchers, with a diverse spread of courses ranging from introductory high-performance computing (HPC) skills to developing advanced, domain-specific capabilities.

Stakeholder feedback indicated there was a particular need for fundamental command line and programming skills. To offer introductory courses quickly and effectively, NCI has partnered with Intersect Australia, a national eResearch support agency. These fundamental programming courses, offering some of the core skills of supercomputing and data analysis, have been a great success, and a variety of these training courses are now offered on a regular basis. To date, 675 attendees have benefited from this Skill Sharpening program, drawn from more than 40 institutions across Australia.

"The introductory courses NCI offers in partnership with Intersect provide an inroad for different cohorts of researchers to accelerate their transition to computational platforms and frameworks."

– Dr Tracy Chew, Bioinformatics Group Lead from The University of Sydney

In parallel to delivering courses focused on core skills, NCI has been developing more tailored training courses in partnership with academics and computational professionals on a number of topics of growing interest to the research community. Artificial intelligence (AI) and machine learning (ML) courses have been developed to showcase the possibilities of this technology. These courses, known as the AI/ML Applications on Gadi series, advise researchers on how to implement these techniques in their respective science domains. In addition, NCI partners with industry collaborators and subject matter organisations, such as NVIDIA and the HPC-AI Council, to facilitate other learning activities. This approach has created hands-on and easy-to-adapt educational pathways for science students and researchers to innovate using AI and ML, fostering collaboration and an inclusive open-source community.

A wide array of training courses and workshops are offered on a regular basis, covering an ever-growing range of topics including introductions to the Gadi supercomputer and the NCI cloud-based Australian Research Environment (ARE), parallel programming, AI and ML, code profiling and data-intensive analytics. In addition to leveraging external expertise, the development and delivery of these courses draw heavily on the knowledge of NCI expert staff. There is a strong focus on applications, with training designed so that users can easily take what they have learned and apply it to their existing workflows running on NCI computing platforms. This ensures researchers experience immediate benefits, including from improved performance of their code and higher accuracy within unchanged resources.

NCI will continue to consult with the Australian research community to ensure our training offering remains aligned to users' needs, raising awareness and building practical knowledge of new technologies to advance science and research. For more details on NCI's training offering over 2021–22, see *Training* on page 74.

3 BUILDING NEW TECHNOLOGIES FOR AUSTRALIAN SCIENCE

Snapshot – New Services Enabled by the NCI Cloud

Officially launched in April 2022 and made possible through the NCI Collaboration, the Nirin cloud computing platform enables interactive workflows and data analysis capabilities. This NCI cloud offering is providing invaluable tools and flexibility to the research community, ensuring the greatest benefit can be derived from NCI's supercomputing capabilities.

Meaning 'edge' in the language of the Wiradjuri people, the name Nirin was selected in consultation with local Ngunnawal and Wiradjuri Elder Aunty Matilda House. It is a fitting title, given the cloud platform is one of the most directly accessible elements of NCI computational capabilities.

Cloud computing refers to on-demand computing platforms that provide access to computing, storage and network capabilities, without the need for direct active management by the user. The Nirin cloud offers a range of compute, memory and storage capabilities suited to a variety of research workflows. Users can optimise their cloud compute environment to meet their research needs, with connections to NCI's global filesystems and scientific data collections, and a window into the Gadi supercomputer.

"The Nirin cloud supports flexible and exciting data analysis workflows and interactive computing tasks. We are pleased to be able to offer Australia's leading researchers more options for their computational and data science projects."

NCI Director
Professor Sean Smith

Nirin is available to all existing NCI users with current compute allocations. It provides support for those aspects of research computing that are better suited to a rapid-response environment. As an Infrastructure as a Service (IaaS) platform with dynamic compute and storage, Nirin importantly offers the ability to scale up and down as needed. The OpenStack cloud management framework provides flexibility, enabling us to optimise the number of physical servers required at any time. This means that energy consumption is reduced in quiet times, without losing the ability to respond to increasing demand in busier periods. NCI and our stakeholders both benefit: more effective distribution of workloads improves productivity and leads to a more sustainable use of NCI's varied infrastructure.

Nirin's novel architecture further demonstrates NCI's commitment to both sustainability and deriving the greatest benefit from the investments of the Australian Government and NCI Collaborators. Much of NCI's previous Raijin supercomputer was repurposed within the Nirin cloud. This technological second life underpins Nirin's massive new capability. This innovative approach demonstrates not only NCI staff members' software reconfiguration capabilities, but also their fundamental hardware engineering expertise.

Nirin provides the foundation for many of NCI's interactive data analysis environments and Virtual Laboratories, and is used extensively for the processing and preparation of data used on Gadi. Significantly, it is enabling more flexible use of NCI systems and lowering barriers to accessing NCI services, including via the recently announced Australian Research Environment (ARE).

"NCI's cloud services generate significant value for its users. Being able to easily access and analyse almost real-time data opens the door to innovation, providing the tools researchers need to develop solutions to tricky research and policy problems," says Dr Balthasar Indermuehle, Senior Experimental Scientist, CSIRO (see Case Study *Flexible cloud computing supporting Australian radio astronomy* on page 52).

A screenshot of the Drishti application running in the Australian Research Environment at NCI, showing a high-resolution, 3-dimensional scan of an ant.

ARE is a web-based graphical interface based on the Open OnDemand application developed by the Ohio Supercomputer Centre. It offers users access to the power of NCI's world-class research HPC capabilities, within the familiarity of their regular desktop or laptop and without the need to install additional tools and utilities. A range of applications are included, including graphical Virtual Desktops, JupyterLab and the Gadi terminal, with more to be offered over time.

Nirin provides researchers with more options to access and benefit from the computational capabilities NCI offers, from tailored environments with domain-specific tools to platforms focused on ease of useability. This is helping to facilitate innovative computational and data science projects, and extending the range of what is possible in Australian scientific research.

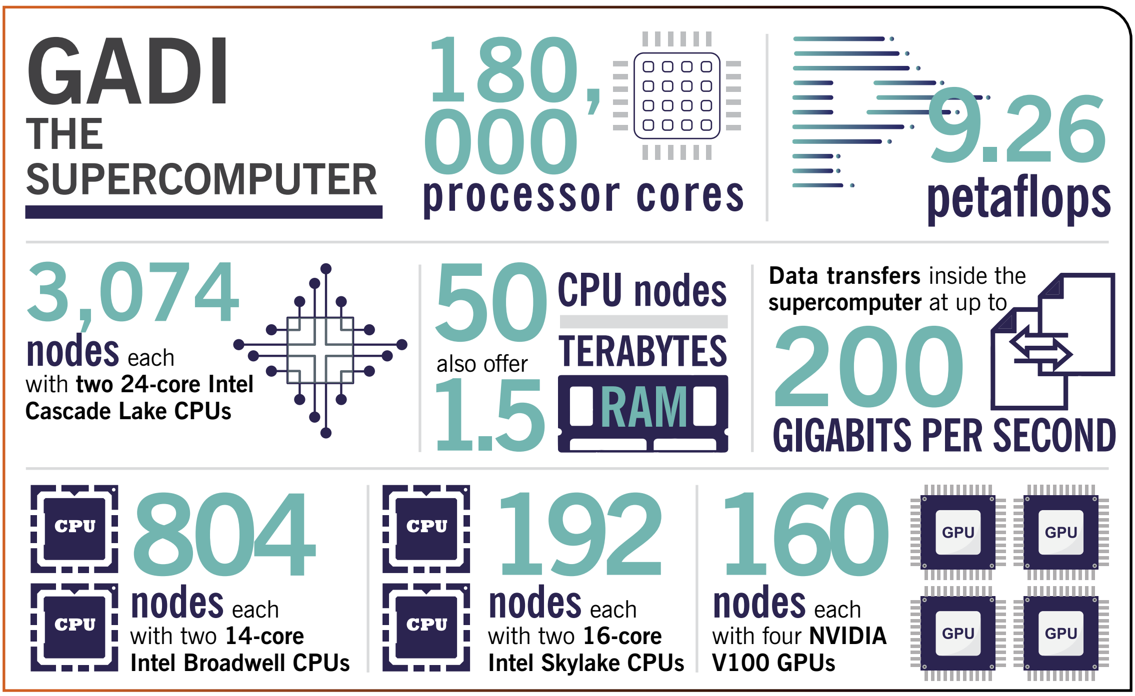
The Gadi Supercomputer and its Filesystems

The Gadi supercomputer is the most powerful supercomputer in the Southern Hemisphere, supporting more than 6,000 researchers and underpinning NCI's diverse data and compute-intensive services. A powerful combination of Central Processing Units (CPUs) and Graphics Processing Unit (GPU) accelerators permits the efficient running of next-generation data-intensive research and software, and allows Australian researchers to run simulations of some of the most complex natural phenomena (see Case Study *Simulating the biggest explosions in the universe* on page 48). First launched in late 2019, Gadi's significant capabilities will be further expanded as part of a mid-cycle upgrade in 2022–23, supported by the Australian Government and the NCI Collaboration.

NCI continually optimises and improves our infrastructure and systems, to streamline users' research experience and respond to the specific needs of particular research communities. This includes incorporating new high-performance computing technologies into our systems as they become available, allowing users to enjoy the benefits of new hardware innovations while also providing insights that inform future NCI investments. Over the past year, NCI made DGX nodes, GPU servers from NVIDIA's advanced technology program, available to our users, providing researchers with optimised performance for the particularly GPU-intensive tasks of artificial intelligence and machine learning workflows.

An ongoing NCI priority is ensuring operating systems and filesystems are available to meet the diverse and evolving computational needs of our users. The new IO Intensive Filesystem gives Gadi users better performance with lower latency than other storage options.

This system has already improved workflow efficiency and permitted new computational methods. It has proved especially useful in research areas such as genomics, where many small files need to be read and written in quick succession. Over the past year, NCI also undertook a successful upgrade of operating systems, moving from CentOS to Rocky Linux. This is now one of the world's largest installations of Rocky Linux. This change to the underlying operating system will help maintain the stability and security of NCI's systems over coming years.



In 2021–2022

Total number of projects 1,000+

Total number of users 6,000+

Total computing time 1.2 BILLION HOURS

For a full list of all compute allocations on NCI in 2021–22, visit www.nci.org.au/about-us/annual-reports

New innovations are also enabling users to manage their data in more sophisticated ways. Facilitated by Australia's Academic and Research Network (AARNet), the Globus data transfer system permits high-speed data transfers on a continental scale. It is now being used by key NCI stakeholders including CSIRO and the National Supercomputing Centre Singapore to transfer data efficiently and securely between systems and institutions. NCI also continues to enhance our high-performance data storage systems, with the 20 petabyte Gdata5 running since late 2021 and now fully allocated, and Gdata6 shortly to be installed. This storage supports both the Gadi supercomputer and Nirin cloud platform, ensuring flexible access to valuable data collections by NCI users, whatever the system they are using.

SIMULATING THE BIGGEST EXPLOSIONS IN THE UNIVERSE

Astrophysics researchers are using the power of supercomputers to better understand what drives supernovae, the explosions of dying stars and some of the most violent events observed in the universe. In doing so, they are uncovering the secrets of how these explosions create the chemical elements that support life and influence the evolution of the cosmos.

When a certain type of star runs out of fuel, the ashes of nuclear fusion pile up to form an iron core at the centre of the star. Eventually the iron core becomes so heavy that the star cannot withstand its own gravitational force and it rapidly collapses, resulting in a giant explosion: a supernova.

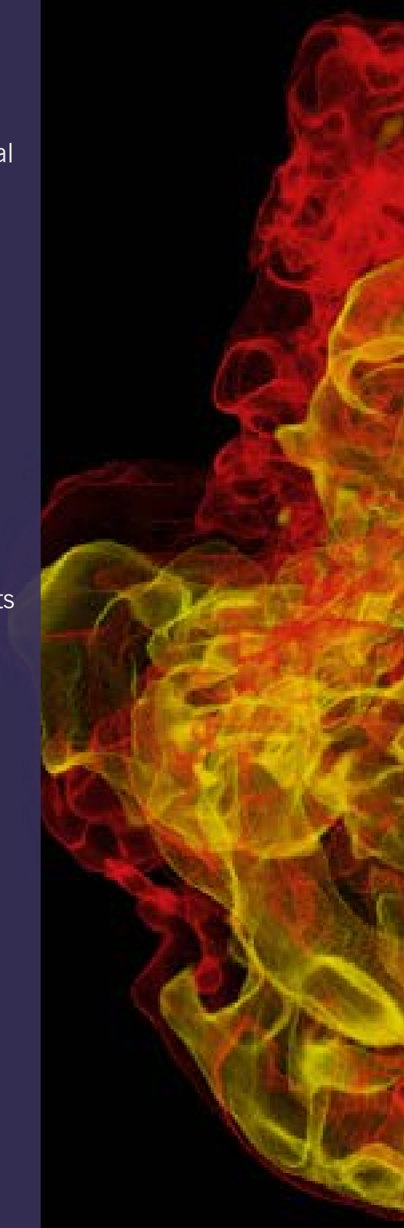
Understanding how this phenomenon occurs is a big challenge for scientists. It is a complicated problem, crossing many areas of physics. While scientists have developed computational models that simulate these events, the size of simulated explosions are only half that observed in actual supernovae. Improved models are needed to account for this 'missing energy'.

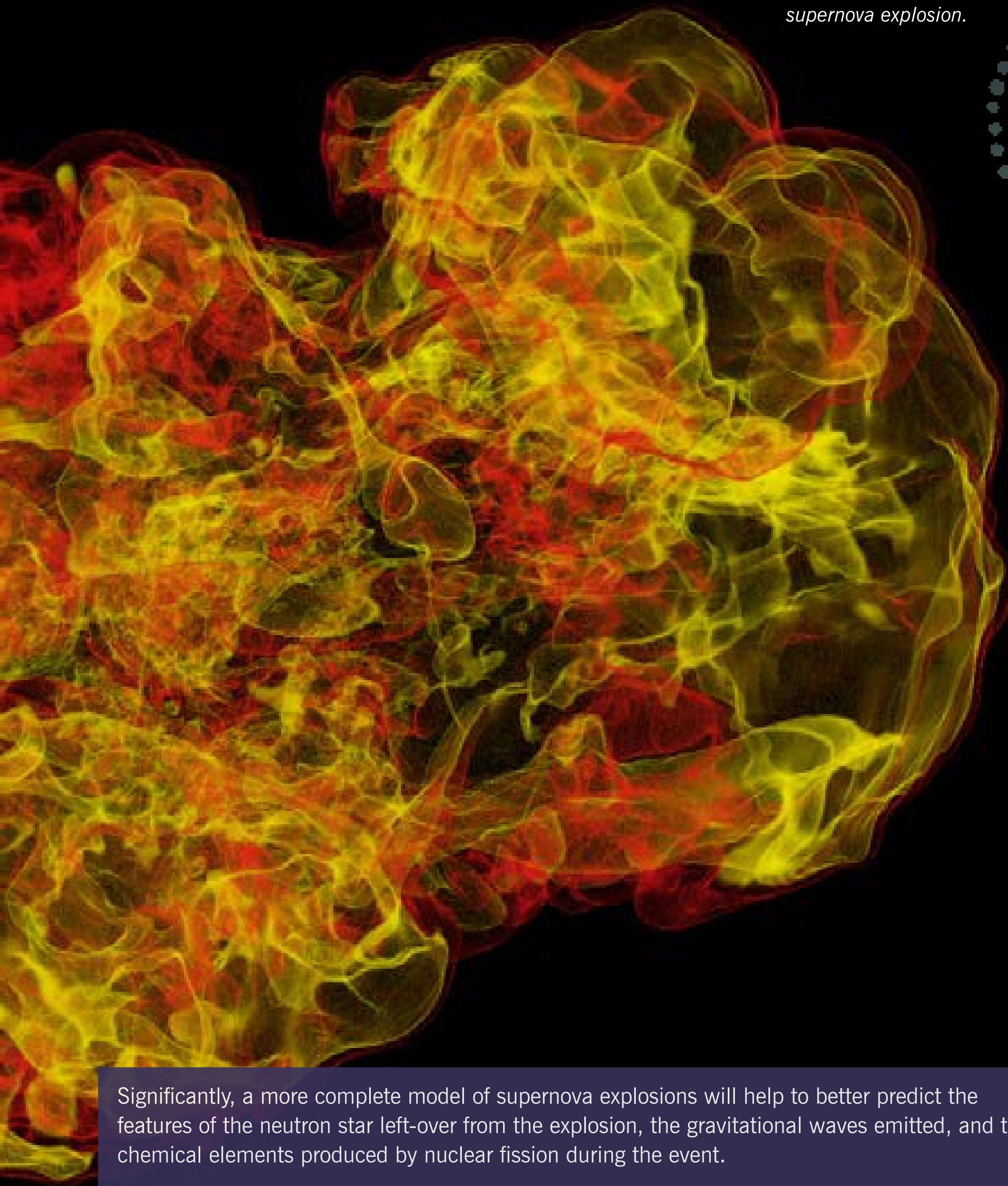
Associate Professor Bernhard Mueller, a former ARC Future Fellow at Monash University's School of Physics and Astronomy, was granted computing time under the 2021 Australasian Leadership Computing Grants (ALCG) to undertake research focused on solving this puzzle. Specifically, his research aims to more accurately model 3 key components of supernova explosions: turbulent fluid motions, magnetic fields and general relativity.

"The ALCG grant allowed us to conduct simulations at sufficiently high resolution to better capture the growth of the magnetic field," he says. "Without the resources from the ALCG grant, it would not have been possible to properly capture these effects, which are driven by amplification processes on small spatial scales."

The simulations run were very large, collectively using 4,000 cores on Gadi over approximately 6 months. Given this scale, the models required continuous monitoring to ensure the project was on schedule, and to manage the 60 terabytes of data produced in the simulations. Analysis and visualisation were conducted on Gadi, with Python-based visualisation used interactively in the Jupyter Notebook environment in the NCI Nirin cloud.

While the immense volume of data produced from the simulations continues to be analysed, an exciting initial finding is that when running the high-resolution model with magnetic fields, the simulation reached a typical supernova energy, whereas in other models a less powerful explosion was observed. This provides evidence of the importance of magnetic field effects that until now have not been a focus in modelling of supernovae.





Simulation showing hot bubbles of expanding plasma behind the shock wave of a supernova explosion.

Significantly, a more complete model of supernova explosions will help to better predict the features of the neutron star left-over from the explosion, the gravitational waves emitted, and the chemical elements produced by nuclear fission during the event.

“Supernovae play a key role in the evolution of the cosmos, responsible for enriching the galaxy with elements such as oxygen, silicon, calcium and iron,” Associate Professor Mueller explains. “They are nuclear furnaces that produce the very building-blocks of our own bodies and many of the technologies we build. Better understanding this phenomenon will enable us to not only understand the origin of the materials that make us up, but also improve the theoretical basis on which we can interpret other phenomena, such as gravitational waves.”

A portion of the Nirin cloud computer at NCI.

The Nirin Cloud Computer

The Nirin cloud provides valuable interactive computing capabilities to NCI users, spread across high-availability and high-capacity computing zones. All of Nirin is closely integrated with the Gadi supercomputer and NCI's multi-petabyte national research data collection. This infrastructure enables the most efficient management of Gadi's load and opens up more ways for researchers to access NCI compute and data resources.

Nirin permits users to establish a tailored compute environment to meet their particular research needs. Users are able to choose the CPUs, memory and disk size they require, and can simply configure and manage this via the Nirin Cloud Dashboard. There is a high-speed connection available to NCI's global data storage, and frequent snapshots of the system allow running environments to be saved and restored as needed.

Further to underpinning a range of user services, Nirin cloud infrastructure also supports a range of internal NCI functions, including data publishing and data management environments. These background tasks are invisible to users while underpinning many key NCI activities.

Nirin comprises a mix of processors and GPUs including:

> 1,856 Intel Broadwell CPU cores with 22 TB memory in a high-availability zone

> 16,640 Intel Sandy Bridge CPU cores with 32 TB memory and 40 NVIDIA K80 GPUs in a high-capacity zone

> 6 PB High-speed Ceph storage.

For further information, please see *Snapshot – New Services Enabled by the NCI Cloud* on page 44.

Nirin artwork by Anthony Best (Canberra-based Indigenous artist). Commissioned for the introduction of NCI's latest cloud computing system, this artwork shows lines of communication, meeting at gathering places represented by the circular features. The arch shapes represent scientific communities researching in the Nirin cloud surrounded by data.

FLEXIBLE CLOUD COMPUTING SUPPORTING AUSTRALIAN RADIO ASTRONOMY

Researchers from CSIRO, Australia's national science agency, are using NCI's Nirin cloud computer to enhance the operation of the ASKAP radio telescope and help protect valuable national science infrastructure.

Located on Wajarri Yamatji Country at the CSIRO's Murchison Radio-astronomy Observatory (MRO) in Western Australia, ASKAP is one of the best instruments in the world for mapping the sky at radio wavelengths. With its wide field of view and powerful survey capability, the telescope is enabling research into some of the big questions in astronomy, such as the origin of mysterious bright pulses of radio waves known as fast radio bursts. Processing and analysis for the influx of data is done in real time at the Pawsey Supercomputing Research Centre in Perth.

The MRO is ideal for radio astronomy for many reasons, including that its remote location means there are extremely low levels of radio-frequency interference from human sources. There are, however, occasions mostly during the summer months, when atmospheric conditions lead to distant human-made radio signals being detected at the site, in a process known as tropospheric ducting.

Dr Balthasar Indermuehle has been working to better understand and model these ducting events, so that they can be more accurately predicted and factored into ASKAP telescope operations. To do this, he requires high cadence data from the meteorological satellites observing Australia. "I was very glad NCI could offer this data in almost real-time and at no additional cost," he says. "Furthermore, rather than having to copy the data to our servers, causing additional latency, I discovered I could use NCI's cloud infrastructure to process the data at NCI itself. This has proved very convenient and greatly supported my research: I can install the operating systems and development and processing frameworks I'm familiar with, and save a lot of time in the process."

Significantly, the NCI Cloud also enables CSIRO to better protect the major research infrastructure located at the MRO. As the site is operated remotely, accurate information on the local weather situation is critical to inform decisions around how assets are managed. However, the absence of local weather observers or radar stations covering the area means that reliable now-casts are not available. Instead, the ASKAP team is now accessing near real-time satellite data products it creates on the NCI cloud. In combination with a range of other measures, this is enabling a very effective response to severe weather systems, such as automatic stowing of telescopes when thunderstorms develop or when frontal systems sweep across the lands.

Support for ASKAP demonstrates the enormous value created by strong and effective partnerships between Australia's leading research and supercomputing facilities. CSIRO researchers are drawing on Australia's 2 Tier-1 HPC facilities in diverse and flexible ways to enable ground-breaking science with ASKAP, from using the Pawsey Supercomputing Research Centre's capabilities to store and analyse the vast astronomical data collected, to drawing on NCI's cloud infrastructure and data collections to optimise facility processes and enhance the protection of assets. Together, Australia's HPC facilities are supporting a strong and innovative Australian research community.

CSIRO's ASKAP radio telescope on Wajarri Yamatji Country in Western Australia is designed to survey the sky faster, and in more detail, than ever before. Credit: CSIRO/A. Cherney

DNA genotyping and sequencing – Photo by National Cancer Institute on Unsplash.

Computational Science Enhancements

NCI's expertise in high-performance computing (HPC) and data innovation is enabling the uplift and modernisation of HPC and data software and methods, and improving the performance and efficiency of scientific model codes and software systems. Improvements to the codes that underpin the biggest computational science projects extend the potential outcomes and impact of the research. Our code porting, optimisation and modernisation activities cover the range of the most important climate and ocean models, and extend to genomics, geophysics, fluid dynamics and astrophysics codes, as well as underlying tools beneficial to all disciplines.

These capabilities are needed to take full advantage of NCI's world-class HPC and data infrastructure, with our experts improving codes and deploying tightly integrated systems to deliver maximum performance advancements. NCI staff help research groups understand the bottlenecks and inefficiencies in their code, as well as plan for future growth in resolution or scale. The solutions developed save time and computational resources, and lead to better scientific outcomes.

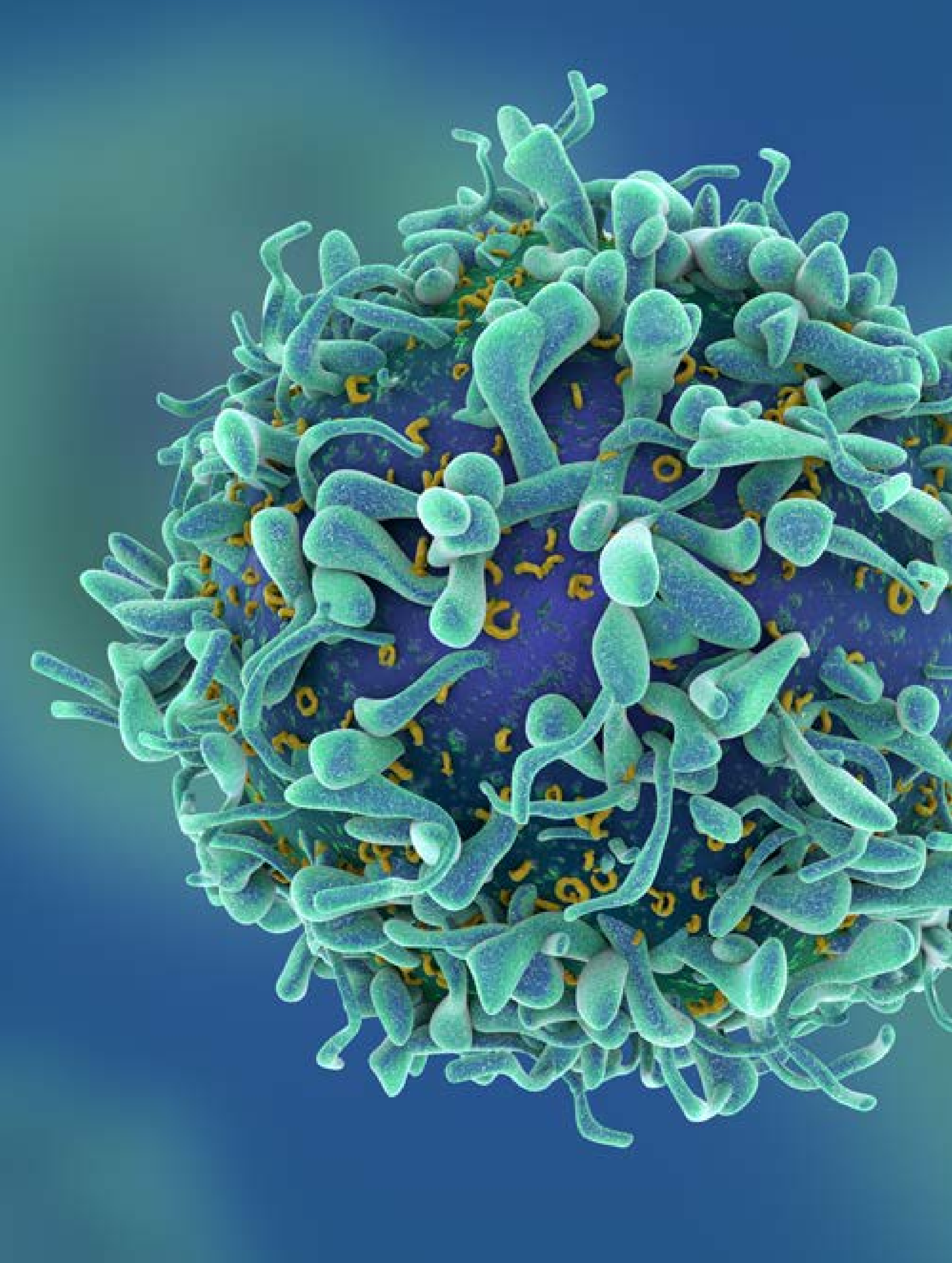
Building on our long history of supporting Australia's weather and climate models to utilise our HPC systems, NCI continues to work with users and partners to ensure relevant software infrastructure and computing environments are continually enhanced. Significant work has been undertaken to successfully transition researchers' models to NCI's systems, enabling scientists to run more efficient applications and modules on our computing systems and online cloud platforms. Furthermore, workflow systems such as Rose-cylc for our coupled weather models help organise and schedule research models, to ensure tasks are effectively coordinated to obtain the outputs desired.

Over the past year, this enhanced software infrastructure underpinned a collaboration with the National Supercomputing Centre (NSCC) Singapore (see *Snapshot – NCI Regional Partnerships and Leadership* on page 68). Importantly, insights and knowledge gained through the project are informing how other NCI users transition to newer climate and weather analysis environments, including researchers from the Bureau of Meteorology. This will help achieve more efficient and controlled running of varied climate models, including those that ultimately inform the deliberations of the Intergovernmental Panel on Climate Change.

NCI is also working to enhance code and analysis environments in many other research domains. In the bioinformatics space, this has included working with users and collaborators at Australian BioCommons to optimise genomic workflows (see Case Study *Helping provide the best tools for children's cancer researchers* on page 56), in addition to developing community-driven command-line tools to assist with the installation of software containers.

As part of the Area 3 computational chemistry porting project, NCI was approached by Fujitsu to port applications to the company's A64FX Processor used in massively parallel computing. Major computational chemistry codes such as Gaussian, Amber, GAMESS and VASP were included. Project outputs will ultimately be used on one of the most powerful supercomputers in the world, the Fugaku supercomputer at the Riken Center for Computational Science in Japan, demonstrating the contribution NCI's expert staff make to the international supercomputing community.

NCI's focus on code optimisation helps create the HPC workflows that lead to national impact and significant scientific advances. Code optimisation produces more efficient jobs, enabling increased resolution and decreased turn-around times on models. Software enhancements ensure the most benefit is derived for NCI users and the community from the massive supercomputing capability available to them.



HELPING PROVIDE THE BEST TOOLS FOR CHILDREN'S CANCER RESEARCHERS

World-class genomics infrastructure is needed to support Australian scientists and clinicians looking to develop novel treatments to improve outcomes for children with cancer. An important element is the availability of genomic data and reference genomes that are well-maintained and up-to-date. Such resources provide a strong foundation for international research collaborations, enabling results to be readily compared and resources used efficiently.

In a project funded through the Australian BioCommons Leadership Share (ABLeS), NCI staff have worked with bioinformaticians from Children's Cancer Institute (CCI) to align their data to the current benchmark reference genome, called hg38. "The update of our genomic data was very important, as over time using out-of-date reference libraries can hamper research outcomes," says Dr Marie Wong-Erasmus, Precision Medicine Informatics Manager at CCI. "We drew on the expertise of NCI staff, who worked with us to develop a tailored pipeline suited to running on NCI's Gadi supercomputer. Furthermore, the project benefitted from NCI's significant computing infrastructure, utilising the I/O Intensive file system and drawing on almost 800 terabytes of storage to temporarily hold our data."

An Australian BioCommons program, ABLeS is designed to build the capacity of life science communities to construct, maintain and gain insights from data assets by providing access to infrastructure and computational resources. In addition, ABLeS communities are offered specialist expertise and centrally supported tools and software, tailored to support the generation of valuable data assets, including reference genome assemblies. The program supports computation at Australia's 2 Tier-1 supercomputing facilities, NCI and the Pawsey Supercomputing Research Centre.

The project delivered the required updated data extremely quickly, with completion over only 2 months. This has helped position Children's Cancer Institute to successfully scale the Zero Childhood Cancer National Precision Medicine Program and associated research, to make comprehensive genomic profiling available to all children in Australia diagnosed with cancer. With more robust data and genomic analyses, scientists and clinicians will be able to better understand the unique biology of the disease and more effectively tailor treatments to suit individual patients.

Curated Data Collections

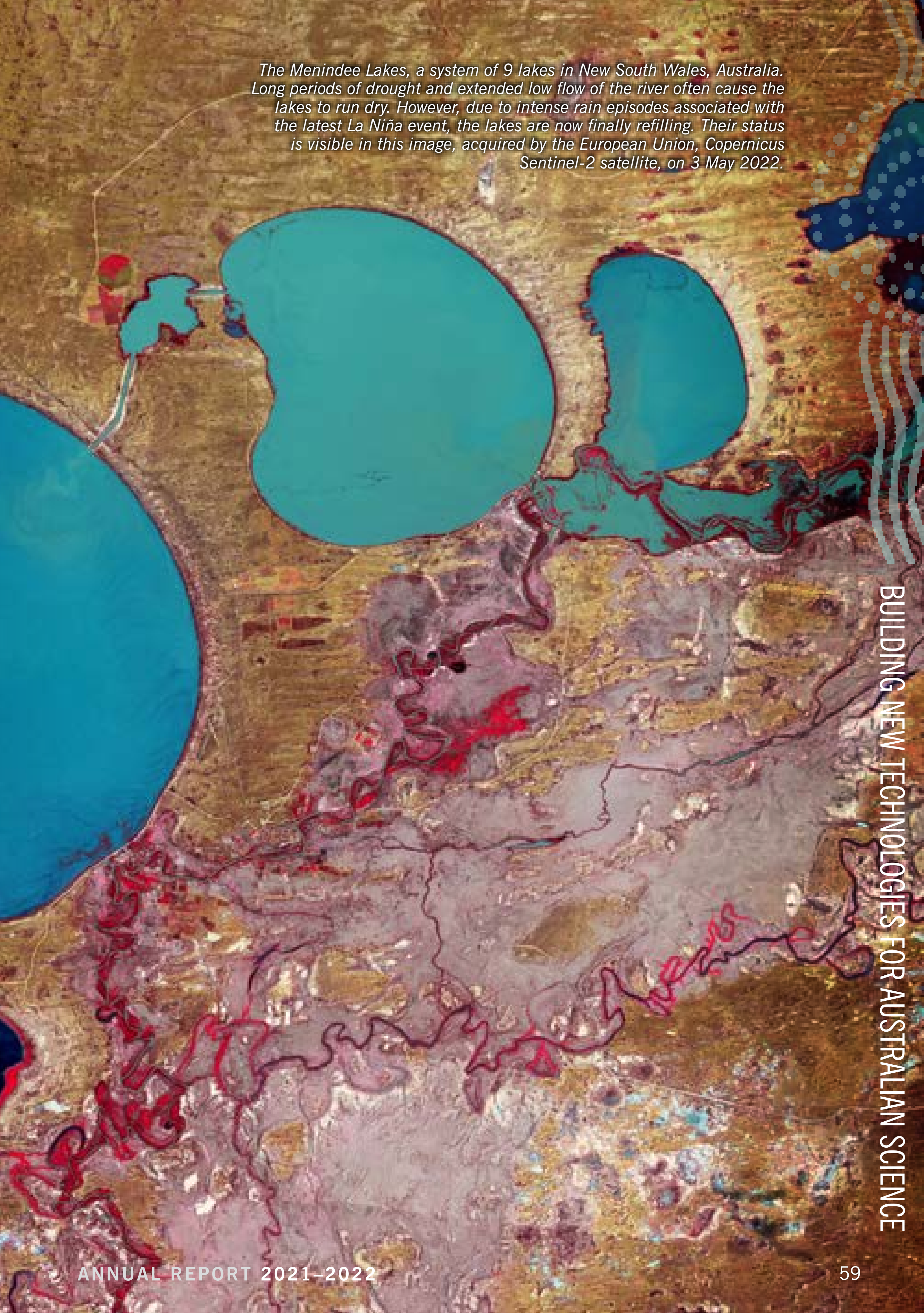
To compete and collaborate internationally, scientists in all disciplines are increasingly reliant on large, distributed data collections for use in their highly specialised workflows. NCI provides this data and all its underlying infrastructure, collating and curating high-quality, analysis-ready datasets (see Case Study *Providing the data infrastructure that Australia's climate scientists need* on page 60). This opens up new research opportunities for scientists, who can focus on their work while NCI maintains the collections in a research-ready state.

We have constantly been adding to our data collections' functionality to help users better access, analyse and share the data they need every day. A core aim is making sure that the data stored at NCI is findable, accessible, interoperable and reusable (FAIR). This has been supported through NCI's partnership with the Australian Research Data Commons as part of the Data Retention Project. The project is delivering significant improvements to the quality of data collections, with controlled and consistent structural metadata driving the FAIR data principles.

All the significant national data collections at NCI, in total making up more than 10 petabytes of data, meet the international FAIR data standard. Maintaining the datasets entrusted to us by the research community and the national science agencies for broad scientific use is a central facet of our data collections work. We are constantly working with our data providers to make sure that their data collections are up-to-date and accurate.

NCI curates a varied collection of valued and well-used data resources that are available to both active NCI users (those with an NCI account) and the broader research community. Over 2021–2022, several additional datasets and releases were made available. The Medical Genome Reference Bank Collection became an NCI-funded national data collection, recognising the important role this dataset plays in assisting researchers to discover new genetic variants underpinning disease, and clinicians to accurately diagnose patients. Other key releases included Bureau of Meteorology satellite products, the Bureau of Meteorology Atmospheric high-resolution Regional Reanalysis for Australia (BARRA), the Bureau of Meteorology Australian Water Outlook Service Data Collection and ANUClimate 2.0.

Improving the availability and state of data is helping researchers extract the relevant pieces of information they require and find creative connections between them, providing a firm foundation for innovative science.



The Menindee Lakes, a system of 9 lakes in New South Wales, Australia. Long periods of drought and extended low flow of the river often cause the lakes to run dry. However, due to intense rain episodes associated with the latest La Niña event, the lakes are now finally refilling. Their status is visible in this image, acquired by the European Union, Copernicus Sentinel-2 satellite, on 3 May 2022.

PROVIDING THE DATA INFRASTRUCTURE THAT AUSTRALIA'S CLIMATE SCIENTISTS NEED

Australian climate researchers are drawing on the rich data collections available at NCI to improve our understanding of droughts and floods.

A central collection is the European Centre for Medium-Range Weather Forecasts' (ECMWF) Reanalysis version 5 (ERA5), a comprehensive reanalysis of weather models and observational data from ground and satellite sensors, covering the period from 1950 to the present. ERA5 includes detailed estimations of air pressure, windspeed, temperature and other climate variables, providing a high-resolution multi-faceted overview of the Earth's weather and climate over the last 70 years.

To unlock ERA5's full potential for climate research, local access adjacent to high-performance computing and data analysis systems is required. With supercomputing infrastructure and robust data services, NCI is the ideal location for Australian researchers to collaborate on and analyse ERA5 data.

Dr Chiara Holgate from the Research School of Earth Sciences at The Australian National University is using ERA5 to investigate how different weather systems contribute to extreme rainfall and the role this may play in breaking droughts. The work builds on previous studies she undertook at the Australian Bureau of Meteorology on the link between drought occurrence in Eastern Australia and the extent of ocean evaporation taking place in the Tasman Sea.

"ERA5 is a great resource for researchers to access reanalysis data from a source that is quality-controlled and easily available," she says. "I use it as the basis for analysis I undertake on Gadi on how the transport of moisture in the atmosphere varies according to major synoptic systems, across both time and space."

Dr Holgate considers the combination of centrally located data, analysis environments and high-performance computing available at NCI critical to her research. "If NCI did not host ERA5, it would be very hard for me to undertake this research. Having everything available in a central location is key, as it means I do not have to copy over significant quantities of data. Further, NCI provides data analysis environments, such as Jupyter notebooks in the Pangeo environment, where I can easily bring in data as needed from other sources, including on cyclone tracking and rainfall observations."

As a national hub for data and compute-intensive research, NCI increases the productivity of the research community by making available robust, high-performance analysis tools with access to some of Australia's most significant data collections.

Data Science Enhancements

An enormous amount of data is being created and made available across scientific disciplines, unlocking new fields of inquiry to researchers and enabling novel approaches to addressing long-standing challenges. However, to support researchers to realise the potential of this data trove, efficient means of analysing and visualising this data are required. This includes multipurpose tools, useful to researchers across disciplines, as well as tools optimised to meet the needs of particular scientific domains.

NCI has designed and set-up specialised environments for managing workflows and analysing data, for both interactive and computationally intensive uses. Dedicated environments exist for a range of scientific disciplines, from bioinformatics and genomics, to geophysics and quantum computing. Collectively these specialised environments represent a unique and highly valued component of the services we offer our users.

In the earth sciences area, a specialised analysis environment is ensuring the 'most is made' of the unprecedented data offering arising from the 2030 Geophysics Collections project. The project, a partnership with the Australian Research Data Commons, AuScope and the Terrestrial Ecosystem Research Network, will make available rawer high-resolution versions of key multi-physical datasets for HPC analysis. This includes magnetotelluric and passive seismic survey data. NCI has created a geophysics module to allow researchers to have a more fluid experience for analysing and processing geophysical data, with a variety of pre-installed software and integration of Python, Julia and R environments. Users can access hundreds of pre-built libraries directly, incorporating these into their workflows without having to build their own software environment.

NCI's specialised environments are also helping more researchers pursue emerging methods of analysis, including the rapidly growing data analysis tools of machine learning. A large benchmark dataset is available, comprising 1.2 terabytes, allowing users to test and compare the performance of different algorithms. The environment includes distributed computing frameworks such as Dask and Ray. Researchers are able to run their distributed machine learning code in an accessible way,

utilising multiple CPU and GPU nodes. As a relatively new research area, NCI is also investing heavily in building the skills of current and prospective users so that they can derive the most benefit from these tools and research methods, including through the use of GPU parallelisation tools such as the RAPIDs software suite.

Scientific visualisations are another valuable tool that NCI facilitates, assisting researchers to gain deeper insights into complex datasets and effectively communicate their results. A number of visualisation packages can now be utilised via the ARE, including Drishti and Paraview. Further, NCI's Vizlab team of specialist visualisation programmers works with researchers and in-house computational and data experts to generate images, videos and virtual reality experiences. These products help scientists see their data in new ways (see Case Study *Bringing ocean models to life: Seeing why ice melts in Antarctica* on page 64).

The combination of supercomputing performance and innovative data services connects and enriches the valuable data stored on site at NCI, supporting new methods of analysis, fields of inquiry and richer science outcomes.

Australian Hits
117,277,867

Australian Downloads
3,019,907,879
MEGABYTES

International Hits
1,975,038,860

International Downloads
2,151,076,931
MEGABYTES

For a full list of data downloads from Australia and around the world in 2021–2022, visit www.nci.org.au/about-us/annual-reports.

BRINGING OCEAN MODELS TO LIFE: SEEING WHY ICE MELTS IN ANTARCTICA

Scientists are developing increasingly realistic models to better understand what is driving the melting of ice in Antarctica. Using these models, researchers can better predict the Antarctic contribution to global mean sea levels rise under different scenarios, informing global responses to mitigate and adapt to climate change.

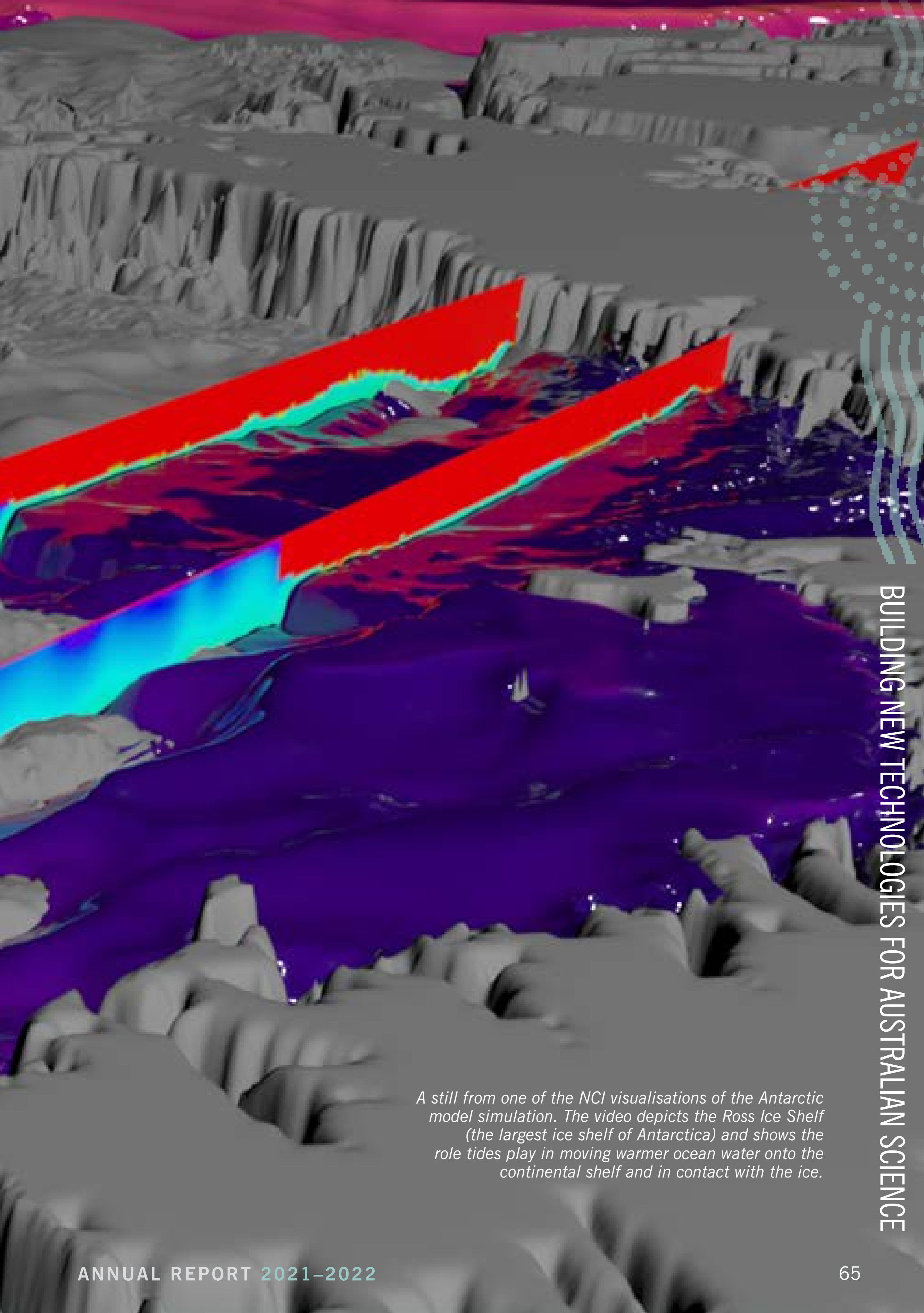
One issue ocean modellers face is finding tools that can effectively capture the outputs of complex simulations, as well as communicate insights gained to diverse audiences. To assist with this challenge, a group of scientists have been collaborating with NCI's Vizlab – an expert team of visualisation programmers – to develop 3-dimensional, dynamic visualisations for their new pan-Antarctic model.

"One of the main benefits of our model is that we are seeking to include the impact of tides across all of Antarctica," says Mr Ole Richter, PhD candidate at the University of Tasmania's Institute for Marine and Antarctic Studies (IMAS). "Tides are comparatively very fast processes and as soon as tidal currents are included, a whole new level of complexity and computational intensity arises. A lot of work is needed to maintain numerical stability in the model. However, these tides have important effects, and including these in simulations will enable scientists to better answer questions and open up new areas of enquiry."

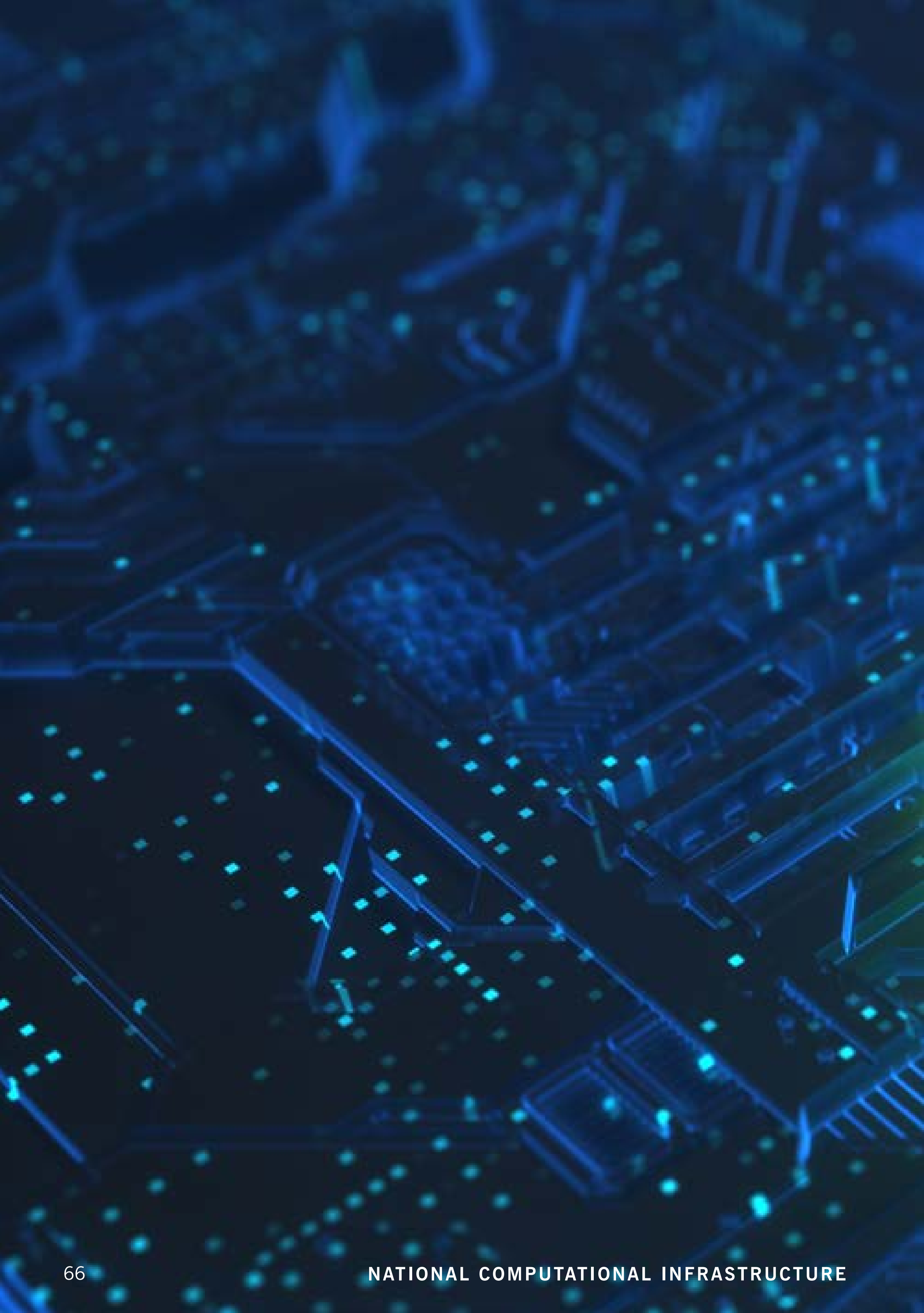
The videos developed with the Vizlab highlight the different processes at play, particularly their varying time scales. "While we can generate lots of plots of simulation data ourselves, these are generally static and 2-dimensional. The videos we made with the Vizlab team show known processes in sharp relief over time, for example the wave-like character of tides circulating Antarctica and the relative speed of tides compared to slower eddies."

By clearly showing how different processes interact at both local and continental scales, the videos are valuable tools for communicating the complex questions that the model can help answer. "Models are long-term projects, the result of different partnerships and the contributions of many scientists. While we focus on sea level rise, other research groups around the world are already picking up what we have created and using this tool to address specific research questions in fields as diverse as biology, sediment transport and water mass transformation."

NCI Vizlab collaborations show the innovative way NCI is working with researchers to create products that extend the science discovery process, and help communicate high-impact results more widely.

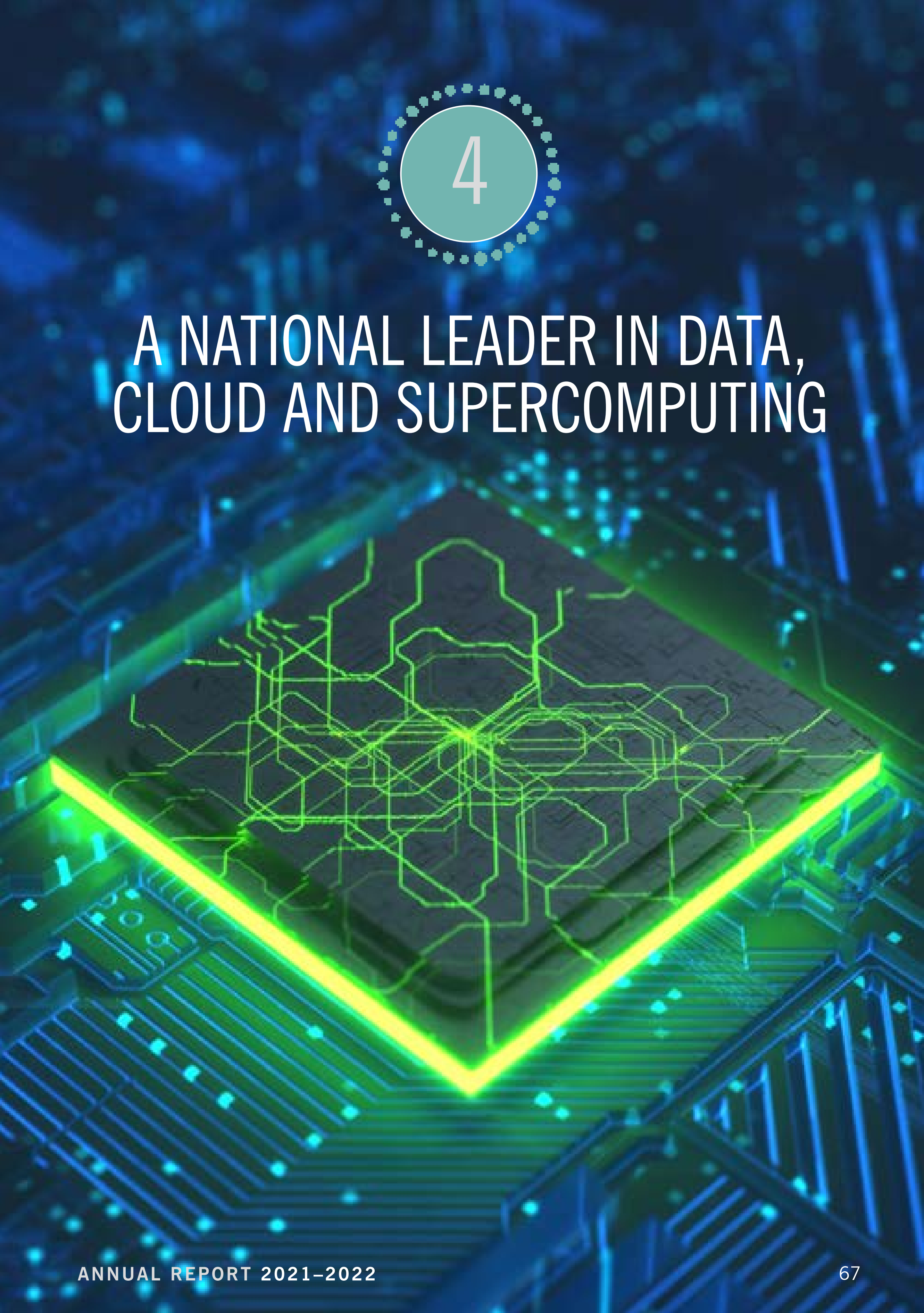


A still from one of the NCI visualisations of the Antarctic model simulation. The video depicts the Ross Ice Shelf (the largest ice shelf of Antarctica) and shows the role tides play in moving warmer ocean water onto the continental shelf and in contact with the ice.



4

A NATIONAL LEADER IN DATA, CLOUD AND SUPERCOMPUTING



Snapshot – NCI Regional Partnerships and Leadership

NCI continues to deepen its relationships with regional counterparts and other stakeholders, helping to further develop the high-performance computing (HPC) community in the Asia-Pacific and establish mutually beneficial partnerships that improve the HPC offering available to users in our region.

In March 2022, NCI and the National Supercomputing Centre Singapore (NSCC) entered into a Memorandum of Understanding (MoU), building on a long history of cooperation. Over many years, the close relationship has enabled each facility to benefit from shared advice and practical support, helping with their respective growth and development. Recently, for example, an agreement between the 2 facilities saw NCI take on some of NSCC's climate science computing load while supercomputer upgrades are completed.

The MoU will deepen joint activities in areas such as technology improvements, software development, staff and user training and data sharing over the next 3 years. NCI and NSCC will also work together to see how each centre's capabilities can be 'geared up' to keep in step with global HPC developments, including the transition to exascale computing. Parts of the MoU will explore collaboration in the area of HPC infrastructure and capability development in areas like green data centre technologies, greater research network connectivity, and more secure data transfer options using quantum encryption technology.

NCI Director Professor Sean Smith said, "This MoU is the next step in a productive working relationship that will grow the capabilities of both centres and lead to advances in computational science software and the technical development of our staff. Even more importantly, users will benefit from improved training materials and future access to novel computing technologies."

The sentiment was shared by NSCC Chief Executive Associate Professor Tan Tin Wee, who said, "This new MoU expands on the already good foundations that the two organisations have laid down and sets our collaboration on a new and exciting path."

Regional conferences also provide a unique opportunity to promote a vibrant HPC ecosystem, fostering exchange of ideas, research results and case studies relevant to all facets of HPC and big data. Supercomputing Asia (SCA) is an annual event co-organised by HPC centres across the region. NCI attended the hybrid SCA conference in 2022 and looks forward to jointly organising SCA 2023 with colleagues from Singapore (NSCC), Thailand (NSTDA Supercomputer Center – Thai SC), Japan (RIKEN Center for Computational Science) and South Korea (RIST).

Another important aspect of NCI's regional relationships is its participation in the Australasian Chapter of Women in High Performance Computing (WHPC). WHPC is a global organisation aiming to better support diversity and inclusion within and across HPC and eResearch sectors. The Australasian WHPC chapter, launched in 2020, continues to grow, supporting networking and knowledge sharing, as well as helping senior level decision-makers adopt practices that support diversity and inclusion within their organisations.

The attendees of the NCI-NSCC MoU signing ceremony, from left: His Excellency Mr Will Hodgman, Associate Professor Tan Tin Wee, Deputy High Commissioner Mr Harold Lee, Professor Keith Nugent and Professor Sean Smith.

NCI continues to look for new opportunities to broaden and deepen our regional ties and is helping lead the formation of the HPC Members' Forum. The Forum is a cooperative venture between several supercomputing facilities, including regionally based centres. An initial focus of the group has been skills uplift strategies and training cooperation, including staff development, and future joint training sessions are planned. Another potential area of cooperation is the development of strategies to attract and retain highly skilled staff with specialist scientific and computing expertise. This reflects the reality that recruitment of talent is a shared challenge across HPC facilities.

By building strong regional relationships and providing HPC leadership, NCI is helping to keep Australia at the forefront of supercomputing, with enormous benefits for Australian researchers and science.

National and International Engagements

NCI is a key pillar of the Australian computational and data science community, with strong partnerships internationally and nationally that help advance Australia's computational science capabilities.

As a critical component of national science infrastructure, we provide the tools that Australian researchers and research institutions rely on for data access and sharing, simulations and modelling. The NCI Collaboration brings together the country's leading research bodies, all contributing funds to support and access the computational infrastructure underpinning their science. NCI Major Collaborators comprises Australia's key national science agencies and leading universities: CSIRO, the Bureau of Meteorology, Geoscience Australia, The Australian National University (ANU) and UNSW. The NCI Collaboration is further strengthened by the contributions of many other universities and research organisations as Minor Collaborators, with new partnerships continuing to form each year.

Funded under the Australian Government's National Collaborative Research Infrastructure Strategy (NCRIS), NCI also works closely with many other NCRIS facilities, providing critical computing and data capabilities that help these organisations deliver on their research missions (see *Supporting National Research Infrastructure* on page 32). This includes membership of key forums and groups such as the National Earth and Environmental Science Facilities Forum and the Digital Data and eResearch Platforms. The Pawsey Supercomputing Research Centre, a fellow NCRIS facility, is our primary technical partner for HPC and high-performance data (HPD), and we collaborate closely on user training, technical development and events.

As a global leader in HPC and HPD, NCI works closely with research organisations from around the world. On the HPC front, NCI is a member of the Accelerated Data Analytics and Computing (ADAC) Institute. Other members include the Swiss National Supercomputing Centre, Oak Ridge National Laboratory in the USA and RIKEN in Japan, all organisations at the pinnacle of extreme-scale supercomputing innovation. Aiming to develop technologies and skills between collaborating organisations, ADAC highlights that the future of supercomputing and big data technologies relies on global cooperation.

The valuable data services NCI offers are a direct result of our strong international linkages. Within the climate research community, NCI plays a central role as one of the 6 nodes worldwide collaborating to share and make available climate data through the Earth Systems Grid Federation. Alongside partners including the US Department of Energy, NASA, the German Climate Research Centre, the UK Centre for Environmental Data Analysis and others, NCI provides access to petabytes of climate model data for the Australian and global research community.

This data is critical for thousands of researchers worldwide and closely informs reports of the Intergovernmental Panel on Climate Change.

NCI also engages closely with the American Geophysical Union and European Geosciences Union, helping guide their annual meetings and contributing to their busy work programs and skills-sharing networks. This work is part of a broad suite of data-related international engagement activities featuring the European Space Agency, the United States Geological Survey, the United Kingdom Meteorological Office and others, that underpin data sharing and scientific outcomes.

Members, allies and friends gathered together for the 13th Annual WHPC Workshop at the International Supercomputing Conference in Hamburg in June 2022.

Diversity and Inclusion

As a prominent member of the Australian scientific and computing community, NCI has an important role to play in increasing diversity across our user base and workforce, and in implementing inclusive practices and procedures across our access schemes.

As part of the WHPC Australasian Chapter, NCI partners with colleagues from the Pawsey Supercomputing Research Centre, the New Zealand eScience Infrastructure, Monash University and Australian eResearch Organisations to facilitate professional development and social events for our community. In 2021–2022, NCI hosted fortnightly meetings to build connections across HPC and HPD centres in our region, and create a safe space to get advice and conversation from peers.

WHPC+ AusNZ was represented by NCI staff at multiple data science and computing conferences in the past year, including at the SCA conference in Singapore in March 2022, the International Supercomputing Conference in Hamburg in June 2022 and at the eResearch Australasia conference in October 2021.

In addition, NCI has continued to work closely with the Office of the Women in STEM Ambassador to implement the National Trial of Anonymising Grant Proposals in the National Computational Merit Allocation Scheme (NCMAS). As Secretariat of NCMAS, NCI has played a central role in redesigning

the application process for the scheme, to remove obvious identifiable applicant details and ensure that the focus is on the proposed scientific projects. The Trial will provide important data to a University of New South Wales study on the effectiveness of anonymising proposals to reduce bias.

As part of our commitment to the advancement of First Nations peoples, NCI has partnered with the ANU to offer, as a complement to the university's existing Reconciliation Action Plan, ongoing, fully guaranteed computing and data storage resources to the National Centre for Indigenous Genomics (NCIG). The storage, processing and analysis of genomic data from the NCIG collection has been taking place at NCI for several years. These new arrangements build on this strong existing relationship, providing direct support for NCIG's critical mission of bringing the benefits of genomic medicine to Indigenous Australians.

These activities and practices highlight the importance of, and benefits associated with, proactively taking steps to achieve greater equity and access to opportunities in the supercomputing community. NCI considers supporting greater diversity and inclusion a key priority and looks forward to building on these efforts in coming years.

Training

Building on existing resources and courses, NCI continues to assist the research community in its aspirational scaling of HPC, AI, and Data Science towards Exascale (see *Building User Capabilities* on page 40). NCI's training program is able to grow and elevate thanks to expert contributions from staff across the organisation. This creates high-quality learning opportunities in various formats to engage diverse user groups at a range of technical levels. Training events are helping to develop and deepen researchers' skills and capabilities across HPC, data science, cloud computing, AI and machine learning, and have been closely informed by conversations with users about where the greatest need exists.

In 2021–22, NCI training courses and events covered a range of topics, including beginner-level introductions to supercomputing and to using NCI data services, as well as more advanced discipline specific cases.

Highlights include the following:

> Digital research training workshops (with Intersect Australia): These workshops, delivered in partnership with Intersect Australia, cover a range of cutting-edge technologies and tools, including Python, R, and Julia programming, machine learning and research computing and parallel programming.

- > NVIDIA Workshops with NCI: These events are hands-on opportunities for NCI users to learn about a range of computational and data science topics, including AI and Gadi's GPU nodes.
- > HPC and Data in Materials Design and Discovery course: This collaborative, graduate-level online course was delivered over 3 months by an array of internationally renowned professors and scientists (facilitated by Intersect Australia).
- > AI and Machine Learning in Drug Design and Discovery Showcase: Consisting of presentations from international and Australian world-leading researchers, this event showcased state-of-the-art research activity in this fast-evolving field.
- > NCI Presents – Mathworks Workshop Series: Offered by NCI and MathWorks, these hands-on workshops focused on machine learning, deep learning with images and parallel computing.
- > NCI Presents – TechTake: This monthly online event provides a forum for international computational and data science leaders to discuss and demonstrate how technology is supporting their research.
- > Introduction to Gadi: This introductory course provides a foundation for those unfamiliar with Gadi and HPC in general.

For a full list of offerings in 2021–22 see *Training Activities* on page 89.

An in-person AI/ML Bootcamp, with group learning, collaborative activities and presentations from NCI expert staff.

Outreach

NCI continues to play an important role in offering diverse groups an introduction to the high-powered infrastructure and services available here. This includes leading researchers, research delegations, parliamentarians, high school and university students and the broader scientific community. We reach hundreds of current and future scientists with in-person and virtual visits of our facility, stalls at school career events, exhibitions at supercomputing and data science conferences, and social media engagement online.

In April 2022, we were honoured to welcome the CEO of Fujitsu Limited, Mr Takahito Tokita, to NCI, in one of his first international trips following the pandemic lockdowns. This visit helped consolidate an important relationship that has been central to the Gadi supercomputer and HPC offering at NCI over many years.

Over the past year, NCI also welcomed representatives from the United States Geological Survey, students as young as 4 from the Radford College Junior School Holiday Program, Indigenous work experience students from the ASTRO3D astronomy immersion program and many more. While providing development opportunities to users alongside training sessions, NCI has continued to host scientific talks from leading researchers and research software engineers around Australia. These TechTake sessions introduce new technologies or research developments to a broader community of scientists.

Higher degree by research (HDR) students from The Australian National University taking a tour of the NCI data centre.

"Thanks so much for your time today and giving the kids a tour of Gadi! You managed to reach a bunch of the kids, in particular the really little ones and at least one of the little girls has said she wants to use Gadi when she grows up!"

Huw Davis, Out of School Hours Care Program Coordinator at Radford College.

NCI's outreach program serves to provide our user community, the general public and everyone in between a look behind the scenes of our nationally significant infrastructure facility. Alongside our training, support and allocation offerings, NCI's outreach activities contribute to business growth and user expertise development.

NCI Director Professor Sean Smith (left), Mr Takahito Tokita, Fujitsu Limited CEO and Mr Graeme Beardsell, CEO of Fujitsu Asia Pacific (right) on a tour of NCI and the Gadi supercomputer.

Financial Report

PREAMBLE

NCI is an organisational unit of ANU. The ANU, as represented by NCI, administers numerous funding contracts that support the operations of NCI. In the interests of providing a comprehensive picture of the NCI operation, a financial report consolidating these funding contracts is presented.

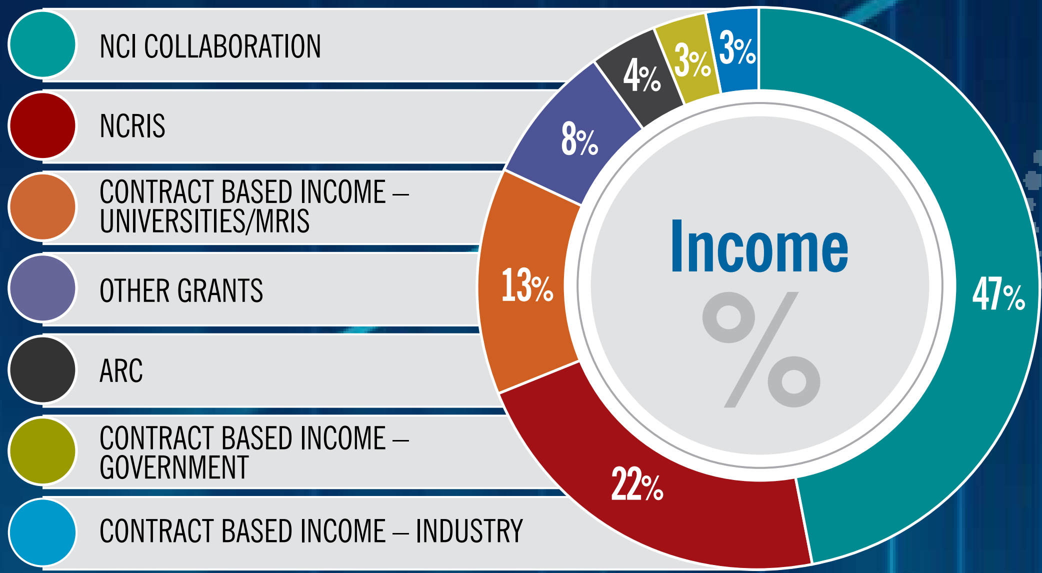
Each funding contract is accounted for in a distinct account within the University ledger, and the University facilitates, and where appropriate, acts on, the NCI Advisory Board's directions and resolutions on NCI matters insofar as they are consistent with the relevant funding contract and not contrary to University statutes and policies.

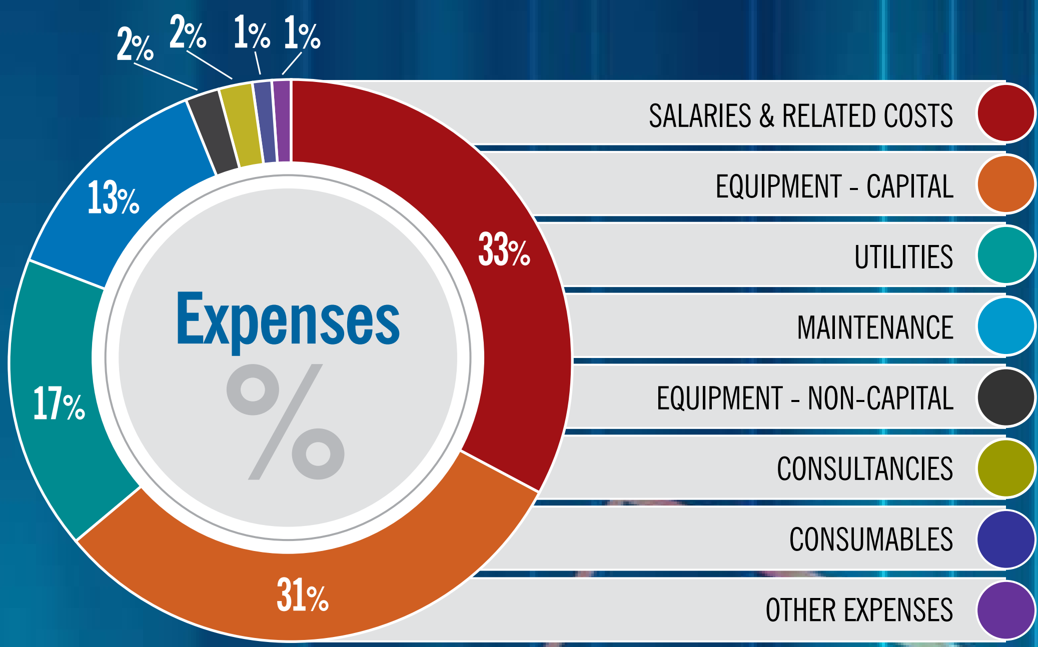
STATEMENT OF INCOME AND EXPENDITURE

For the period 01 July 2021 to 30 June 2022

For the NCI collaboration and associated project accounts

	NCI $
Income	
NCI collaboration income	13,223,191
Other income	14,788,403
Total income	28,011,593
Expenditure	
Salaries and related costs	9,052,741
Equipment - capital	7,532,359
Equipment - non-capital	539,855
Utilities and maintenance	7,807,674
Travel, field and survey expenses	103,990
Expendable research materials	19,777
Contributions	65,000
Consultancies	418,551
Consumables	416,099
Internal purchases	12,117
Other expenses	122,859
Transfers to other	362,000
Total expenditure	26,453,021
Surplus / (Deficit)	1,558,572





A NATIONAL LEADER IN DATA, CLOUD AND SUPERCOMPUTING

NCI COLLABORATION INCOME

The NCI Collaboration Agreement enables many of Australia's leading research-intensive universities and science agencies to collectively fund a capability beyond the capacity of any single institution. The Australian Government, through the Department of Education, together with these institutions (including ANU, CSIRO, Bureau of Meteorology, Geoscience Australia, UNSW, the Australian Research Council, and a range of other research-intensive universities and consortia) funds a significant proportion of NCI's operating costs. Operational funds from the Department of Education are received under the National Collaborative Research Infrastructure Strategy.

A small but growing proportion of NCI Collaboration income comes from the commercial sector, which has a need for NCI services to support innovation and their growing research needs in driving the economy. Under the strategic guidance of the NCI executive and NCI Advisory Board, income that has not been expended in the current year is strategically used in subsequent years to meet contractual obligations and expenditure on emerging infrastructure requirements.

NCI administers a number of grants and contracts outside of the NCI Collaboration and NCRIS accounts. These special purpose arrangements fund clearly defined projects, infrastructure and services that provide synergistic benefits not only to the NCI Collaboration, but also to the wider Australian research community and, through them, the nation.

EXPENSES

NCI, as Australia's national research computing service, provides world-class, high-end services to Australia's researchers. In order to do this, NCI invests significant amounts of money in its expert team of staff and HPC and data storage infrastructure – with these 2 factors driving NCI's expenditure profile.

REVIEW/AUDIT

Each funding contract held by the ANU as represented by NCI has specific financial reporting and auditing requirements, and NCI, in conjunction with the University's Finance and Business Services Division and Corporate Governance and Risk Office, acquits individual project funds in accordance with these requirements.

This consolidated statement has been reviewed by the ANU Finance and Business Services Division and certified that:

The statement accurately summarises the financial records of these grants and that these records have been properly maintained so as to accurately record the Income and Expenditure of these grants.

The NCI Collaboration

SUPPORTED BY

MAJOR COLLABORATORS

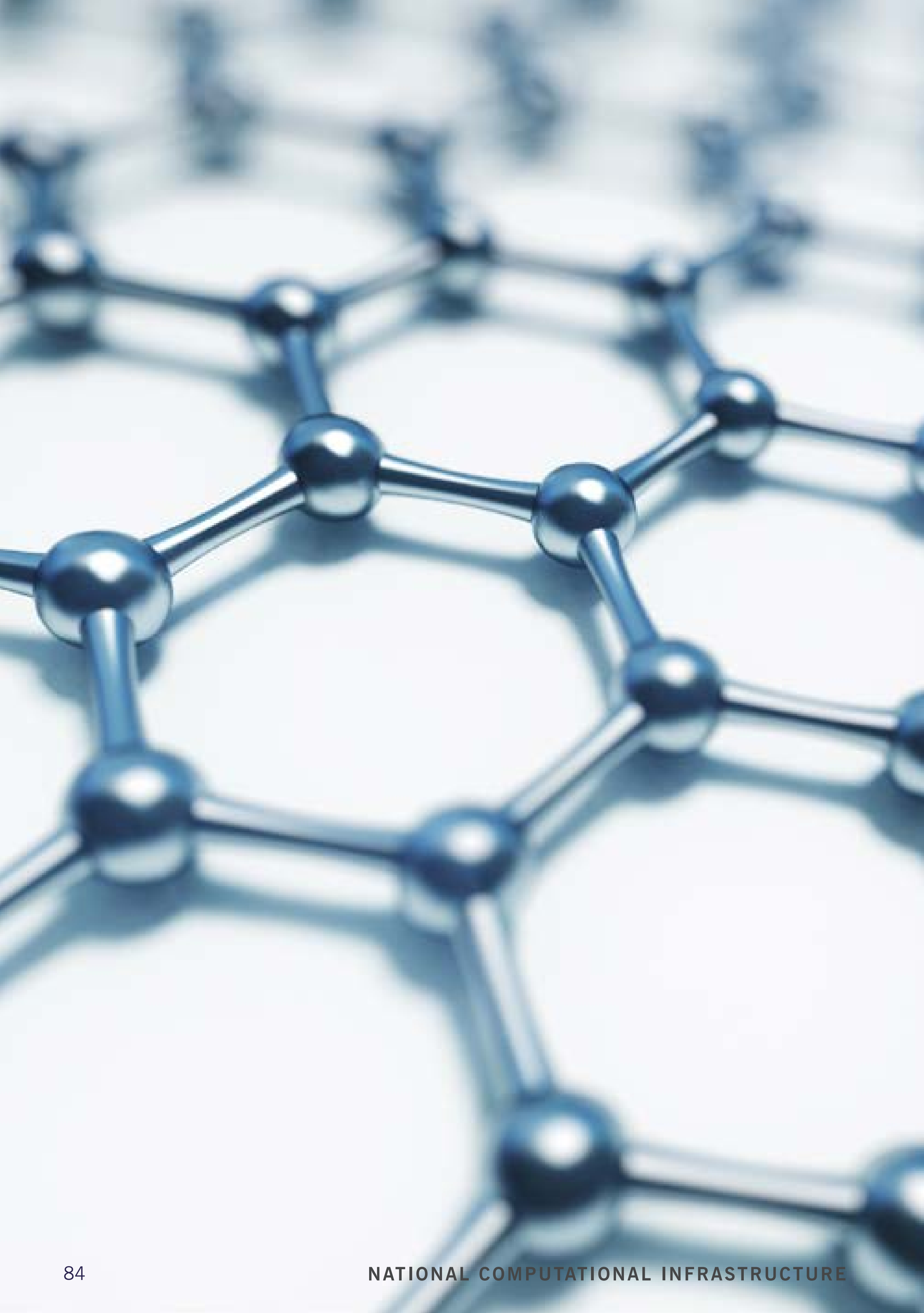
COLLABORATORS

OTHER CONTRACTS

MERIT FLAGSHIPS

Our Vendors

A NATIONAL LEADER IN DATA, CLOUD AND SUPERCOMPUTING



5

APPENDIX

NCI Links to Government Portfolios

The table below highlights, for illustrative purposes, select examples of how Australian Government priorities are underpinned by activities and projects supported by NCI. Many additional links to government priorities exist, including to State and Territory Government policies and programs.

Australian Government Department	Relevant Program/Agency	Activities/Projects supported by NCI
Department of Education	Australian Research Council (ARC)	Dependencies from more than 557 projects funded by ARC's National Competitive Grant Programs (NCGP) comprising in excess of $819m in research investment
	National Collaborative Research Infrastructure Strategy	Support for services provided by numerous other NCRIS Capabilities (AuScope, TERN, ACCESS-NRI, ALA, AARNet,…)
Department of Industry, Science and Resources	CSIRO	Australian Community Climate and Earth System Simulator (ACCESS)
		Climate Systems Hub of the National Environmental Science Program
		Climate and Weather Science Data-Enhanced Virtual Laboratory
		Marine Virtual Laboratory
		AuScope Virtual Research Environments Geoscience Data Enhanced Virtual Laboratory
		CMIP6 Climate Dataset
	CSIRO and the Australian Institute of Marine Science (AIMS)	eReefs
	Geoscience Australia (GA)	Digital Earth Australia
		Copernicus Data Hub
		National Reference Grid
		Exploring For The Future Initiative
		Water Observations from Space (WOfS)
		Australian Natural Hazards Data Archive
		AuScope Virtual Research Environments Geoscience Data Enhanced Virtual Laboratory
Department of Foreign Affairs and Trade	Policy development for, and by, the tourism sector	eReefs (through CSIRO)
		International collaboration (ESGF, ADAC, NeSI, NSCC)

Australian Government Department	Relevant Program/Agency	Activities/Projects supported by NCI
Department of Climate Change, the Environment and Water	National Environmental Science Program (NESP)	Climate Systems Hub
	Environmental policy development	eReefs
		Coupled Model Intercomparison Project (CMIP)
	Bureau of Meteorology	ACCESS
		Climate Systems Hub of the NESP
		Climate and Weather Science Data Enhanced Virtual Laboratory
		Marine Virtual Laboratory
		BARRA Reanalysis
		Bushfire Model Development
	Australian Antarctic Division	Australian Antarctic Program Partnership
		ACCESS Southern Ocean and cryosphere models
	Policy development for the agricultural industry and water resources	ACCESS development with BoM/CSIRO/ACCESS-NRI
		Development and hosting of Digital Earth Australia Collaboration with GA and CSIRO
		Hosting and curation of WOfS data with GA
	Murray-Darling Basin Authority	Exploitation of Digital Earth Australia data
		Development of water-related Digital Earth Australia capabilities in collaboration with GA
Department of Health and Aged Care	National Health and Medical Research Council	Dependencies from more than 37 NHMRC funded projects and fellowships comprising in excess of $52m in research investment
		Australian Genomics Health Alliance (AGHA)
		Support for urgent research related to public health, including COVID-19
Department of Defence	Australian Geospatial Intelligence Organisation (AGIO)	Onshore topographic data and products provided by GA
	Australian Hydrographic Service (AHS)	Raw and processed bathymetric data collections provided by GA
Department of Infrastructure, Transport, Regional Development, Communications and the Arts	Australian Maritime Safety Authority (AMSA)	Managing risks to marine vessels in Australian waters (undertaken with consultant DHI)
	Aviation Programs	Development of specialist weather reporting products for the aviation industry (with BoM)
	Transport Infrastructure Programs	National Reference Grid

Outreach Activities

Event	Date
University of New South Wales Virtual Tour	6 July 2021
WHPC Birds of a Feather Session – Collaborative Conference on Computational and Data Intensive Science (C3DIS)	6 July 2021
Canberra Computer Science Enrichment Tour	13 August 2021
National Science Week NCI Virtual Tour	18 August 2021
SC21 Conference (Virtual)	14-19 November 2021
National Youth Science Forum Digital Tour	20 January 2022
Radford College Junior School Holiday Program Tour	21 January 2022
Science Meets Parliament 2022	28 February - 4 March 2022
SCA: SupercomputingAsia 2022 Conference	1-3 March 2022
Women in High-Performance Computing Birds of a Feather Session	1 March 2022
Professor Alex Brown, ANU and Telethon Kids Institute Professor of Indigenous Genomics,Visit	3 March 2022
Geoscience Australia and United States Geological Survey Delegation Visit	25 March 2022
Fujitsu Japan CEO Visit	6 April 2022
ANU Higher Degree Research Students Visit	8 April 2022
Australian Academy of Science Elizabeth and Frederick White Research Conference NCI Tour	11 April 2022
Mr Johnathan Davis ACT MLA Visit	27 April 2022
ASTRO3D and ANU School of Astronomy Visit	4 May 2022
ANU Research School of Computer Science – Undergraduate Student Tour	19 May 2022
The ARC Centre of Excellence for Climate Extremes Winter School 2022 Tour	23-25 May 2022
ASTRO3D Indigenous Work Experience Program – Visit	25 May 2022
ISC High Performance 2022 Conference – Online Exhibitor	29 May - 2 June 2022
ACCESS-NRI Community Tour	22 June 2022
Department of Industry, Science, Energy and Resources Visit	27 June 2022
Ms Elizabeth Kikkert ACT MLA Visit	29 June 2022
Stakeholder Visits to NCI Vizlab	Multiple dates

Training Activities

Training	Date
MathWorks Workshop Series - Machine Learning	19 July 2021
MathWorks Workshop Series - Deep Learning	20 July 2021
MathWorks Workshop Series - Parallel Programming	21 July 2021
HPC and Data in Materials Design and Discovery Course	1 September 2021 - 26 November 2021
NVIDIA Workshop with NCI: JupyterLab Setup with Gadi Bootcamp	28 October 2021
NVIDIA Workshop with NCI: Accelerated Data Science GPU Bootcamp	19-20 April 2022
NVIDIA Workshop with NCI: Distributed Deep Learning	4-5 May 2022
NVIDIA Workshop with NCI: CUDA Python GPU Bootcamp	18-19 May 2022
AI and ML in Drug Design and Discovery Showcase	26 May 2022
Parallel Programming for HPC	29 March 2022
Introduction to Machine Learning using R: Introduction and Linear Regression	16-17 May 2022
Introduction to Machine Learning using R: Classification	23-24 May 2022
Introduction to Machine Learning using R: SVM & Unsupervised Learning	31 May 2022
Version Control with Git	2 June 2022
Data Visualisation in Python	3 June 2022
Beyond the Basics: Julia	9 June 2022
Introduction to Gadi	Multiple dates
NCI Presents: TechTake	Multiple dates
Unix Shell and Command Line Basics	Multiple dates
Getting started with HPC using PBS Pro	Multiple dates
Learn to Program: Python	Multiple dates
Learn to Program: R	Multiple dates
Learn to Program: Julia	Multiple dates
Data Manipulation and Visualisation in Python	Multiple dates
Introduction to Machine Learning using Python: Introduction and Linear Regression	Multiple dates
Data Manipulation and Visualisation in R	Multiple dates
Introduction to Machine Learning using Python: Classification	Multiple dates
Introduction to Machine Learning using Python: SVM & Unsupervised Learning	Multiple dates
NCI HPC Toolkit: Introduction to OpenMP	Multiple dates
ML4AU Machine Learning Showcase	Multiple dates