New Frontiers in Science and Discovery

Pawsey Supercomputing Research Centre
Annual Report 2020–2021

Pawsey

TABLE OF CONTENTS

A national research infrastructure and leader in science and innovation

Making Tomorrow Happen, Today

Chairman's Foreword

Welcome to the Pawsey Supercomputing Research Centre 2020–21 Annual Report

> "It is a potentially transformative period for universities and Australia's wider research landscape, and we have a major part to play in maintaining our critical digital infrastructure and the expertise that continues to enable it.

Working through a pandemic has thrown into sharp focus the reliance we have on the infrastructure that connects us, and the data and communications that support research and evidence-based decision making. We acknowledge the foresight and strategic commitments made by both the Western Australian and Australian Governments to invest in Pawsey, enabling us to continue to deliver significant value against both State and national priorities while also tackling emerging problems like the COVID-19 pandemic with our international partners.

Facilities like Pawsey work with our core research partners, with government, and increasingly with industry to enable greater national technological capacity and to provide strategic direction in support of their own digital capability initiatives.

We have delivered against major milestones in our $70 million capital refresh program and the 130 PB upgrade of our storage infrastructure will be live by the end of the year. Excitingly, 'Setonix' Phase 1 will also be operational by year end and fully commissioned in 2022. I am delighted that Pawsey's flagship system has been named Setonix, with the quokka becoming recognised internationally as an iconic goodwill ambassador for Western Australia.

This massive investment in Australia's supercomputing infrastructure has only been possible with the support of both the Australian and the Western Australian Governments and our founding research partners: CSIRO and Western Australia's four public universities. CSIRO and Curtin University's contribution to the Capital Refresh through governance, procurement and legal support has been significant.

Pawsey's board has been particularly engaged with risk management and oversight of this complex joint venture through this challenging and evolving period. The commitment of our partners to effective risk management and the wellbeing of Pawsey and all of its staff through this period is to be commended.

This major multifaceted procurement and technical project has been delivered through the challenges of a global pandemic and will provide an historic uplift to national research supercomputing and data services, providing opportunities for the entire nation.

Pawsey and the National Computational Infrastructure (NCI) are essential digital infrastructure in the National Collaborative Research Infrastructure Scheme (NCRIS). We welcome the appointment of NCI's new Chairman Dr Greg Ayers this year and will continue to work together to lead both national and international High Performance Computing (HPC) and data research efforts. Our efforts are strategically aligned with national and regional research priorities that reinforce our future scientific capability and support economic, social, societal and environmental advances.

Our investment in digital infrastructure and computational capacity is also central to the long-term commitment Australia has made to the Square Kilometre Array (SKA) project. Continuing this commitment, Pawsey hosted the Prime Minister in April who affirmed Australia as a member of the SKA Observatory (SKAO), created in December as only the second statutory intergovernmental organisation in the world dedicated to astronomy. The Prime Minister announced a further $387 million towards constructing the SKA, including $64 million for supercomputing. Pawsey is an integral part of Australia's involvement in this mega-science project.

Pawsey is also committed to support Australia's emerging space industry through our role in the Australian Space Data Analysis Facility (ASDAF).

It continues to be a challenging time for the research sector in Australia, as it is for many other sectors through the COVID-19 pandemic. It is a potentially transformative period for universities and Australia's wider research landscape, and we have a major part to play in maintaining our critical digital infrastructure and the expertise that continues to enable it.

I would like to thank the Board Directors for their work and support throughout the year, and again congratulate the entire Pawsey team for continuing to build Australia's digital future through a time of global upheaval. Their on-going efforts are recognised throughout this report, and as always, I am proud to be associated with them.

As I will be leaving Pawsey in October 2021 to become Western Australia's Agent General for the United Kingdom and Europe – a position which is based in London – this will be my final report and I wish everyone associated with Pawsey well for the future.

John Langoulant AO
Chairman of the Board

Executive Director's Report

Looking back and despite the challenges posed through the COVID-19 pandemic, Pawsey has had a very good year. Pawsey and its staff have continued to deliver and expand the programs and activities supporting the research communities and key stakeholders that we serve.

> "Pawsey is the sum of its many different parts, and we are steadily growing richer.

With the $70 million capital refresh program nearing completion and 'Setonix' rapidly becoming a reality, we launched the Pawsey Centre for Extreme Scale Readiness (PaCER) program in late 2020. This $3 million program is allowing researchers to become exascale-ready and be the first to exploit the computational scalability offered by Setonix. Ten projects are underway and represent the leading edge of HPC-powered software development for exascale computing in Australia. Setonix will enable the computational science and data intensive approaches increasingly required by Australia's leading researchers to tackle the world's wicked problems.

'Business as usual' with our existing systems has continued to progress in a flexible, remote, and Covid-safe manner. Online delivery of our services, particularly training, support and outreach to potential new users, has expanded in scale, thanks to the efforts of our communications and education staff. They have provided an accessible platform for our operations not only under current circumstances, but also in support of our increasingly national and international activities.

To be more reflective of who we are and where we are going, we relaunched Pawsey's brand in May 2021. Becoming simply 'Pawsey', with the new 'P' and data cube logo, reflects our origins and history with radio astronomy and computational innovation, while expanding our relevance beyond traditionally regional research domains to capture our increasingly national and international outlook and capabilities.

We have also contributed to several national and international milestones over the last year. The Australian Space Data Analysis Facility (ASDAF), hosted by Pawsey in partnership with the Australian Space Agency and the Western Australian government, was launched in May. This facility is helping businesses access space and Earth-observational datasets and computing infrastructure, tools and training, to develop new products and services. Through ASDAF we continue to grow support for Australia's emerging space industry.

Pawsey also remains an integral part of Australia's ongoing commitment to the SKA, with active involvement in the operation of the Murchison Widefield Array (MWA) and the Australian Square Kilometre Array Pathfinder (ASKAP), and Australia's new SKA Regional Centre.

As a national supercomputing facility, while we're leading investment in next-generation computing, it's not enough to just service the immediate needs of our research community. We are also trying to anticipate their future needs, and realise potential opportunities in HPC. To this end, we are partnering with innovative researchers and companies to develop and pilot quantum computing technologies. We've been testing software algorithms and industry applications on quantum emulators in our Quantum Pioneers program and in the coming year we will be one of the first to test quantum hardware from Australian start-up company Quantum Brilliance in an operational HPC centre. Science fiction is rapidly becoming an industrial fact in quantum computing.

Pawsey's true potential is only unlocked because of the people I am privileged to represent, and the people we are fortunate to work with. Throughout the last year we've remained focused on the wellbeing of not only our staff but also the research communities we support, and I'm exceptionally proud of the way we've continued to grow and deliver through a globally challenging time. I would like to acknowledge the tremendous contributions of everyone at Pawsey to support each other and our shared ambitions in such unprecedented conditions. Many of the ways we've adapted to work flexibly and remotely will stay with us as we expand in an uncertain future, as is appropriate for an innovative and progressive science and technology organisation.

Our on-going commitment to equity, diversity and inclusion is more important than ever, and we continue to invest in this from the Board, through the Executive and senior management, through our communities of practice to our internship program. Pawsey is the sum of its many different parts, and we are steadily growing richer. Throughout this report you'll find examples of the great expertise and diversity of the many faces of Pawsey.

I'd like to acknowledge the support of Pawsey's Board, our foundation partners, the Western Australian and Australian governments as we continue to evolve, innovate and lead essential national supercomputing and data infrastructure for the benefit of the communities we serve.

Mark Stickells
Executive Director

Systems and Capabilities

Over 1,500 hours of training was delivered to teachers and students as part of the STEM strategy

5,500 people were reached with training

2020

September

Garrawarla goes live – The new 78-node cluster provides a dedicated system for astronomers to process in excess of 30 PB of MWA telescope data.

October

HPE awarded the contract for the Pawsey Supercomputer System (PSS).

November

23 new interns joined Pawsey Summer internship out of 300 applications.

December

The ASKAP ingest nodes were delivered by HPE.

2021

January

Over 512 million hours requested across all allocation schemes.

143 projects received an allocation on Pawsey supercomputers.

Contracts awarded to Dell and Xenon for the multi-tier Storage procurements.

February

Pawsey announced the new PSS has been named 'Setonix' after the species name of the quokka, Setonix brachyurus.

The new firewall was procured from Palo Alto. The firewall configuration will support a throughput of 100 Gbps into AARNet, and hence to other Research and Education (R&E) networks and into the NCI.

The network upgrade was procured from CISCO. Pawsey has moved away from a monolithic single core router to a spine-leaf architecture with a 400 Gbps backbone and 100 Gbps links to host endpoints.

March

Quantum Brilliance made a quantum emulator available to be hosted at Pawsey for staff and early adopters from academia and industry to test.

April

May

Pawsey rebranded.

Setonix Phase 0 (Test and Development System) arrived at Pawsey and was named 'Joey' (a baby quokka).

June

People and collaborations

over 2,000 people
joined Pawsey events

980 people
were reached by Pawsey's first virtual tour during National Science Week

2020

JULY

AUGUST

Pawsey and NCI joined the COVID-19 High Performance Computing Consortium led by the White House Office of Science and Technology Policy, the U.S. Department of Energy and IBM.

SEPTEMBER

OCTOBER

Pawsey launched new partnering program to achieve HPC research at scale, the Pawsey Centre for Extreme Scale Readiness (PaCER).

A new Australasian Chapter of Women in HPC (WHPC), a collaboration between Monash University, AeRO, NCI Australia, Pawsey, and NeSI, was announced.

NOVEMBER

DECEMBER

2021

JANUARY

FEBRUARY

Ten PaCER projects were awarded.

MARCH

Pawsey and Quantum Brilliance joined forces with other Australian industry leaders and researchers as part of Pawsey's Quantum Pioneer Program

APRIL

A $387 million investment to build the SKA radio telescope, including $64.4 million to establish the Australian SKA Regional Centre (AusSRC), was announced at Pawsey by the Prime Minister Scott Morrison and the Minister for Industry, Science and Technology Christian Porter.

MAY

Australia Space Data Facility opened their Expressions of Interest to SMEs.

JUNE

Two new virtual communities of practices on HPC – Computational Fluid Dynamics (CFD) and bioinformatics – were created for knowledge sharing.

ABOUT PAWSEY

Pawsey is a world-class high-performance computing facility accelerating scientific discoveries for Australia's researchers.

Pawsey started its transition to become a world-leading supercomputing facility in 2018, beginning the $70 million Capital Refresh Program delivered through a grant from the Australian Government. This upgrade to Pawsey's supercomputing infrastructure is helping us stay competitive in relation to other world economies. We are transitioning to new technologies, new ways of doing, and helping build new research communities around HPC.

This transition has become increasingly evident with Pawsey's rebranding, the arrival and commissioning of new systems and services, and expanded and refocused training programs. These are all leading to the deployment of Setonix supercomputing and all of the supporting infrastructure needed for the success of Australian researchers working at scale.

COVID-19 imposed several challenges but also accelerated organisational changes worldwide. Many new connections were forged by Pawsey during this period, and existing partnerships became stronger. The Centre moved from a face-to-face to a more virtual presence, increasing its reach and developing new relationships as a result. Pawsey has become much closer to its research base, moving from a technology focus to a service-focussed facility.

Since Pawsey's inception, radio astronomy research has been central to our activities, however during the last few years space, AI and bioinformatics became emerging areas of activity and strength. The rise of 'big data', particularly in astronomy and the life sciences, continues to underpin our investment in data storage, keeping researchers at the enabling edge of this research.

We continue to create new opportunities to contribute to science with real impact and demonstrate the relevance and benefits provided by high performance computing facilities.

PAST AND PRESENT

EPIC at Murdoch

Fornax at UWA

Pawsey Centre construction started

2013 Pawsey Centre opened its doors

2013 Magnus phase 1 arrived

Along with tape storage

2014 fully commissioned Magnus was made available to researchers

LOOKinG FORWARD

PAWSEY IS ON THE CUSP OF A STEP CHANGE IN INFLUENCE

The Centre's growing capacity, and our successful engagement with Australian researchers, peer facilities around the world and stakeholders in a broader sense, is the beginning of a new and stronger chapter and a greater leadership role in HPC and research for Australia.

PAWSEY IS POISED FOR GROWTH

The capital refresh allows us to look beyond our previous limits, identify new partnership opportunities, increase the number of scientific projects we support and outcomes we help create, and supports new services in training and industry engagement. We are building the foundations to drive GPU uptake by Australian researchers and contribute to Australia's investment in AI and space.

PAWSEY IS PARTNERING WITH RESEARCHERS ON THIS GROWTH JOURNEY

Pawsey researchers have played a key role in the Centre's growth and Pawsey experts will continue working closely with them to help them take full advantage of the new infrastructure and achieve computational research at unprecedent scale.

PAWSEY IS ENGAGING ON THE FRONTIERS OF SCIENCE

Recent work, particularly that involving COVID-19 research, the SKA, and quantum computing, has expanded our reach. We will continue working at the forefront of technology and science on our path towards innovation and as a discovery enabler. The Centre will help grow the HPC community in Australia and will work with emerging research and industry partners to accelerate breakthroughs on Australia's path to exascale.

Improving our 'green' credentials to operate has been an important part of the tender process to procure Setonix, and it continues to be a priority area at the Centre moving forward.

COLLABORATION IS PART OF PAWSEY'S DNA

Pawsey and Australian researchers are benefiting from the collaborations in HPC and science that the Centre has established. We will continue developing partnerships that help accelerate researchers' work and build a stronger HPC community.

Mind-blowing
solutions
to universal
problems

Amazing
outcomes

CASE STUDY 1

QUANTUM

Partner Institution:

QUANTUM BRILLIANCE

THE CHALLENGE

Most quantum computers in existence today are very fragile – they need vacuum systems, supercooling and lasers to stay stable enough to operate. It limits them to very specialised mainframe computing installations with big buildings, big budgets and teams devoted to their operation. This in turn severely limits the opportunities to learn how to best use these systems, integrate them with other computing facilities, and work out how to program and operate them to do something useful.

Making quantum computing hardware more easily accessible is a necessary first step to demonstrating their capabilities and allowing their wider use.

THE SOLUTION

Quantum Brilliance is a start-up company from the Australian National University building diamond-based quantum accelerators. By using the extremely stable properties of synthetic diamonds, their quantum computers can operate at room temperature with much simpler control systems, making them much easier to miniaturise.

FIRST diamond-based quantum accelerator

25 QUBITS+ to be run in the emulator

4 early adopters currently testing the technologies

THE PATH TO UBIQUITOUS QUANTUM COMPUTING

PROJECT LEADER: Dr Andrew Horsley, Dr Marcus Doherty and Mark Lou/ Quantum Brilliance

Quantum computers promise a revolution in computing ability, but there are many hurdles to be overcome to turn this promise into reality. Some of these challenges includes high upfront capital cost to meet ultra-low temperature requirements, high on-ongoing operational and reliability costs needed to maintain vacuum and cooling infrastructure, and limited synergy with current state-of-the-art HPC data centre practices. Start-up company Quantum Brilliance and Pawsey are partnering to show how the future of supercomputing can include not only CPUs and GPUs, but diamond-based quantum accelerators.

Mark Luo, Chief Operating Officer at Quantum Brilliance explains: “We want to make quantum computers useful by dramatically lowering the barrier to entry by providing quantum utility, rather than building a massive mainframe quantum computer… the key is to make them more accessible, and mass installed in supercomputing facilities which enables deep integration with classical compute nodes. This enables the entire computing community to start developing programs to run on them.”

“Once we can integrate and run our diamond quantum accelerator prototype with supercomputing infrastructure, we will understand how to integrate ten or a hundred of them, leading to massively parallelised quantum computing.”

Quantum Brilliance and Pawsey have partnered to create Australia’s first Supercomputing-Quantum Innovation Hub, becoming one of the first few supercomputing centres worldwide to host a quantum computer onsite and creating a community of researchers and industry participants to develop and test software for the application of quantum accelerators for quantum utility problems.

OUTCOMES

Quantum Brilliance’s Quantum Emulator is already operating at Pawsey, and the qbOS, the next release which will work with the Quantum Accelerator prototype, is due in 2022. Researchers in the Pawsey Quantum Pioneer Program are already using the emulator to develop different quantum circuits and algorithms, and exploring how they will eventually run on the hardware.

Mr Luo explains: “The emulator replicates how the hardware will perform, including the number of qubits, how they’re connected, and the error and noise characteristics of our diamond-based accelerators.”

“Seeing how the quantum emulator behaves allows all of us to experience what the real hardware will eventually do, but we can work with it today. Using Pawsey supercomputing allows us to run an emulator containing 25 qubits or more, while for smaller numbers of qubits, the quantum emulator runs well on the Pawsey Nimbus cloud.”

A particular goal with the emulator is to accurately estimate the run time of quantum circuits on the diamond-based quantum accelerator. “We want to compare the quantum computing time against classical computing time,” says Mr Luo. “Demonstrating ‘quantum utility’ means the industry needs to compare computational power comparable in terms of physical size, weight and power consumption as well.”

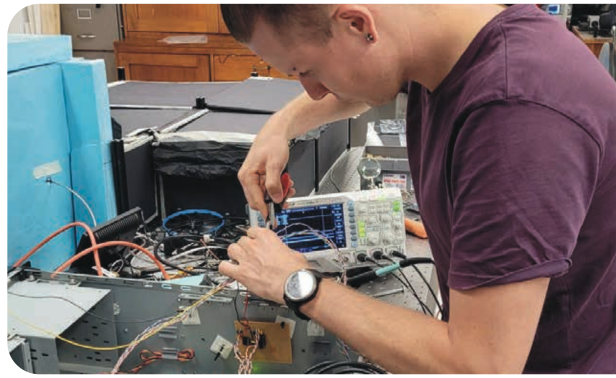
In 2022, the diamond-based quantum accelerator prototype will be installed at Pawsey, allowing Quantum Brilliance and Pawsey to learn how to integrate it with Pawsey’s existing hardware and software. Software integration is already occurring, as a result of the work being done through collaboration on the Quantum Emulator. This will then be followed by a Quantum Accelerator installation which will contain more qubits and a next-generation control system.

“Our eventual goal is to provide a full-suite of quantum computing hardware systems, which Pawsey can host and operate as part of its heterogeneous computing architecture. Quantum accelerators can then be deployed to be part of the solution to give better overall computing performance.”

“At that point, the only question remaining is how many accelerators you want to plug in.”

Four groups of Quantum industry leaders and researchers have been granted access to the already installed Quantum emulator at Pawsey as part of Pawsey’s Quantum Pioneer Program – an initiative for selected researchers to develop cutting-edge quantum applications in machine learning, logistics, defence, aerospace, quantum finance and quantum research.

- **Quantum South**, a team working with complex optimization problems for cargo in airlines and ships leveraging quantum computing software.
- **Trellis Data**, an Australian company delivering Machine Learning capability across Government, Commercial and Not-For-Profit organisations.
- **Prof Jingbo Wang**, Head of the Physics Department at the University of Western Australia and a pioneer of cutting-edge research in the development of quantum walk-based algorithms, aiming at solving complex problems of practical importance otherwise intractable.
- **Dr Casey Myers**, a senior research fellow in the School of Computing and Information Systems at the University of Melbourne, playing a leading role in establishing a new quantum research and education program within the Faculty of Engineering and Information Technology.



CASE STUDY 2

COMPUTATIONAL FLUID DYNAMICS

Partner Institution:

THE UNIVERSITY OF MELBOURNE

PROPELLING THE ENVIRONMENTAL EFFICIENCY OF JET ENGINES

PROJECT LEADER: Prof. Richard Sandberg, Chair of Computational Mechanics, University of Melbourne

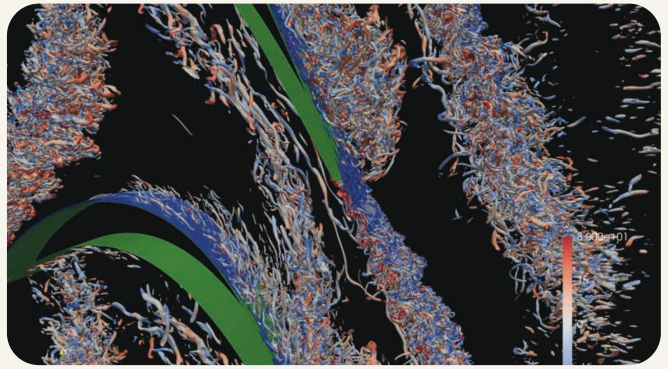
PROJECT LEADSYSTEM: Magnus

AREAS OF SCIENCE: Turbulent Flows

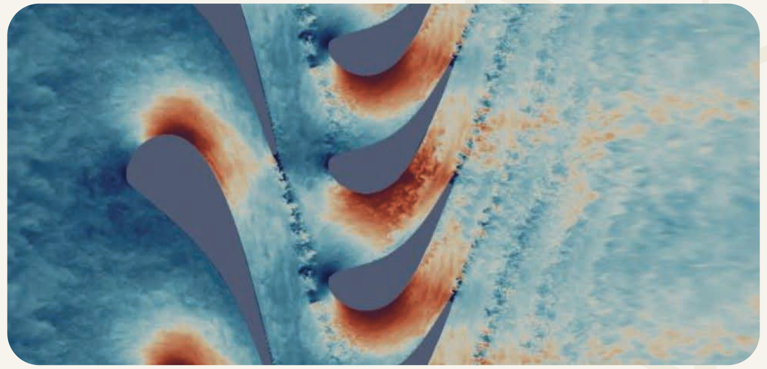
APPLICATIONS USED: HiPSTAR

Global fuel consumption is continuing to climb, with commercial airlines alone reaching an all-time consumption high of 363 billion litres in 2019. Increasing the efficiency of an aircraft by even 0.1 per cent would have a massive global impact on fuel demand, costs, and the environment.

Prof. Richard Sandberg and a team at the University of Melbourne are working towards this impact, using high-fidelity simulation approaches to examine the effects turbulence has on aircraft engine efficiency, ultimately creating cleaner methods of air travel.



Multi-scale wakes being distorted when passing between turbine blades, thereby generating efficiency loss. Image courtesy, Richard Sandberg, University of Melbourne



High-fidelity simulation of a rotor-stator interaction in a high-pressure turbine at real engine conditions

20,700,000 core hours allocated

$500,000 worth of reduced CO2 emissions

1% boost to engine efficiency

THE CHALLENGE

Making aircrafts more efficient is a large-scale problem that demands an equally large-scale, significant solution. Prof. Sandberg and his team at the University of Melbourne are taking on the challenge, working to better understand turbulent flows in jet engines and how they affect efficiency and noise generation.

Prof. Sandberg's work combines numerical methods and high-performance computing to predict gas flows in turbine engines and translate the data generated into predictive models.

The work is unique in that it is conducted and resolved on full-scale gas turbine flows. This type of high-fidelity work where turbulent flow is represented by simulation requires tremendous computing power. Performing calculations at real engine conditions, rather than at half-size model scale, drives the computational cost up eight times.

THE SOLUTION

The team's predictive modelling methods use around 100 million computer hours per year — the equivalent of approximately 3,000 years on conventional desktop computers.

Backed by Magnus' computing power, Prof. Sandberg's high-performance computing (HPC)-optimised compressible fluid solver is able to produce predictions of the detailed turbulent motions that affect turbine efficiency. The data generated by these computational fluid dynamics simulations is vast, with a single simulation generating tens of terabytes, which need to be down-sampled in order to store and subsequently deliver research findings.

"I'm focused on solving equations that are very non-linear and complex that we have not been able to find general solutions for to date. With this technology, we can now use high-fidelity simulation approaches to conduct world-leading research for industrially-relevant problems," explains Prof. Sandberg.

OUTCOMES

Evolving such complex, large-scale machines requires a focus on the most minute details. As a leader in the study of turbulent flows, Prof. Sandberg's team is paving the way for further understanding and developing the way gas turbine engines perform under various conditions.

The data Prof. Sandberg generates is used to obtain direct physical insights for his own project, as well as to help improve industry models. Prof. Sandberg's team has developed machine learning tools that translate the large-scale data he generates into plug-and-play models. These tools have already attracted the attention of the US Navy, General Electrical, and Mitsubishi Heavy Industries.

With a secret superpower in supercomputers, Prof. Sandberg has been able to contribute significant outcomes to the complex issue of turbulent flows. It's just the start of a new avenue of research in the field, but the early results from his simulations are paving the way for huge global impact.

CASE STUDY 3

Health Science + AI

Partner Institution: The University of Western Australia

Predicting medical complications in time for treatment

PROJECT LEADER: Dr Janis Nolde, School of Medicine, Royal Perth Hospital and the University of Western Australia

SYSTEM: Nimbus

AREA OF SCIENCE: cardiovascular research

APPLICATIONS USED: python3 and tensorflow

Where machine learning and big data meets continuous monitoring of at-risk hospital patients, there is the possibility for artificial intelligence to identify the risk of medical complications before they occur. Researchers from the Machine Intelligence Group at the University of Western Australia, clinicians from Western Australia's first inpatient remote monitoring service, the 'HIVE' at Royal Perth Hospital (RPH), and Pawsey supercomputing experts are working together to identify and better treat the most vulnerable patients in our healthcare system.

>20,000 patients to train initial models

> 5000 planned for validation projects

5 collaborating institutions

THE CHALLENGE

Cardiac complications are common after major non-cardiac surgery, and are the primary reason for death and disability in this setting. More than 10 million adults experience a major cardiac complication, or myocardial injury, within one month of a non-cardiac operation annually.

Post-operative myocardial injury is caused by a range of mechanisms, differing between individuals and over time, making it difficult to predict and prevent. Current methods to identify patients at risk of cardiac complications rely on pre-operative clinical factors and the intended surgical procedure but ignore factors occurring during surgery and recovery, which limit their value.

However, the onset of myocardial injury can be detected through increased levels of specific heart proteins in the blood, and the presence of these markers is strongly associated with adverse patient outcomes.

THE SOLUTION

Dr Janis Nolde from the School of Medicine, Royal Perth Hospital and the University of Western Australia, points out that there are increasing efforts to use technology to continuously monitor at-risk patients. "The Health in a Virtual Environment (HIVE) system began operation at RPH last year. It collects patient physiological data like continuous blood pressure monitoring, respiration and oxygenation levels from monitored beds, as well as healthcare notes made by bedside staff, medical imaging and pathology test results, and streams them to a central command centre staffed by clinical experts."

With a future full capacity of up to 1,000 patients continuously monitored at a time, the HIVE is supported by an artificial intelligence (AI) platform that detects subtle changes in patient condition and the earliest signs of clinical deterioration, using machine learning algorithms to continuously estimate the risk of complications and alert medical staff.

The AI in use is a commercial product, developed on large international datasets of patient information to monitor several specific health conditions. But the HIVE is also intended for medical research, as it is now generating large amounts of detailed patient data (collected with consent).

"The existing AI model hasn't been developed to identify all of the adverse medical outcomes we may want to focus on, and doesn't take into account potential differences between our local population and the international populations it was developed with," says Dr Nolde. "But we can use the data collected by the HIVE to develop our own machine learning models to target risks and adverse events relevant for our local population."

Machine learning has many advantages where large amounts of data are being acquired, and where causal links between events are complex or unclear. It can identify patterns predicting injury, recognise specific triggers for events, and can continually improve risk estimates as new data becomes available and interactions and patterns become more apparent.

OUTCOMES

Dr Nolde is using Pawsey facilities to train machine learning algorithms using datasets of patient condition from before, during and after surgery, including markers of myocardial injury in the blood, to see if they can accurately estimate an individual's risk of post-operative cardiac complications. "We anticipate that machine learning will be able to out-perform existing risk scores for cardiac complications, and identify the complex patterns and specific precursors to myocardial injury. This could ultimately improve post-operative care and reduce the incidence of cardiac injury."

The first step, training machine learning algorithms with existing international datasets, is already underway. Once the HIVE has generated an even richer pre-and post-operative dataset from up to 10,000 Australian patients, Dr Nolde will be retraining and adapting the AI prototype to better improve its predictive capabilities for cardiac complications in the local population.

Once the AI algorithm is developed, it will run alongside current HIVE operations but separated from patient treatment, to see how well it predicts cardiovascular injuries as they develop after surgery in real time.

"If we can establish that the algorithm is a clinically useful risk prediction tool, only then will we be able to embed it into the decision-support AI at the HIVE. It is a long-term goal to change the way we both assess and reduce the risk of cardiac complications for the millions of patients undergoing surgery every year."

With thousands of patients accumulating medical data at the HIVE and Pawsey supercomputing able to train machine learning networks that consider the hundreds of different variables embedded in them, Dr Nolde is hopeful that we can develop tools to assess many types of patient risk, guide their management, and improve health outcomes.

CASE STUDY 4

ARTIFICIAL INTELLIGENCE

Partner Institution:

Curtin University

KICKING GOALS WITH GPUS

PROJECT LEADER: Dr Sajib Mistry and Associate Professor Aneesh Krishna, Curtin University

SYSTEM: Topaz

AREAS OF SCIENCE: Artificial Intelligence, Neural Networks

APPLICATIONS USED: Forest Deep Neural Network, Python, CUDA

Soccer is the biggest game in the world. With over 250 million active players, followed by four billion fans, it's a huge, competitive industry.

While the game is on the pitch, a massive data industry sits behind it. This industry looks for the secrets to victory in player actions. The biggest teams hire secretive data consultants for an edge in this multi-billion-dollar industry.

Now, three researchers at Curtin University have developed a tool to level the playing field. The project was supported by Jordan Makins via Pawsey Summer Internship.

26
Field zones

25000
Allocation hours

1000 GB
storage allocation

THE CHALLENGE

Graduate student, Jordan Makins is a member of Curtin's High Performance Intelligent Systems research group. With guidance from Assoc. Prof. Aneesh Krishna and Dr Sajib Mistry, he set out to make an open-source tool for analysing player actions during soccer matches.

The 2003 Michael Lewis book, Moneyball, popularised sports analytics by following the data-driven approach of the Oakland Athletics Baseball team. Since then, many sports have turned to data to get an advantage in games, but soccer has been resistant to this approach.

"Soccer is the most popular game in the world, but soccer analytics are so complicated, not a lot of work is done in this space," says Dr Sajib.

The sport is set on a constantly evolving field where many actions occur at once. Any computer analysis needs to cut through all that noise to the actions that lead to winning soccer matches.

"The big teams like Barcelona and Liverpool invest a lot in analytics, but that knowledge is proprietary, so others don't have access to it. Our findings are open-source, so people may use them however they like."

THE SOLUTION

Supercomputing was vital for analysing the field. Each second of game time was logged using event streams. Soccer experts logged these streams, keeping track of every action from passes to dribbles and goals. The trio divided the field into 26 zones and used the logs to analyse plays.

"We were able to aggregate the different action types per zone and assign them values. We used a random forest feature selector to do this, then fed it into a deep neural network. This is the first time this approach has been used in soccer analytics," says Mr Makins.

A random forest is a machine learning model that uses many decision trees to sort data by values. Random forest sorting is already used in genetics, so the team studied work in this field and decided to combine it with a deep learning system to trial a novel soccer analytics approach.

"If you think of a DNA strand, it's a long, complex vector of values, much like our soccer data. The Forest Deep Neural Network was used in a recent genetics paper, so we figured it could be a good crossover," says Mr Makins.

"It's a great technique for selecting features. In our soccer analytics, certain actions we wanted to value were nested in a lot of constant, changing noise. This feature selector helped us cut through that noise and find the actions that were really valuable to us," says Dr Sajib.

With a large amount of data coming in and two years' worth of European games to analyse, the team needed access to a graphics processing unit (GPU) supercomputer, and gained access to the Topaz GPU cluster.

"It was important we had access to this kind of HPC power. If you think about the fluid nature of this data, we needed to process it as close to real-time as possible. Eventually we'd like to train the software to analyse a live soccer game," says Dr Sajib.

Pawsey helped Mr Makins upskill to work on the Topaz cluster. After completing introductory supercomputing and parallel data coding modules, he applied the knowledge to his research project.

"I would remote connect from my PC, so I didn't need to go to the Pawsey building. I would run my Python code from home on my PC connected to the cloud and I would get the superior performance of the Pawsey supercomputer," says Mr Makins.

OUTCOMES

"Curtin's partnership with Pawsey trains our students and researchers to use HPC resources. That training is very thorough. It's carefully updated every year and goes through a step-by-step process. It also gives us access to computing power for this kind of exploratory research." – Assoc. Prof. Aneesh Krishna.

The open-source model is usable by anyone. This means soccer teams no longer need huge budgets to hire consultancy firms to make data-driven decisions for their players. It levels the playing field in the sporting world, while exploring powerful new data analytics techniques.

"In the realm of machine learning, this is a newer ensemble algorithm. It's the first time somebody's combined these two widely-tested algorithms. The mass adoption of a new algorithm is always an exciting concept. In the same way we got the random forest feature selector from a genetics study, other fields could use this approach to identify valuable data within a lot of noise."

The model assigns values to players based on their actions. Unlike other models, defensive and mid-field actions are considered alongside goals, so defensive players are valued. With these player values, future research could explore in-game, data-driven decision making. It could help managers pick players on a limited budget and analyse opponent tactics in near-real-time.

"There are still a lot of analytics left to do and we're looking at making all of this open-source, so everyone has access," says Dr Sajib.

The team are using Pawsey supercomputing to democratise soccer data and give everyone a fair go.

CASE STUDY 5

CHEMISTRY

Partner Institution:

CSIRO

This image shows the centre of the Milky Way as seen by the Galactic Centre Molecular Line Survey. Credit: Chenoa Tremblay (ICRAR-Curtin)

TO BOLDLY GO WHERE NO SUPERCOMPUTER HAS GONE BEFORE

PROJECT LEADER: Dr Chenoa Tremblay, Postdoctoral Fellow in Dark Magnetism, CSIRO

SYSTEM: Magnus

AREAS OF SCIENCE: Astrochemistry, birth of stars, search for extra-terrestrial intelligence

APPLICATIONS USED: Python (Numpy, Scipy, Astropy), AIOFlagger, AOtools (Cotter, Calibrate), WSClean, Miriad, Aegean, MWA_Tools, CASA

For thousands of years, we have stared at the stars and wondered 'are we alone in the universe?' As our understanding of our own Solar System has developed in the last 50 years, we know that finding life requires us to look beyond our immediate neighbours to the stars beyond.

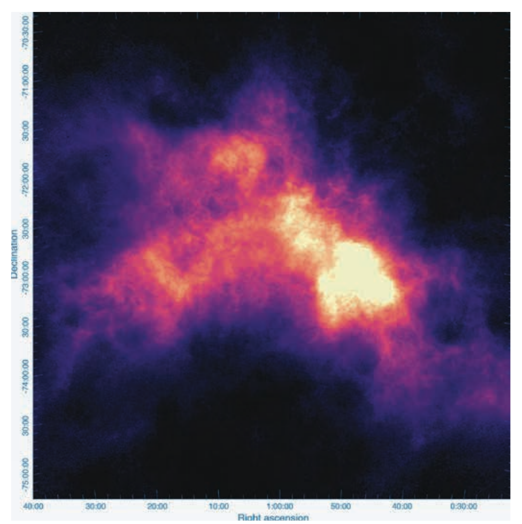
Dr Chenoa Tremblay and her team are using Pawsey's systems to analyse data from radio telescopes in an attempt to finally answer one of the most fundamental questions of our existence.

500,000+ images of the sky to analyse

9 team members

0 technosignatures

10 publications



An image of the Hydrogen gas in the Small Magellanic Cloud (SMC), as observed with the Australian Square Kilometer Array Pathfinder. The SMC is one of our nearest and smallest galactic neighbours. The image is from a work published in Nature led by Professor Naomi McClure-Griffiths from the ANU Research School of Astronomy and Astrophysics and Professor John Dickey of the University of Tasmania. Image Credit: McClure-Griffiths et al. and CSIRO

THE CHALLENGE

Dr Tremblay uses spectroscopy, the study of the interactions between light and matter, to look for molecules in the gas layers around stars. As molecules in space are bombarded with energy from nearby stars, they can absorb and re-emit that energy at specific frequencies based on their structure, with each molecule's energy signature being unique.

She uses radio telescopes to collect signals across a range of narrow-frequency bands to build up and identify those full-spectrum energy signatures. This is a relatively new science, beginning in 2014 with her PhD work at Curtin University. With little historical perspective, Dr Tremblay needs to rely on modelling, theories, and ideas to predict the impact of her work.

Searching for technosignatures and molecules means, as part of a recent study, compiling more than 17 hours' worth of information on the sky, equivalent to hundreds of terabytes of data. On a standard laptop, processing and analysing this data would take more than 25 years.

THE SOLUTION

Dr Tremblay uses the Murchison Widefield Array (MWA) radio telescope, a precursor telescope to the Square Kilometre Array (SKA), to collect observations and detect astronomical signals that help her investigate the hypotheses she has set out. These volumes of data are stored on Pawsey's systems until they're ready to be used.

In order to be able to image and analyse the astronomical signals, Dr Tremblay needs to use different pieces of software, including Python, AIOFlagger, MWA Tools, and CASA, among others. This is where Pawsey's systems come in, allowing Dr Tremblay to switch between software packages and run multiple datasets simultaneously through each, combining all of this disparate data in the last step.

As Dr Tremblay explains: "There isn't a pre-formed software package to analyse this data, so we have to download the data, and either write our own software or use software packages written by other astronomers or software developers to complete each step of the process."

With more than 500,000 images of the sky to analyse, there is a need to process multiple sets of data at the same time, as well as organise the imagery and make it easily discoverable for future analysis and comparison.

OUTCOMES

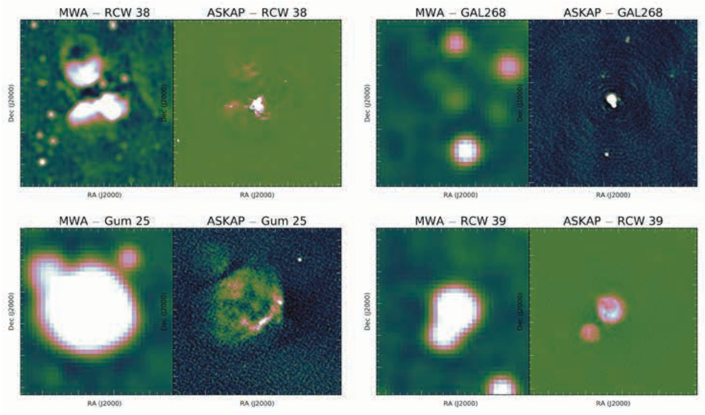
As Dr Tremblay sets out, now is an exciting time for her field: "With this recent dataset, there are a number of different publications completed and soon to be finalised, including the largest ever summary on technological signals."

Together with Prof. Steven Tingay from the Curtin University node of the International Centre for Radio Astronomy Research (ICRAR), Dr Tremblay recently used the MWA telescope to explore hundreds of times more broadly than any previous search for signs of extra-terrestrial life. Scanning a patch of sky with over 10 million stars, they found no technosignatures — that is, no obvious sign of intelligent life.

Dr Tremblay continues her search for signs of alien life, in particular the vital building block nitric oxide, which hasn't been detected as often as expected and will be a focus area for future research.

Key to the future success of the research is the ability to generate and process even more data. Dr Tremblay explains: "We're still not hitting the sensitivity limits, and with the signals being very weak, finding new ones requires even more data".

With the power of the Pawsey supercomputer behind her, Dr Tremblay has accelerated the outcomes of her research from years to days. In doing so, she may also accelerate the most defining discoveries of our lifetimes.



An image compilation of star-forming regions that are currently under investigation with the MWA and ASKAP telescopes. The view from different frequencies (and different telescopes) give us clues as to what is going on in the region and how old it is. The morphologies of each region are different with shapes ranging from "snowmen" to "ballerinas". Image courtesy, Dr Chenoa Tremblay, CSIRO.

CASE STUDY 6

MANUFACTURING

Partner Institution:

RMIT UNIVERSITY

DESIGNING BETTER BATTERIES

PROJECT LEADER: Professor Michelle Spencer, RMIT University

SYSTEM: Magnus

AREAS OF SCIENCE: Chemistry, Environmental Science, Physics

APPLICATIONS USED: VASP, Siesta, Quantum Espresso

Lithium-ion batteries revolutionised portable electronic devices, making them smaller and lighter than ever before. But they've hit their limit. More powerful and lightweight batteries with even higher energy densities will be required to further miniaturise devices, or to power larger vehicles like trucks and aeroplanes.

Prof. Michelle Spencer, from Applied Chemistry and Environmental Science at RMIT University, is using computer modelling to understand exactly what happens between the anode and cathode of next-generation lithium-metal batteries, in search of the best materials to ensure safe and reliable operation of these powerful batteries.

>200ps
of AIMD simulations related to batteries

10 times
increase in energy density

2,700,000
core hours on Magnus

THE CHALLENGE

Despite the widespread adoption of lithium-ion battery technology, limitations in their energy and power density still prevent their use in many advanced devices and vehicles. Efforts to develop next-generation rechargeable battery technologies are still based on lithium-ion as the moving charge, but use a lithium metal anode rather than a graphite anode. These lithium metal batteries can provide a 10-fold increase in energy density, potentially unlocking high-powered battery applications from transport to miniaturised electronics.

Unfortunately, the conventional organic solvent-based electrolytes used in lithium-ion batteries are flammable and reactive with lithium metal, causing continuous reactions that compromise both safety and long-term battery performance.

Ionic liquids have been proposed as new electrolytes as they are inherently safe (very low flammability), have high conductivity, and good thermal and chemical stability. Prof. Spencer explains: "It is essential that the repetitive deposition and stripping of lithium ions at the lithium metal surface, which is the charge movement that occurs during normal battery operation, must occur without significant reaction or decomposition of the electrolyte in contact with it. So, the goal is to design an ionic liquid that will form a stable passive layer at the lithium metal surface but still allow lithium ions through."

Several different ionic liquid electrolytes have demonstrated stable battery performance, but it is not known why some perform better than others. "They are likely to form a solid electrolyte interphase (SEI) layer on the lithium metal surface," says Prof. Spencer, "but the complex composition and structure of these SEIs are very difficult to characterise experimentally – they are thin films only nanometres thick. This makes designing the best electrolyte difficult."

THE SOLUTION

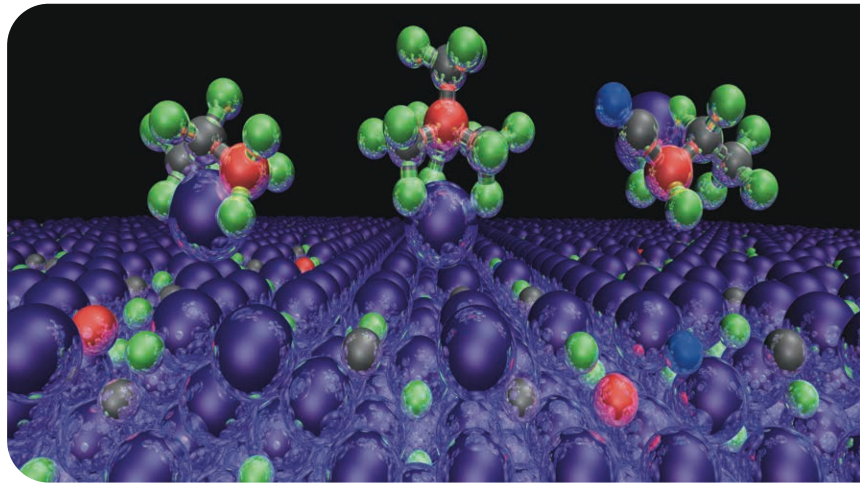
Prof. Spencer and her colleagues Dr Michael Breedon at CSIRO Manufacturing and Dr Thomas Rüther at CSIRO Energy are using a combination of experimental testing, physical characterisation and molecular modelling to understand the formation and structure of effective SEI layers. But with COVID-19 limiting access to physical laboratories, computational experiments using Pawsey's supercomputing facilities have allowed the research to progress.

The team used density functional theory calculations and ab initio molecular dynamics simulations to determine how the various components of ionic liquid electrolytes preferentially adsorb and then react on a lithium metal surface. They have also examined the properties of the lithium salt that forms part of the electrolyte mixture. "Using these simulations, we could better understand the stability and electronic properties of the electrolyte, the most likely reactions occurring at the anode surface, identify any decomposition products, and determine how the resulting SEI layer can stabilise the lithium anode," says Prof. Spencer. "If no gaseous (expanding) products form and the SEI is stable and inert after that initial contact and surface reaction, it's likely to enable safer cycling of the battery. Our calculations can also provide estimates of the ionic conductivity and the oxidation and reduction potential of the electrolyte components, which can determine their effective working voltage range in a battery."

OUTCOMES

"Although we're only modelling 100-200 atoms at a time, the calculations are very computationally intensive, as we're making and breaking chemical bonds – if I was running this on my desktop computer, it would still be calculating long after I'd retired. But with Pawsey facilities we've been able to perform tens of thousands of calculations to study hundreds of electrolyte components and their many complex interactions and reactions with the lithium metal surface at a range of operating temperatures."

Prof. Spencer's modelling work has confirmed that a new class of ionic liquids based on boronium cations are viable candidates for rechargeable lithium metal batteries, showing high electrochemical and thermal stability as well as forming a relatively inert SEI layer. Insights gained into their SEI formation and structure at a molecular level are now being used to guide the experimental development of safe, long-life, high-capacity rechargeable lithium metal batteries that could power the transport of the future.



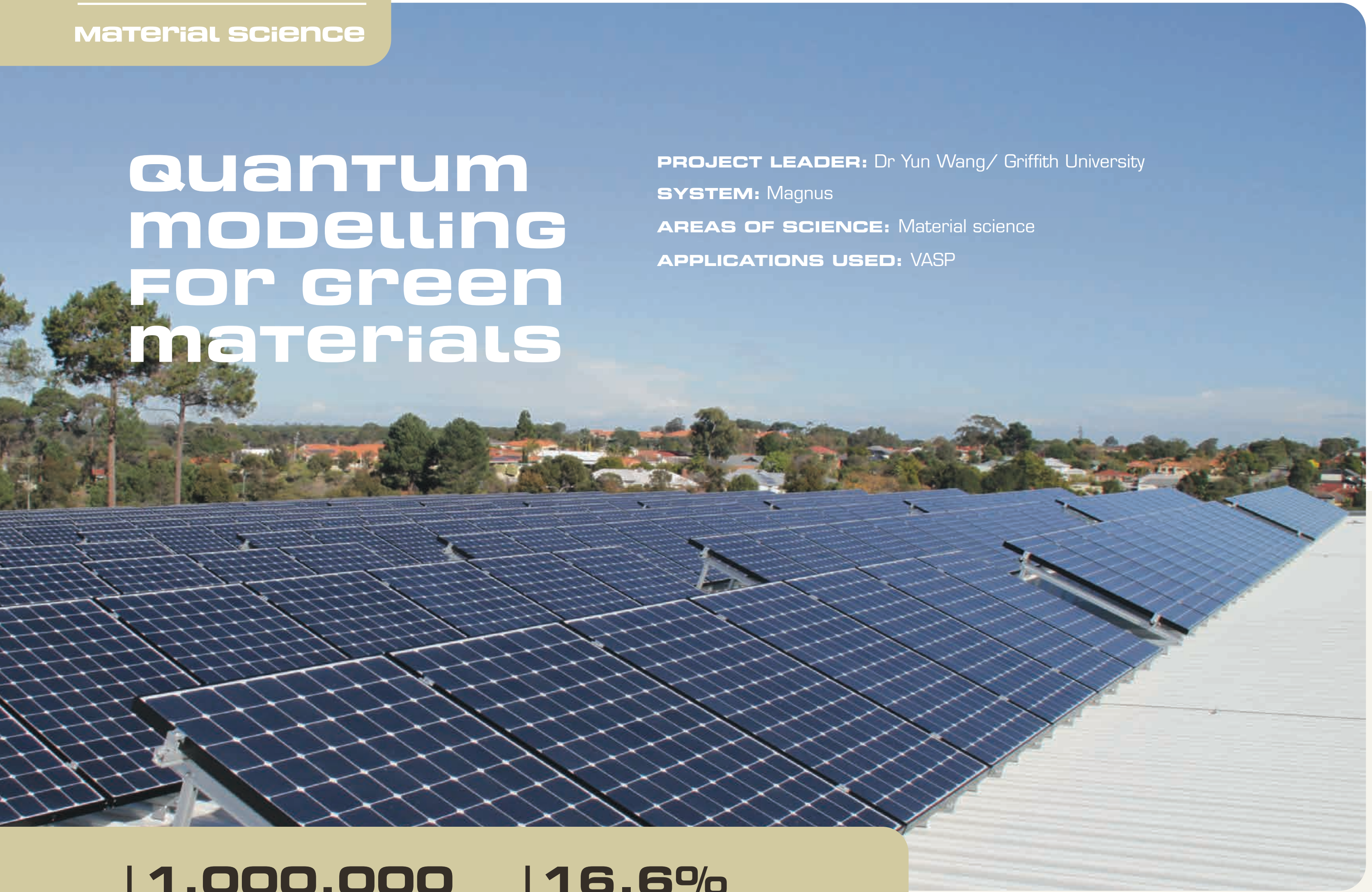
AIMD simulation of the reaction between borate anion electrolyte molecules with a Li(001) anode surface. (Image created by Dale Osborne and Jonathan Clarke-Hannaford)

CASE STUDY 7

Material science

Partner Institution:

Griffith UNIVERSITY
Queensland, Australia



Quantum modelling for green materials

PROJECT LEADER: Dr Yun Wang/ Griffith University

SYSTEM: Magnus

AREAS OF SCIENCE: Material science

APPLICATIONS USED: VASP

1,000,000
core hours Magnus supercomputer

16.6%
quantum dot power efficiency achieved

In 2020, an estimated 24 per cent of Australia's total electricity generation came from renewables, but each state is moving to make renewables their main energy source before 2050. A changing energy industry needs new materials to keep power prices low. Dr Yun Wang's team at Griffith University is probing how atoms behave when absorbing energy to find these new materials.

THE CHALLENGE

Dr Wang models how green materials work at an atomic level. Mathematical models show how atoms change their behaviour when they receive energy. How well atoms can absorb energy changes solar cell efficiency and hydrogen production, two of Australia's biggest future industries.

One of the most promising solar cell technologies available on the market is quantum dots. These are particles over 1,000 times smaller than the width of a human fingernail. Quantum dots absorb more sunlight than any other material. Dr Wang wanted to understand how this material behaved to improve it.

"Quantum dot materials are promising solar performers. We collaborated with groups at the University of Queensland synthesising these quantum dots, modelling how they worked at an atomic level. This understanding would help improve the molecular designs of these dots," says Dr Wang.

Atoms and energy behave in ways predicted by quantum mechanical laws. This means dealing with large amounts of probabilities, variables, and uncertainty. Even a high-powered laboratory computer would struggle to simulate this well enough to find new materials. To overcome this challenge, Dr Wang used Pawsey's Magnus supercomputer.

THE SOLUTION

The team partnered with Pawsey to do their quantum simulations. They were able to narrow down their research to a handful of specific green materials, based on their efficiency and cost.

"We used over one million core hours through Pawsey. We could create oversimplified models to study this mechanism with our own resources, but only a supercomputer can give us a more realistic model. So Pawsey was vital for exploring these mechanisms," says Dr Wang.

The energy mechanics Dr Wang modelled for quantum dots was also applied to different green energy materials. This included finding cheaper versions of the rare metals needed for hydrogen production.

"Atomic modelling is time-consuming. We need a supercomputer to help us solve the Schrödinger equation. This equation is central to the quantum behaviour of atoms and helps us find better green materials."

OUTCOMES

Dr Wang's team calculated the power conversion of quantum dots in solar cells. They found quantum dot power efficiency was 16.6 per cent. This shows they're the best sunlight harvesters in the solar cell industry. The team are now testing cheaper alternatives. Y Dr Wang un was also able to model replacements for expensive platinum in hydrogen production.

Platinum speeds up hydrogen gas production, saving time and energy. But platinum is expensive at nearly $1,300 an ounce. Nickel within an organic molecular framework could become the future of cheap hydrogen fuel. This will help Australia meet its National Hydrogen Strategy to become one of the top hydrogen producers in Asia by 2030.

"It's difficult to synthesise these kinds of precise materials. Our future research considers how solvents and different catalytic conditions affect energy conversion efficiency. This will be a more complicated model, so we'll continue to need supercomputing resources," says Dr Wang.

"If we want to develop new green energy, we need to find cheap materials with better performance."

Dr Wang's research is contributing to Australia's clean energy future. Thousands of cheaper, more powerful quantum dots will line the surface of solar cells. Meanwhile, Australia's hydrogen industry will move from platinum to cheaper metals like nickel. Materials research has the power to transform Australia's energy industry. High-performance computing is vital for these material explorations. Better green materials will help create a safer, more sustainable world.

CASE STUDY 8

AGRICULTURE

Partner Institutions:

The University of Western Australia

CHARTING HOW WE FEED A FUTURE WORLD

PROJECT LEADER: Prof. David Edwards and Mr Philipp Bayer, University of Western Australia

SYSTEM: Magnus, Topaz, Zeus

AREAS OF SCIENCE: Bioinformatics

APPLICATIONS USED: MAKER, Samtools, Velvet, VelvetOptimiser, SOAPaligner

Climate change is a threat to national and global food security, with generally decreased agricultural water supplies and rising temperatures affecting major agricultural yields including staple crops like canola or bread wheat. This has serious implications for Australia's own food supply as well as its export market. Australian canola exports are forecast to increase by 42 per cent in 2020–2021, which would represent AU$1.35 billion to the Australian economy.

To help transform crop practices and develop more resilient crop varieties, we need to better understand the genetic make-up and breeding capabilities of these crops. Working under Prof. David Edwards, Mr Philipp Bayer and a team at the University of Western Australia are using genome sequencing to speed up the time it takes to develop, test, and cultivate new crops to feed us into the future.

THE CHALLENGE

Mr Philipp Bayer and the Applied Bioinformatics team under Prof. David Edwards at the University of Western Australia are tackling the challenge of how to feed a future world, using next-generation genomic sequencing technology and bioinformatics.

This research involves coming up with ways to breed plants in our changing climate and develop new plants that are less prone to problems caused by a changing climate. But conventional plant breeding and cultivation processes — something that has historically taken decades — will not be enough to help feed a growing global population. The Applied Bioinformatics team is aiming to make it possible to develop new plant cultivars within a matter of years.

The team focuses on large plant genomes such as wheat, brassica, and chickpea. As part of this work, they develop custom algorithms and analysis pipelines in order to gain a greater understanding of the complex plant genomes and their expressions.

Looking at thousands or hundreds of thousands of individual plant genomes at the same time generates massive amounts of data and takes time to analyse. Each plant genome must be pulled apart and compared to the reference genome to determine how variants might transform the gene.

THE SOLUTION

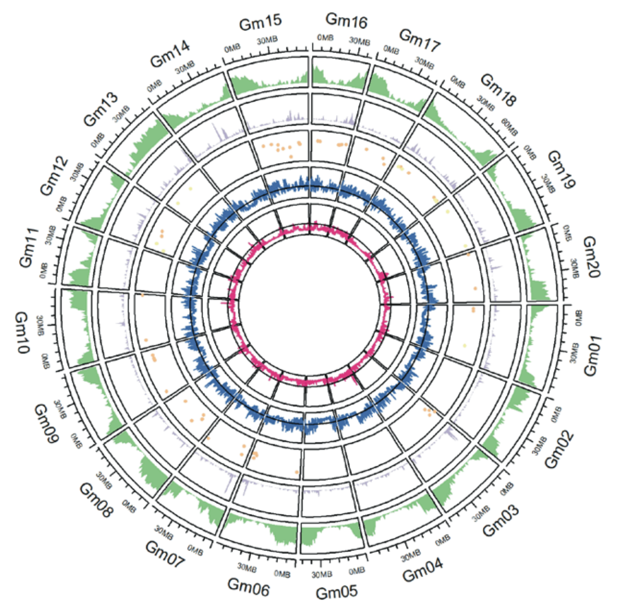
Supercomputing is critical to Mr Bayer's research because of the scale of data he analyses and the complexity of the work.

"We're working with huge amounts of data — in the hundreds of terabytes of data. Just five years ago, we wouldn't have been able to do the work we're doing today. Supercomputers and this scale of computational power are essential to things like our genome assembly and pangenomics work," Mr Bayer says.

The team recently finished a project with 1,000 genomes which generated a few hundred terabytes of data. In collaboration with Prof. Rajeev Varshney at the International Crops Research Institute for the Semi-Arid Tropics in India, the team is looking to expand its work even further through a dataset of 30,000 chickpeas, which is going to generate approximately one petabyte of data.

4,900,000 core hours allocated	**5** grants
13 team members	**231** publications

A Circos plot of the soybean genome showing various measurements of selective pressure.



The Applied Bioinformatics team is also deploying machine learning and deep learning techniques to further the impact of their research. Using Pawsey's Topaz GPU cluster, Mr Bayer is able to deploy machine learning algorithms that link genetic variants with traits of interest to farmers and breeders.

"We're working with industry partners to do genomic prediction using deep learning. This involves five million data points per individual, 100 times over — then running machine learning techniques on that data. We run this on Topaz since the GPU compute power allows us to run state-of-the-art deep learning algorithms to achieve the highest prediction accuracy possible."

OUTCOMES

The scale and complexity of Mr Bayer's work have only been made possible by using supercomputers, bringing computations that are normally impossible on a laptop computer down to a few hours of work. This has expanded what his team is able to undertake and opened the door to bigger and more sophisticated projects.

"We've been able to scale our research to the next level. It's now possible for us to look at the genomes of hundreds of thousands of individuals at the same time, which means we can map all the genetic variations of a whole population in just a few weeks."

CASE STUDY 9

GENETICS

Partner Institution:

THE UNIVERSITY OF WESTERN AUSTRALIA

Mapping DNA to protect an iconic Australian species

PROJECT LEADER: Associate Professor Parwinder Kaur, Australian DNZ Zoo

SYSTEM: Zeus, Data Portal, Magnus, Topaz

AREAS OF SCIENCE: Genetics and zoology

APPLICATIONS USED: Juicer, 3D-DNA, W2Rap-contigger, Maker annotations, LAST

The Australian DNA Zoo node has mapped the first-ever genome of Australia's cutest mammalian critter: the quokka, aka Setonix brachyurus. A genome is the library of DNA for an animal or species.

DNA Zoo is a watershed initiative, leading the world in rapid generation and release of high-quality genomic resources. DNA Zoo Australia at The University of Western Australia is the Australian node of the global project.

These efforts will promote and enable evidence-based decision making in all aspects of biodiversity and environmental research, policy and operational outcomes.

Credit: Rottnest Island Authority

20
team members

200,000
dollars of direct funding

1,000,000
core hours allocated in Magnus

THE CHALLENGE

Mapping an animal's genome can help conservationists understand more about their life code, health and diet, as well as help develop treatments for any diseases that may endanger the animal population.

For quokkas, this effort is doubly important. Rottnest Island in Western Australia, a popular tourist destination, is home to the only remaining high-density sub-population of the quokka. While the mainland quokka populations continue to decline, the Rottnest Island sub-population is well managed and may be the only source of reproductive potential for the species in the future.

Because of their limited habitat, quokkas are very vulnerable to natural disasters. In 2015, the quokka population around Northcliffe, mainland Western Australia was almost decimated, going from 500 animals to 39, following a large bushfire. If a disease or natural disaster were to occur on Rottnest, quokkas might face extinction.

Associate Prof. Parwinder Kaur is the Director of the Australian DNA Zoo node and a passionate advocate of the quokka. She led the team that successfully mapped a chromosome-length sample of the quokka genome.

"This project gives us insurance for the future. Anything can happen any time; any disease could come and wipe out the population. But now we have the DNA code, we're better equipped to deal with future problems," Associate Prof. Kaur says.

Like humans, a quokka's genome consists of approximately 3 billion bases. Each base acts as a building block of our DNA strand. While a coiled DNA strand is nearly ten times smaller than the width of a human hair, it is so long that uncoiled it would reach two metres. Because of this huge amount of data, the DNA Zoo team used the Pawsey Supercomputing Centre to process and store the quokka genome.

"There's no technology that can accurately read such a long information train, so we have to chop it down. We chop it into millions of pieces. Then we read each piece hundreds of times to make sure each letter is correct," Associate Prof. Kaur says.

THE SOLUTION

By cutting up the genome into many smaller segments and parallel processing it in 3D, the DNA Zoo team was able to accurately read an entire quokka genome. They were interested in creating a chromosome-length copy of the quokka genome for a more meaningful handle on the biology.

Chromosomes are major segments of your DNA, inherited from your mother and father. Humans have 23 chromosome pairs, while quokkas have 11. By mapping a chromosome-length strand of DNA, the DNA Zoo team were able to map the first quokka genome in history.

"Mapping this genome, we were working from a blank slate. We had no other quokka DNA to compare it to." Associate Prof. Kaur says.

To accurately read this chromosome, each DNA base had to be read 50-to-100 times to make sure it was accurately recorded. Pawsey systems were a vital part of processing these multiple readings of the huge DNA strand.

"To accurately read a single animal's genome, we're looking at roughly 172 GB of data. On a desktop computer, a single run of this data may have taken years. With Pawsey, it was close to a week. It's not just checking the data, there are a lot of mathematical equations and comparisons to other genomes that need to be made to ensure accuracy. That's why Pawsey was critical to this project," Associate Prof. Kaur says.

"The DNA Zoo team uses Pawsey to process and store the data."

OUTCOMES

The quokka joins a list of 118 mammal species with their high-quality genomes mapped by DNA Zoo. Mapped genomes are used by local conservation teams to protect vulnerable species, to guide population management and breeding strategies.

With humans occupying an ever-increasing portion of the Earth, species like the quokka are pushed to smaller habitats or forced into human spaces. This can make the species vulnerable to extinction. With many mammals and insects, occupying human spaces may also mean more diseases can be spread between humans and animals. We have seen the devastating effect this can have with the SARS-CoV-2 coronavirus.

"We have a duty to protect these species but it's also in our own best interest. If we can keep their populations healthy, we protect ourselves from zoonotic diseases. By mapping their genomes, we may also discover superpowers these animals have that can be used in our own medicines," Associate Prof. Kaur says.

"Quokkas, for example, can halt their pregnancies when the environment doesn't suit having a joey. Understanding the genetics behind that could help our own reproductive science."

As the world's climate changes and more species than ever face the threat of extinction, projects like DNA Zoo are vital to maintaining those vulnerable animal populations that remain. Pawsey's support is enabling researchers to map dozens of endangered mammal species, helping protect them for future generations.

CASE STUDY 10

STEM

Partner Institution:

AUSTRALIAN DATA SCIENCE EDUCATION INSTITUTE

Students work with big data, survey teen sleep

PROJECT LEADER: Dr Linda McIver, Australian Data Science Education Institute

SYSTEM: Visualisation

AREAS OF SCIENCE: Data science

APPLICATIONS USED: Python

How many hours of sleep do you lose due to your mobile phone each night? The Australian Science Data Education Institute (ADSEI) might have an answer.

ADSEI created a sleep survey on teenage mobile phone use before bed. It asked respondents around the world to rate the hours of sleep they get and how long before bed they put away their phones.

The sleep survey is also upskilling high school students in data literacy and data visualisation. It's overcoming the fear of big numbers and getting students to understand how data can help themselves and their communities.

THE CHALLENGE

The 2018 Longitudinal Study of Australian Children Annual Statistical Report found nearly a third of Australian teenagers with internet access in their bedrooms were losing sleep. The blue light from portable devices, like phones and tablets, suppresses melatonin release. Melatonin is a vital hormone that gets your body ready for bed. Sleep deprivation lowers mood, increases stress, impacts health and decreases academic performance.

Dr Linda McIver is the Executive Director of the ADSEI, a charity that trains kids and teachers in data science.

"I ran a workshop for teachers at Pawsey which Ms Marianne Beattie from All Saints' College attended. She wanted a project for her Year 9s and 10s to explore the impact of mobile phones on sleep. She wanted to combine that with the opportunity to work with big data," Dr McIver says.

Data literacy is the ability to understand and communicate facts to solve a problem. Around 20 per cent of Australians are data literate. We fair better than Europe and Japan, but worse than the USA and India. Poor data literacy costs Australian industries roughly $13.9 billion per year.

"We're working with real data that kids can use to explore data science and build their data literacy skills."

But it's not just students who lack data literacy skills. Linda says few teachers have the knowledge to manipulate big data packages or understand HPC.

"The fear factor is the biggest obstacle. When people hear terms like data science and HPC, they think they're way outside anything they can cope with. The fact is, it's actually really accessible if it's presented in the right way."

THE SOLUTION

Ms Ann Backhaus, Education and Training Manager at Pawsey, had a vision for STEM projects. She wanted to combine real world data with Pawsey's HPC and visualisation expertise to make STEM subjects more engaging to students. She approached Linda to use the sleep project as the focus for a summer internship to build a platform that would support data science projects. They worked with Pawsey Senior Visualisation Specialist, Dr Yathu Sivarajah and visualisation specialists Mr Ali Zamani and Mr Jesse Helliwell to supervise a Pawsey Summer Intern, Mr Thai Nguyen. Mr Nguyen developed the full stack web site to collect data into a database and create visualisations on the fly to allow students to explore the results.

"From a technology perspective, Pawsey is coordinating all the technical expertise and hosting the platform, so we're heavily involved. Our plan is to facilitate ADSEI throughout this process, not just by hosting the platform but by providing guidance," Dr Sivarajah says.

"Pawsey is providing the funding and support to train teachers to use this platform as well," Dr McIver says.

The survey is being used by students to practice visualising big data. Students usually deal with easy-to-visualise data sets. A few dozen points can be easily placed on a graph and outliers are obvious. When hundreds or thousands of points become involved, it's more difficult to read what the data are saying.

The team had Mr Nguyen build a visualisation platform in React. The platform helps students compare different questions in the sleep study, such as participant country and age. This way, students get familiar with using big data without needing an in-depth understanding of coding.

"Python graphing libraries are really difficult for kids to use. We wanted a way for kids to play with the data and go, 'What if I compare this field against that field?' We want them to be able to download and play with the data, to look for relationships and build their programming skills," Dr McIver explains.

The sleep survey isn't just for data training though. Dr Sivarajah says schools can use it to explore different variables, such as the impact of exercise or stressors, on students' sleep patterns.

"You could have students contribute at different times. For example, you could get them to answer questions about how they're sleeping before and after exams, then maybe trial ways to prepare students for test taking by directly involving them in the discussion."

OUTCOMES

So far, over 500 participants, mostly Australian teens, have responded to the survey. It will remain open indefinitely at this point. Dr McIver hopes to get thousands more students contributing to the dataset, so it's a true representation of working with big data.

Dr McIver is also hosting a series of workshops for teachers to train them in using ADSEI data sets in schools. These workshops feature Pawsey staff introducing teachers to the concept of HPC.

"We're planning to do these workshops online to make them accessible right across Australia. And the goal is to run four of them this year."

Pawsey Head of Scientific Platforms, Dr Mark Gray was featured in one of the recent workshops.

"He talked about some of the applications of big data and the teachers were just blown away by it. Just hearing the different ways supercomputing impacts the world is so powerful. A lot of teachers said that was the best part of the workshop."

The data manipulation tools will be used for a variety of future studies for students. They're designed to need little reworking to adapt to new big data sets, so national or international education programs can use them.

"The same workflow can be repeated across studies. We want this platform to be widely accessible, so it can be used across Australia, potentially globally, for school curriculums," Dr Sivarajah says.

Pawsey is continuing to host the ADSEI sleep survey and is working with Dr McIver to empower students with data literacy. By knowing how to look at data, students are preparing themselves for their future careers and giving themselves the tools they need to contribute to their communities.

4
Countries surveyed

80
The oldest person to complete the survey

500
People completed the survey

PROJECTS AND PUBLICATIONS

During the reporting period, 4056 researchers used Pawsey facilities to progress 143 projects in scientific fields spanning chemistry, physics, engineering, bioinformatics, health and the life sciences. Their breakthroughs were shared across the scientific and broader community through 397 publications, a publication rate that has been increasing steadily over several years.

International journals Pawsey researchers have published in include Nature, Nature Communications, Nature Nanotechnology, the Journal of Physical Chemistry, Angewandte Chemie, Physical Review, Astrophysical Journal and Astronomy and Astrophysics.

Over 30 per cent of the publications were from chemistry research investigating new materials for renewable energy and storage, nanotechnology and porous materials. This was followed by radio astronomy and astrophysics papers, accounting for 22 per cent of the total. Physics and engineering in combination account for a further 22 per cent of the publications acknowledging Centre support. Life sciences research enabled by Pawsey continues to increase, and now accounts for almost 10 per cent of listed publications.

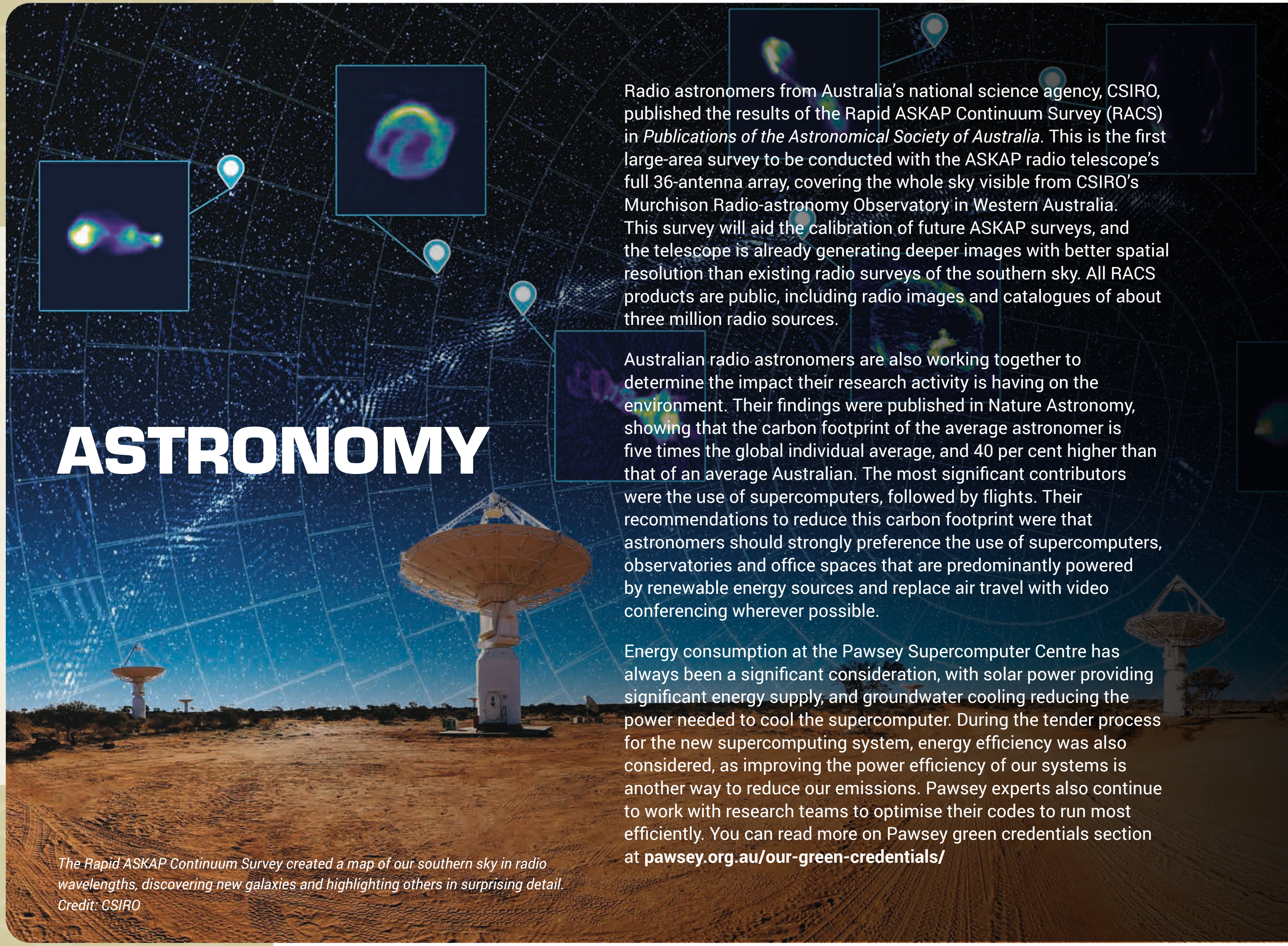
The full list of publications can be found online at **www.pawsey.org.au.**

4,056 researchers

143 projects

397 publications

Examples of research achievements and breakthroughs from this year that were made possible with Pawsey support include:



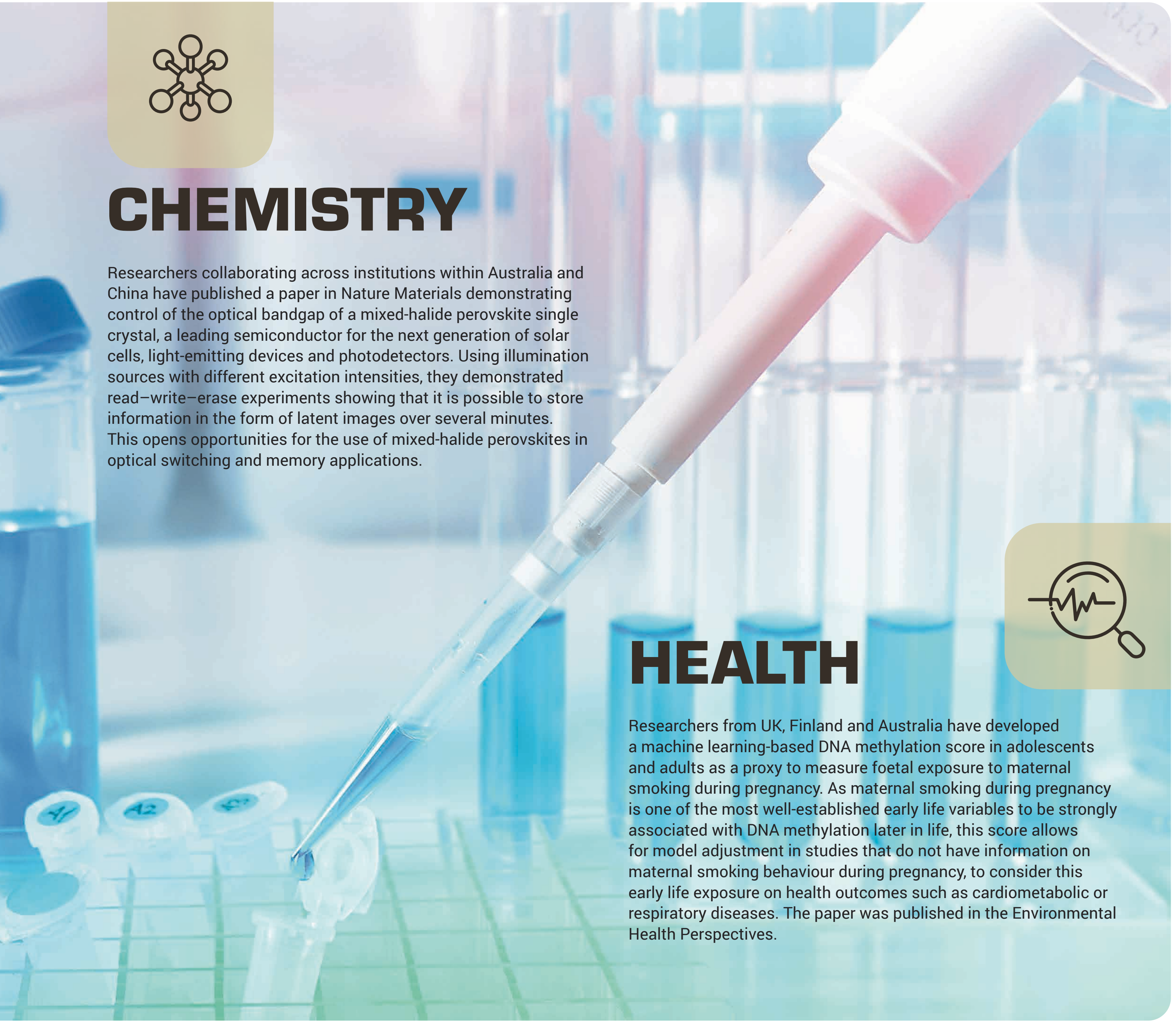
ASTRONOMY

The Rapid ASKAP Continuum Survey created a map of our southern sky in radio wavelengths, discovering new galaxies and highlighting others in surprising detail. Credit: CSIRO

Radio astronomers from Australia's national science agency, CSIRO, published the results of the Rapid ASKAP Continuum Survey (RACS) in *Publications of the Astronomical Society of Australia*. This is the first large-area survey to be conducted with the ASKAP radio telescope's full 36-antenna array, covering the whole sky visible from CSIRO's Murchison Radio-astronomy Observatory in Western Australia. This survey will aid the calibration of future ASKAP surveys, and the telescope is already generating deeper images with better spatial resolution than existing radio surveys of the southern sky. All RACS products are public, including radio images and catalogues of about three million radio sources.

Australian radio astronomers are also working together to determine the impact their research activity is having on the environment. Their findings were published in Nature Astronomy, showing that the carbon footprint of the average astronomer is five times the global individual average, and 40 per cent higher than that of an average Australian. The most significant contributors were the use of supercomputers, followed by flights. Their recommendations to reduce this carbon footprint were that astronomers should strongly preference the use of supercomputers, observatories and office spaces that are predominantly powered by renewable energy sources and replace air travel with video conferencing wherever possible.

Energy consumption at the Pawsey Supercomputer Centre has always been a significant consideration, with solar power providing significant energy supply, and groundwater cooling reducing the power needed to cool the supercomputer. During the tender process for the new supercomputing system, energy efficiency was also considered, as improving the power efficiency of our systems is another way to reduce our emissions. Pawsey experts also continue to work with research teams to optimise their codes to run most efficiently. You can read more on Pawsey green credentials section at **pawsey.org.au/our-green-credentials/**



CHEMISTRY

Researchers collaborating across institutions within Australia and China have published a paper in Nature Materials demonstrating control of the optical bandgap of a mixed-halide perovskite single crystal, a leading semiconductor for the next generation of solar cells, light-emitting devices and photodetectors. Using illumination sources with different excitation intensities, they demonstrated read–write–erase experiments showing that it is possible to store information in the form of latent images over several minutes. This opens opportunities for the use of mixed-halide perovskites in optical switching and memory applications.

HEALTH

Researchers from UK, Finland and Australia have developed a machine learning-based DNA methylation score in adolescents and adults as a proxy to measure foetal exposure to maternal smoking during pregnancy. As maternal smoking during pregnancy is one of the most well-established early life variables to be strongly associated with DNA methylation later in life, this score allows for model adjustment in studies that do not have information on maternal smoking behaviour during pregnancy, to consider this early life exposure on health outcomes such as cardiometabolic or respiratory diseases. The paper was published in the Environmental Health Perspectives.

ENGINEERING

The ability to filter out oil mist and dust is crucial to the operation of equipment from respirators and compressors to advanced machining equipment and metal-cutting tools. Optimising the microstructure of filter media to maximise collection efficiency while avoiding clogging and pressure drops is becoming increasingly reliant on computational simulation of dry and wet filtration processes, but computational fluid dynamics (CFD) simulations demand significant computation effort, especially for fine filters. A team from Curtin University and TU Dortmund University in Germany have developed a dynamic pore network model (DPNM) framework to predict residual saturation and pressure drop during mist filtration, and have shown that the model can successfully predict these critical properties of filtration processes to within 15 per cent of experiment and CFD, but with an overall computational cost 2–3 orders of magnitude lower. DPNM can be considered an efficient tool for the early-stage design and optimisation of better non-woven mist filtration media, providing fast predictions of overall filter performance. The paper was published in Separation and Purification Technology.

Inspiring the scientists of tomorrow, accelerating the science of today

Voices of Science

Geoffrey Harben

Geoff Harben has been an independent member of Pawsey's Board for at least 10 years, predating iVEC's transition into the Pawsey Supercomputing Centre. A lot has changed for Pawsey's infrastructure and operations in that time, but Geoff has been embedded in the ICT industry for even longer, witnessing the transition from punch-card programming to the new era of diamond-based quantum computers over 45 years. Having worked through technical, analysis, sales and business management positions at Fujitsu, then progressing to major consulting firms in the ICT space, while being a long-term council member for the Australian Information Industry Association, he has a wealth of knowledge of the ICT industry in Western Australia, and how government engages with and develops it. Corporate management in the ICT industry is where all of his career experience intersects.

"I was invited to join the board by the late Mal Bryce," remembers Geoff. "We were just starting to raise funds to develop iVEC into supercomputing facilities of a decent scale in a central data centre. I thought it would be a really exciting opportunity to work on that – I wasn't wrong! We've come such a long way since then. We're about to install the southern hemisphere's most powerful supercomputing facility, and more importantly, we're enabling genuine research, research that just wasn't possible before, right across Australia."

Part of Geoff's role on the Pawsey board concerns risk management. He notes: "We've got a different risk profile to many other organisations, as an unincorporated joint venture largely dependent on State and Federal government funding, and on government-funded authorities in the research space. So, it's incredibly important that Pawsey can deliver against the outcomes that are expected of us, and continuously demonstrate a return on that investment."

"It's not enough to work hard and achieve great things, those outcomes need to be demonstrated to the government and the community."

Geoff considers it a responsibility of the Board to lift Pawsey's profile, something he's been doing since the start. "I'm very proud that we've moved from being a small computing facility on the west coast with a vision, to being a world-leading facility, and a critical part of the nation's infrastructure in High Performance Computing (HPC). We got here because we've continuously demonstrated our capability and responsibility, and our good governance in managing a supercomputing environment. There's no question that we are a significant facility now, and our new supercomputer is going to generate enormous value for Western Australia and Australia. But we need to keep demonstrating that."

Ten years in, and the work is ongoing. "We engage with the academic and research community extremely well," says Geoff, "but we still need to better engage with the commercial world to do research. We need to develop a closer working relationship with industry, and with the public service, because that's how to further demonstrate our impact. We're pushing very hard for that as our capability expands with the supercomputing upgrade."

"The world is facing enormous issues – climate change, pandemics – and supercomputing facilities are becoming increasingly important as the world does the research it needs to survive. Government organisations, health institutes, major commercial industries all have a part to play in addressing these challenges, not just academic researchers. So, we need to be working with all of them."

> "It's not enough to work hard and achieve great things, those outcomes need to be demonstrated to the government and the community.
>
> **Geoffrey Harben**
> Board Member

Exploring the solar system without leaving the ground

The Desert Fireball Network uses a collection of fixed cameras across Western Australia and South Australia that continuously take pictures of the night sky, in search of fireballs. When more than one camera detects a fireball, the combination of observations from different locations can be used to work out where the meteoroid came from, where it was headed, and where the resulting meteorite might eventually be found.

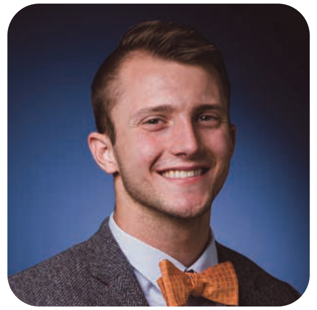
All of that information, and the meteorites themselves, let us piece together the composition and working of the inner solar system, without ever having to leave Earth. This is the team doing it.

DR HADRIEN DEVILLEPOIX

PLANETARY SCIENTIST, SCIENCE LEAD

Hadrien wanted to use his skills in data science and computing for something more 'fun' than the corporate world. Through an internship at the Paris Observatory, he started developing computing tools to process spacecraft data to discover new asteroids. It led him to a PhD in planetary science with the Desert Fireball Network, developing data reduction tools to make sense of the more than 50 terabytes of data the early network collected every week.

He now helps the rest of the team process the massive amounts of data coming out of the expanded camera network, turning data points into meaningful scientific observations. A particular focus is using fireball trajectories to extrapolate back to where in the solar system the meteoroids came from.



MR PATRICK SHOBER

PHD CANDIDATE

An excellent science teacher in school put Patrick on the path to study physics and engineering, and then an engaging lecturer introduced him to geology. His first research project involved studying meteorites, looking at the alterations occurring at the rock surface to understand the forces that have acted on them and decipher their history. His current research is fundamental to the work of the Desert Fireball Network, doing the orbital modelling to work out where fireballs may have come from in our solar system.

"There are only two ways to figure out what asteroids are made of – you can look at them through a telescope and infer a few things about them, and you can study meteorites on the ground like a geologist can study any rock to determine its composition and history," Patrick explains. "But you don't know where it came from."

"By tracking enough fireballs back to their origins, and finding enough meteorites on the ground to study, you can make connections between the asteroids and the meteorites they generate, and start to work out the composition of our inner solar system."

RESEARCH WITH SUPERCOMPUTERS – TRACKING METEOROIDS BACK THROUGH SPACE

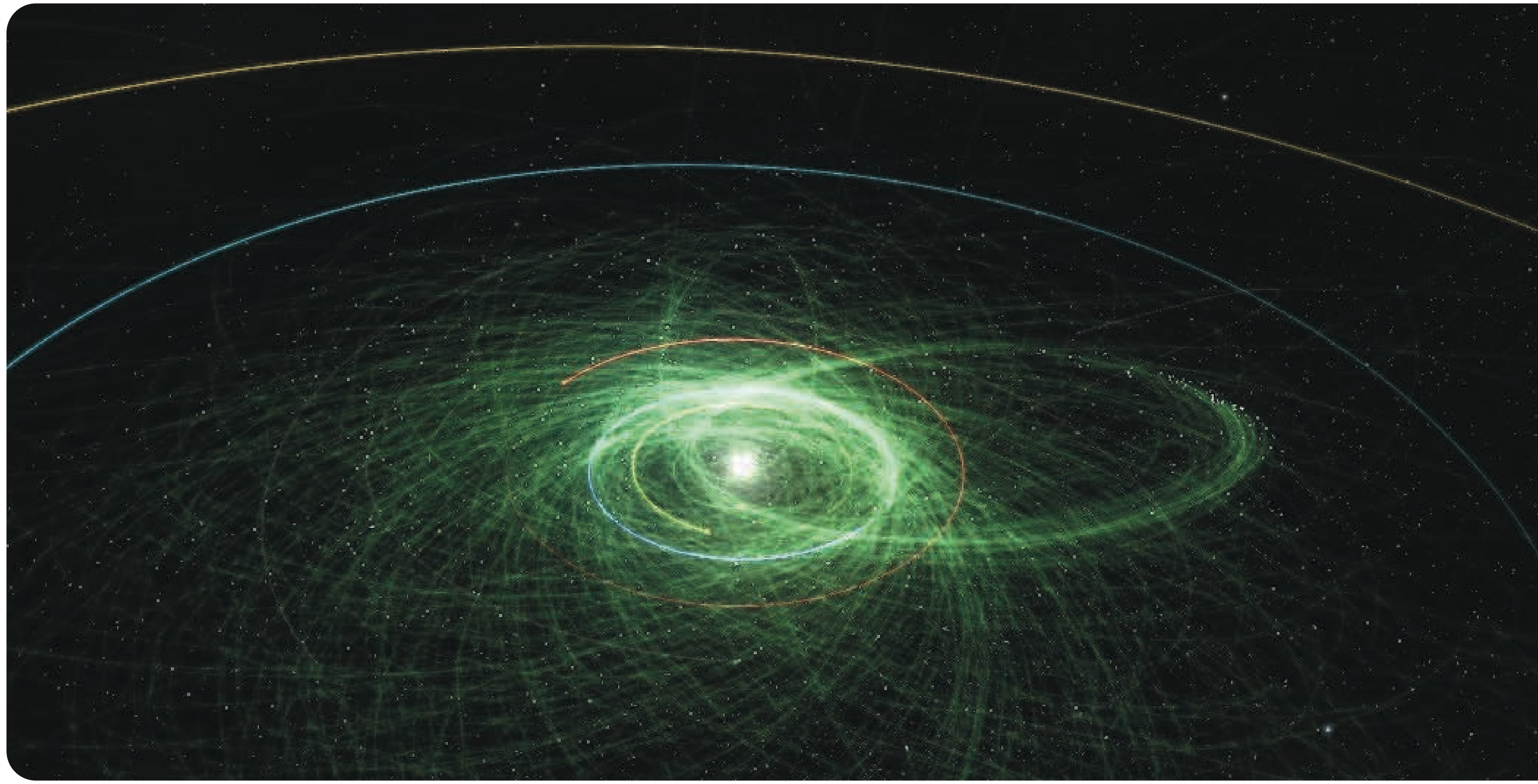
Extrapolating the track of a fireball backwards through space to work out where it came from is complicated by the fact that everything in our solar system is constantly moving. With even the best measurements of our planets and their orbits, there are small uncertainties in their exact trajectory and speed. It limits our view back through the solar system to about 60 million years – earlier than that, we can't tell where the planets were with enough certainty. This only covers the recent geological history of our solar system, as most meteorites that make it to Earth have travelled from their ejection point for millions to tens of millions of years.

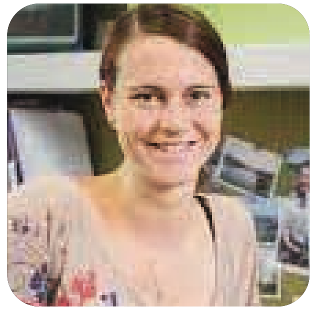
"Because we can't compute an exact trajectory back to a definite point in space and time, we use supercomputing to calculate lots of realistic possibilities," explains Patrick.

"We create lots of copies of our meteoroid trajectory, representing the range of uncertainties we have in its measurement, and run them backwards through multiple solar systems spanning the range of uncertainties we have in planetary movement."

Small differences in the starting observations can cause big differences once you extrapolate back, but by calculating thousands of potential orbital trajectories back through time, the probability that a meteorite can be traced back to a specific source can be calculated.

"The best supercomputer in the world can't give you the exact answer for a meteorite trajectory," says Hadrien. "But by calculating the probabilities on multiple trajectories, we can narrow it down to the most likely origin."



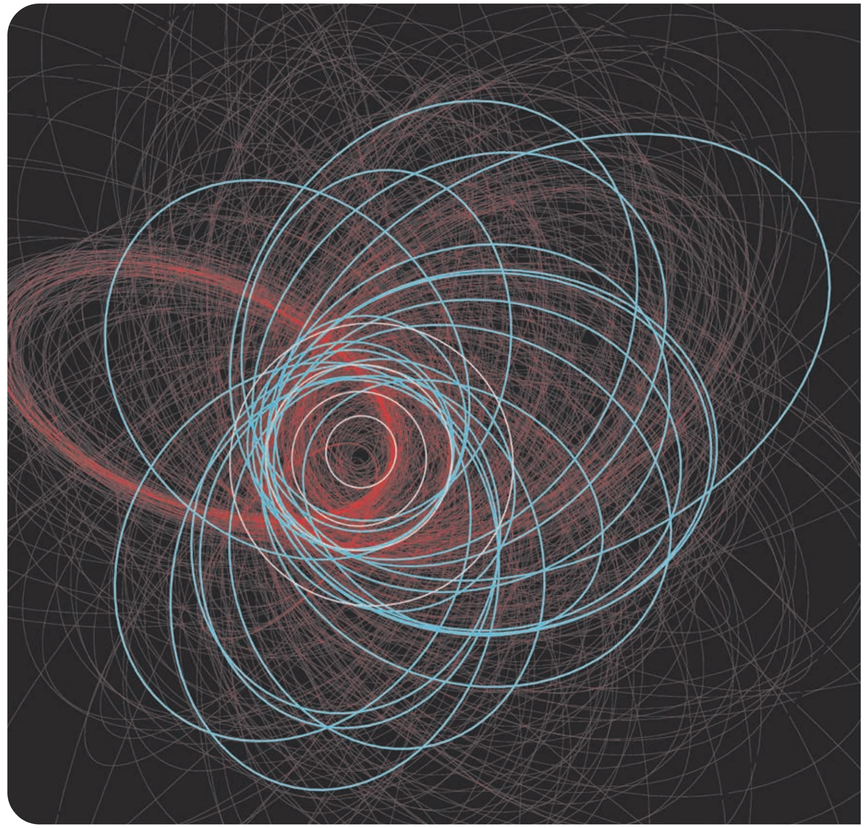


DR ELLIE SANSOM
GEOPHYSICIST, PROJECT MANAGER

Ellie was inspired by an astronaut that visited her school, who told her that if she wanted to be an astronaut, the best thing to do was find something you enjoy, and become really, really good at it, as space exploration needs experts in every field. With an interest in rocks and geology among many other things, Ellie decided to study geophysics.

"I did my PhD with the Desert Fireball Network group. Although I'd studied geophysics, I'd always been keen on space and astronomy, and this blends the two. We're getting so good at astronomy that we can apply geological techniques to other planets in our solar system, and study space without leaving Earth."

Ellie now manages the Desert Fireball Network, creating new international opportunities as it expands into the Global Fireball Observatory, with even more cameras in multiple countries covering even bigger sections of sky. On the science side, she analyses the atmospheric dynamics of fireballs and meteoroid entry.



RESEARCH WITH SUPERCOMPUTERS – WORKING OUT WHAT IS IN A FIREBALL

A surprising amount of information can be gleaned about the meteoroid falling to Earth just from the observations of the fireball burning up in our atmosphere. "The behaviour of the meteoroid is based on very fundamental equations describing atmospheric drag," says Ellie, "and we can work out how fast it is falling. The major problem is that we don't know the object's size, mass, shape, or if it's strong enough to avoid breaking into fragments."

Ellie uses supercomputing to create millions of simulated meteoroids with a range of different characteristics, and then calculates how each responds to atmospheric drag, comparing which ones best match the observations of the fireball over time. By matching observations with the myriad of possibilities at every step of the fireball trajectory, she can calculate the initial size and shape of the meteoroid on atmospheric entry, any pattern of fragmentation with estimates of rock strength, and the size of any final meteorite that makes it to the ground.

Knowing the likely size of a fireball as it descends is crucial to predicting where it will land and if it's big enough to find, so Ellie's calculations are used to prioritise which fireball sightings may turn into a successful meteorite recovery. Even without recovering a meteorite, the information gleaned about each meteoroid can be combined with its calculated orbit to continue building our picture of the inner solar system.

MR MARTIN CUPAK
SOFTWARE DEVELOPER

Martin's forte is computer systems, making the connections between hardware and software work, especially for fully automated systems. From developing software for instruments on satellites, he turned to automating camera systems, developing the control software for film-based automated cameras and eventually the digital cameras that became the basis of the early Desert Fireball Network.

From building and installing the cameras, Martin is now providing technical support for groups running camera networks across the world as part of the Global Fireball Observatory. He points out: "The observatory is a network of devices and instruments collecting data, all working independently for months at a time. You can't do that only with scientists – you need the technical people who can operate the network and keep it running too".

Martin also coordinates the data transfer, processing and archiving of all of the digital photography at Pawsey, as images are periodically collected from the remote cameras. When fireballs are identified, Martin refines the atmospheric models used to predict how local weather conditions alter meteoroid trajectory, allowing more accurate prediction of the final meteorite impact sites.

DR MARTIN TOWNER
FIELD OPERATIONS CHIEF

Martin grew up with the Thunderbirds and Star Trek, but science fiction rapidly became science fact in his career. After studying physics and working in industry building satellites, Martin returned to research, analysing data from the Cassini–Huygens space research mission to Saturn and developing instruments for several missions to Mars.

His skills in designing instruments that work autonomously in space was turned to developing digital cameras to work autonomously in the desert in 2012, and Martin has been running the Desert Fireball Network's field operations since then. "I do a lot of the maintenance work, the field work, and the search and recovery operations when we identify a meteorite impact location," says Martin.

While COVID-19 restrictions on travel are hampering fieldwork, Martin also works on the inner atmosphere trajectory modelling of fireballs, trying to pinpoint their final impact sites. The list of meteorites to search for at a later date is steadily growing.

RESEARCH WITH SUPERCOMPUTERS – THE PATH TO THE GROUND

Tracking a meteorite to the ground isn't as easy as following the path of the fireball. As the meteoroids burn up in our atmosphere, they slow down to the point where they no longer glow, disappearing from view. Martin Cupak explains: "During the 'dark flight' section of the meteoroid's trajectory, the last 20–30 km to the ground, we have to consider how the wind will affect its fall".

Data from meteorological stations all over the world is assimilated in the US and put into a global atmospheric model, so Martin uses that as a starting point. "Once we know when and roughly where a meteorite has fallen, using Pawsey we access the appropriate 500 x 500 km 'chunk' of the global model, and run it with the local meteorological data from that time period to get a more precise understanding of the winds at the time, down to a resolution of 1 km. It's an important part of calculating the final meteorite's impact site, because winds can easily blow it 5 km off course."

REAL WORLD SOLUTIONS

The team at the Desert Fireball Network are continually adding to our understanding of the inner solar system, and have classified and calculated orbits for around 2,000 fireballs to date. Seven meteorites have been recovered on the ground, with several tens of meteorites still out there waiting for the next field trip to be discovered.

Director, Australian SKA Regional Centre, ICRAR-UWA and CSIRO

Area of Science: Astrophysics

Deciphering the Data of the Universe

Dr Karen Lee-Waddell is studying how the Universe changes over time, using data from the ASKAP radio telescope. At the same time, she's setting up the Australian SKA Regional Centre so future scientists will be able to do the same thing on a much larger scale – using data from the world's most powerful telescope, the Square Kilometre Array (SKA), to further our understanding of the Universe.

ABOUT DR KAREN LEE-WADDELL

Karen has already juggled two successful careers, which have positioned her uniquely to move her research to the furthest reaches of the Universe using the world's most powerful telescopes. While studying physics and astronomy in Ontario Canada, she first heard about plans for the SKA during her PhD. "I'd already fallen in love with research, radio astronomy, and travelling around to work with the best telescopes in the world," she reminisces. "So, hearing about this amazing world-wide research project that would bring nations together in the common pursuit of science and understanding the Universe – I wanted to be a part of that."

As soon as she finished her PhD, she moved to Sydney to work with ASKAP at CSIRO, one of the precursor telescopes to the SKA. By mapping the hydrogen distribution across sections of the sky, her team is examining how galaxies form and evolve in much greater detail than ever before. They are already working with hundreds of 'new' galaxies that were unresolved blurs in previous sky surveys. As she progresses in her career, Karen continues to lead the Widefield ASKAP L-band Legacy All-sky Blind survey (WALLABY): a neutral hydrogen survey of 75 per cent of the sky that will detect hundreds of thousands of galaxies.

While still at university, Karen also joined the Royal Canadian Navy as a Logistics Officer. Throughout her study and research she continued as an active Reservist, serving with units in Canada and then in Australia, as a foreign exchange officer. "During the day I'd be doing astronomy, but on certain evenings and weekends I'd be dealing with personnel, military support services, and general logistics for the Reserves."

WHAT DREW HER TO SCIENCE?

Growing up in a big Canadian city, the light pollution meant Karen never really saw the night sky. She remembers: "My sister was driving me home from a visit to a more rural area one night and her car had a sunroof. She opened it and started pointing out a few constellations. I was about nine years old at the time and I had NEVER seen the night sky like that – I fell in love. I started reading astronomy books and never looked back."

RESEARCH WITH SUPERCOMPUTERS

Karen relies on Pawsey supercomputing to process and store all of the raw data collected by ASKAP, turning it into useable information for astronomers. "We're almost at full science operation of ASKAP, generating over four terabytes of data per hour," she says. "When you're working with one terabyte cubes of processed astronomical data, you need much bigger computers than conventional laptops."

REAL WORLD SOLUTIONS

Owned and operated by CSIRO, ASKAP and the research it enables isn't just a stepping stone to the full SKA telescope, it's also building a community that is developing the data processing capabilities that will be required for SKA-sized datasets.

"The SKA will serve up more astronomical data than anyone has ever dealt with, 600 petabytes per year. It's one thing to process and store it, but it's another to get that data to the scientists so they can do their research," explains Karen.

SKA Regional Centres are being created to allow scientists to access and handle the eventual SKA data, storing processed datasets together with the analytical tools and algorithms needed to do research at the required scale. Karen is the Director of Australia's SKA Regional Centre, with the task of creating this $63 million facility for Australian researchers.

"The goal is to make sure make sure our scientists can actually do research with the corrected, calibrated, and imaged datasets that will be coming from the SKA telescopes. With Pawsey as a partner, we can test workflows and tools with ASKAP data as we scale up. In learning how to deal with full science operation for the WALLABY survey and other projects, we're learning how to deal with eventual SKA research."

Karen is combining her radio astronomy and supercomputing expertise with 13 years of military leadership and logistics training to make the Australian SKA Regional Centre a reality over the next ten years.

"I remember going into work each day, knowing that I could be the first person to set eyes on a new galaxy. The discovery aspect of my work really drove me as a researcher. But now I'm working on the logistics so others can do that. With SKA Regional Centres, researchers will be able to see galaxies that have never been seen before and understand things no-one has understood before. My job now is to help entire teams of scientists use the SKA to figure out what is 'out there' and answer some fundamental questions about how the Universe works. That's still my passion."

Senior Research Fellow and Senior Lecturer, School of Computing and Information Systems, The University of Melbourne

Area of Science: Noisy intermediate-scale quantum technology, quantum algorithms

DESIGNING THE PROGRAM TO DEMONSTRATE QUANTUM ADVANTAGE

Current quantum computing technologies have limitations: quantum bits (qubits) don't last anywhere near as long as classical transistors do, units of information are not transferred with absolutely 100 per cent accuracy, and noise can accumulate in circuits before an algorithm can complete, drowning out the answer. Dr Casey Myers is designing quantum algorithms to specifically work within these limitations – to demonstrate that our first generation of quantum computers can still do useful calculations faster than classical computers.

ABOUT DR CASEY MYERS

Casey was studying physics at the University of Queensland when part of Australia's first research centre for quantum computation and communication technologies was being established there. It put him in the right place at the right time to see the first steps being taken in the world of quantum computing, and led him to complete a PhD and then eight years of postdoctoral research in the field.

But the realities of life, location and occupation prompted a move out of academia and into 'real world' jobs. Theoretical physicists can use their skills in a wide range of areas, and Casey worked in seismic imaging for the oil and gas industry, and as a finance and risk analyst in the banking sector while continuing to follow research progress in quantum computing from afar. A chance to work in a quantum computing start-up company then gave him the opportunity to return to the field and catch up on the latest developments, letting him move back into full time research.

"When I left academia for industry in 2015, quantum computing was moving into a really exciting phase," recalls Casey, "but machine learning hadn't really made a mark on physics research yet, and access to small quantum computing systems meant caging favours from the experimental physics team working in the basement. Now I've come back to a research world where access to physical quantum computers exists, and cloud access to a small diamond-architecture quantum computer and HPC-scalable quantum simulator will be available at Pawsey next year. I can now log in and put my quantum computing algorithms into the queue to test how they will perform on actual quantum-based hardware. It is a fantastic capability that is really progressing our knowledge of what quantum computers may eventually be able to do."

WHAT DREW HIM TO SCIENCE?

"In high school, I just started reading science books. I came across one of Stephen Hawking's books, and it led me to ask questions about the Universe in general, and find out what theoretical physicists like him actually do. By the time I finished school, I was absolutely sure I wanted to study physics, because that was the way to get the questions answered that I was really asking."

RESEARCH WITH SUPERCOMPUTERS

'Quantum supremacy' – demonstrating that a quantum computer can solve a problem faster, more efficiently, or more accurately than a classical computer – has already been achieved. But Casey notes: "These demonstrated problems are fairly contrived, they're designed for the way quantum computers work. To make quantum computing a reality we need to demonstrate 'quantum advantage' – solving a useful, practical problem faster or more accurately than any classical computer."

Casey is using Pawsey supercomputing to run his experimental quantum algorithms on the Quantum Brilliance Quantum Emulator, which uses classical computing to simulate how a quantum computer would operate. The aim is to design the algorithm so it can deal with the current physical limitations of quantum computing hardware, primarily the introduction of noise and errors during the time it takes to perform a calculation.

Casey is working on two complementary approaches to designing quantum algorithms: reducing the size of the algorithm so it can execute before being overwhelmed by noise, and incorporating quantum error mitigation techniques to limit the introduced errors and recover the quantum advantage that would otherwise be lost to noise.

REAL WORLD SOLUTIONS

The problem Casey is working to solve is drawn from the world of finance – working out the right price for an option on the stock market. "Using the quantum emulator I can see how these programs would run, to get them working and see if they can perform faster than a classical computer."

Once Quantum Brilliance's diamond-based quantum hardware is installed at Pawsey next year, he'll be able to queue up his algorithms over the cloud and test them on the real thing. "That's the ultimate goal, to see these algorithms run on a quantum computer, and beat a classical computer with a practical problem. Even if we only show a low-level polynomial improvement in speed, it's significant. If we can cut the run time or increase the accuracy of even simple problems, the intractable problems that have too many calculations for even our supercomputers to handle start becoming approachable."

Professor, Oceans Graduate School, UWA Oceans Institute

Area of Science: Coastal engineering and physical oceanography

Calculating coastal interactions to protect our land girt by sea

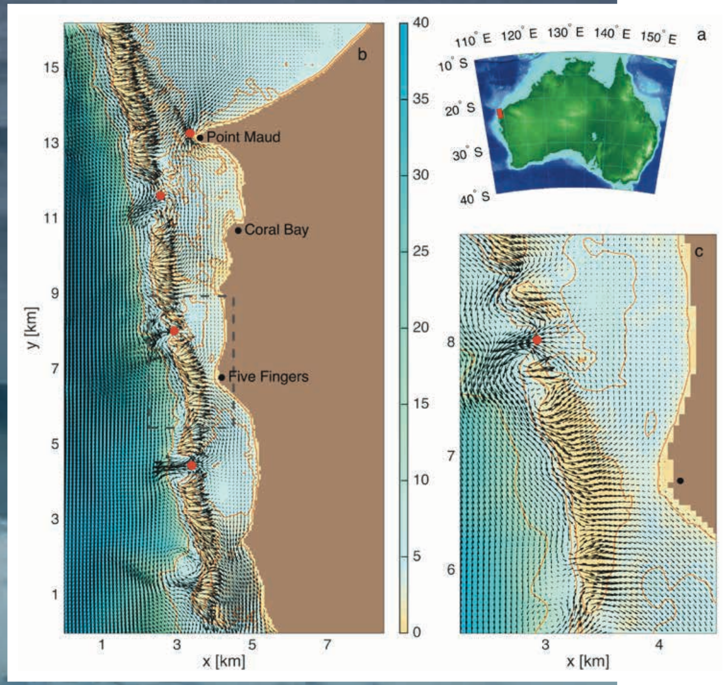
Professor Ryan Lowe focuses on understanding and predicting coastal processes. By defining the physics of ocean waves and currents and how they interact with complex coastlines, his work has applications ranging from predicting beach erosion and coastal flooding to mapping nutrient flows or pollution spread in the environment.

ABOUT PROFESSOR RYAN LOWE

Ryan studied mechanical engineering at the University of California, but quickly discovered an interest in physical oceanography and ocean science more generally, completing a PhD in Civil and Environmental Engineering at Stanford University. His early research focused on understanding how coastal water flows driven by waves interact on local scales with the roughness of many coastlines, like coral reefs, seagrass meadows or mangrove forests. Since commencing at the University of Western Australia in 2007, his research has increasingly focused on improving predictions of wave-driven processes and hazards in coastal environments. As part of this work, Ryan also identifies strategies to protect coastlines from flooding and erosion by incorporating both natural features of ecosystems and coastal engineering structures.

WHAT DREW HIM TO SCIENCE?

"I grew up on the beach," admits Ryan. "The ocean was a part of everything I did growing up." Enjoying science and mathematics in school, Ryan gravitated to engineering, and his interest in environmental flows and fluid mechanics steered him into physical oceanography. "I wanted to understand the physics of the ocean."



Ningaloo circulation

RESEARCH WITH SUPERCOMPUTERS

Water motion in the ocean occurs over a vast range of scales – from large-scale currents in ocean basins, to coastal current systems, down to the small-scales of ocean wind-waves and turbulence," explains Ryan. "So, understanding transport and mixing processes in the ocean requires models that can capturing ocean physics over a range of spatial and temporal scales that vary by application. That range of spatial and temporal scales makes predicting ocean flows inherently challenging. It's not practical to measure everything out in the field, so we often create large numerical models using supercomputers that are usually supported by field or laboratory measurements of ocean flows."

Experimental measurements of local interactions in the coastal zone, whether on a beach or in a wave tank, allow Ryan to develop methods to parameterise small-scale hydrodynamic processes that can then be incorporated into larger scale ocean models. Equally, the numerical models are often used to improve the field work: by running models before heading to the coast, Ryan can gain insight into likely ocean conditions and plan the optimal places and frequencies to take useful measurements.

REAL WORLD SOLUTIONS

Ryan's work is allowing coastal ocean models to more accurately account for how coastal topography and habitats modify the coastal environment, and affect processes like beach erosion and coastal flooding. Accurate models are also allowing better predictions of future coastal hazards, incorporating climate change effects such as sea level rise and the changing frequency and intensity of storms.

"Coastal flooding is a worldwide threat that is growing with climate change," notes Ryan, "and 85 per cent of Australia's population lives near the coast. 'Hard' engineering flood defences such as seawalls and breakwaters work but come at significant cost, including to coastal ecosystems and general coastal amenity. We can't build them everywhere."

Ryan's predictive modelling is showing how ecosystem features such as seagrass meadows, coral reefs and mangroves can provide effective coastal protection by dissipating wave energy and reducing extreme water levels at the coastline. This is leading to practical guidelines and proposals to preserve and restore these valuable ecosystems in place of conventional engineering structures to protect coastlines.

MANAGING BLEEDING-EDGE TECHNOLOGY TODAY FOR LEADING-EDGE RESEARCH TOMORROW

As Pawsey's Chief Technology Officer, Ugo Varetto not only manages Pawsey's technical operations and scientific computing support, he also leads Pawsey's overall technical strategy, putting him at the centre of Pawsey's capital refresh and the integration of new technologies like quantum computing.

Ugo brings a wealth of experience to the role, with a background in software engineering and a more than 25-year history of developing software, designing IT solutions and leading research and development (R&D) projects. Working in Silicon Valley start-ups was only the beginning. "I have always had an interest in code optimisation, real-time computer graphics, low-latency networks and machine learning," says Ugo. "Then all of the technologies I'd been working with for many years in commercial companies started to be applied to supercomputing. I got pulled into HPC applications in 2006 at the Swiss National Supercomputing Centre and have been contributing to R&D projects in HPC ever since."

Pawsey has a number of groups that support researchers through both short- and long-term engagements using the HPC facilities. Ugo's involvement with the senior scientists and university research leaders is pivotal to understanding the resources they need, allowing the various groups at Pawsey to best meet their requirements. Ugo notes that filling both an executive and a technical lead role can be challenging: "There is such diversity within the research community and the projects they undertake. I have an active role in planning and co-designing solutions with researchers in computing, visualisation and data management for their various research projects, and then also implement processes at Pawsey to make sure their resource needs can be met".

Ugo is well-placed to advise on the best software or hardware technology to apply to a given problem, as he puts significant time into assessing and testing the latest technology. "At the Swiss Supercomputing Centre I contributed to evaluating and designing the first ever HPC solutions based on graphics accelerators. They're now mainstream, and a significant part of our capital refresh. But that type of work is ongoing – I'm really proud that at Pawsey, we're deploying and testing new technologies together with supercomputing centres around the world. From exploring quantum computing, to bringing our users into the exascale through activities like our PaCER program, it's great to be able to offer the best-in-class support to our scientists."

"Working at Pawsey (and HPC centres in general) is special because through collaborations with both researchers and computing vendors we can enable cutting-edge science through bleeding-edge technology. It's really rewarding to make a tangible contribution to Australia's research outcomes, and actively contribute to the design of future national supercomputing infrastructure at the same time. Pawsey in particular is unique because of the focus on large-scale real-time data processing and management brought in by the various radio astronomy facilities we support – the Murchison Widefield Array (MWA) and ASKAP – and in the near future the SKA."

Radio astronomy is showing the potential for our next industrial revolution – Industry 4.0 – the digital transformation of production industries and industrial control systems. Its requirement for high-fidelity simulations and data analytics to manage the increasing size of data sets will require computational capabilities orders of magnitude above what is currently available in Australia, even with Pawsey's capital refresh. Ugo expects to stay in the middle of it: "Exabytes of data are going to need exaflops of computing power...."

DESIGNING THE SYSTEMS FOR SUPERIOR DATA STORAGE

Luca Cervigni is the IT infrastructure architect at Pawsey, and for the past five years has been working extensively with Pawsey's cloud and storage technologies. He describes his days as spent "writing notes, drawing diagrams and coding", but with his responsibilities for systems design, he also works closely with other supercomputing centres (such as the National Computational Infrastructure (NCI) in Canberra and the Swiss National Supercomputing Centre (CSCS)) and major research facilities (such as the European Organisation for Nuclear Research (CERN) and South Africa's SKA pathfinder telescope MeerKAT) to ensure that Pawsey's systems stay at the cutting edge. "Exchanging point of view and building on the experience of other experts worldwide to better understand the strengths and weaknesses of each system is very important in my work, to ensure Pawsey invests in appropriate future technologies, while minimising the risks of implementing new systems."

His scientific interests span physics and astronomy, and he is no stranger to mega-science projects – he was working at CERN in Geneva when the elementary particle the Higgs boson was first detected in 2012. "That was one of the most exciting moments of my career," Luca remembers. "It is why I always love to support and enable research: I can see that what I do is so important for so many others."

"Pawsey is arguably the most cutting-edge research-enabling centre in Australia. We may be known for big projects like the SKA, but we support an extremely broad range of science disciplines. It would be impossible for many of these researchers to perform their jobs and achieve what they do without Pawsey. It makes working here pretty special."

Luca has recently been designing, and is the technical leader for, Pawsey's new 90 PB object storage cluster. This is the largest part of Pawsey's $70 million Capital Refresh apart from the new supercomputer Setonix itself. The new storage system accommodates not just the sheer volume of storage required, but also considers the speed at which data can be initially accessed, and moved in and out of storage.

"This new object storage will completely revolutionise the way Pawsey users access their data," explains Luca. "Our object storage will use S3, a standard Hypertext Transfer Protocol (HTTP) protocol that will allow researchers to access their massive datasets more easily via the internet. It will move Pawsey even closer to its future goals of extremely large but very accessible data collections."

"It's been very rewarding to have the trust of Pawsey's management and be empowered to suggest new approaches and change the way we design our internal technologies as we upgrade," says Luca. "HPC of the future will only move further towards complete adoption of containerised workloads and object storage. It allows us to completely decouple where the data is stored (Pawsey) from where it is analysed and worked with (everywhere the researchers are)."

With Luca's assistance, Pawsey is already making massive moves in this direction.

BIOLOGY AND BIG DATA NEED BIGGER COMPUTERS

Sarah Beecroft joined Pawsey in 2020 to increase support for our bioinformatics research community. With a research background in the genetics of neuromuscular disease, she had already experienced firsthand how genetics research is becoming increasingly reliant on computational power. She regularly needed to apply coding skills to examine the complexities of biological phenomena, entering the realm of bioinformatics. In a happy professional development, Sarah discovered that she thrives on the analytical challenge of coding, and now extends that expertise to help other researchers who may be less comfortable with the newest HPC tools available in genetics research.

"Supercomputing has traditionally not been used in the life sciences, so getting biology users onto supercomputing resources is still a challenge," says Sarah. "But Pawsey is a science enabler. With my expertise in medical research, I understand both the scientific concepts that bioinformatics and biology researchers are working with, and how their research efforts can be fast-tracked by HPC."

As Pawsey's Bioinformatics Applications Specialist, Sarah provides training and development about using Pawsey resources and upskills researchers to apply them to biological questions. She also directly assists researchers with problems through the helpdesk, implements computational pipelines that are easy to run and well-documented, and generally builds the supportive infrastructure around the developing bioinformatics 'community of practice'.

"It is very gratifying getting biology researchers running on our systems. I've been 'on the other side', and I know first-hand how helpful it is when someone can give you the solution or the tool you so desperately need for your research."

"I feel very privileged to work here, it's a great blend of intellectual freedom, new projects, helping people achieve their research goals, coding-based work, problem solving, and training others. It's not a 'siloed' workplace at all, so there are lots of interactions across all areas at Pawsey, and I have ongoing collaborations with a significant sub-section of our users. And I'm constantly learning myself, which is great."

Sarah can see that HPC will become increasingly important both for researchers and industry in future. "Some genetic computational work that would take decades to run on a laptop can be done in a day using our resources – we provide computational capacity on a scale that is hard to access in any other way. The scale of the work that needs to be done will just continue to increase, as our understanding of complex biological systems continues to grow. Data analytics will be an important area for a long time to come."

Sarah expects the research community as a whole to become much more HPC-savvy in the next decade, because she's helping to make that happen.

KEEPING RESEARCHERS AT THE EVOLVING EDGE OF HPC

Pascal Elahi could always see that HPC would revolutionise scientific research by enabling experiments that wouldn't be possible in the physical world. "I've always been fascinated by the history of the cosmos, and how physics governs its evolution," he explains. "But that's difficult to explore experimentally. The idea of using HPC systems as a virtual laboratory for building test universes to explore the physics that govern our own Universe was very appealing, and what drew me to the field of computational astrophysics".

Pascal is now a Supercomputing Applications Specialist at Pawsey, using his science background and understanding of how researchers use HPC to help Pawsey users run their software on large supercomputing systems efficiently. If your model system spans the Universe, you've got to run it efficiently, no matter how big the computer is.

His fluency in programming languages like C++, CMAKE, Python, and GPU programming, and communication and training skills are being used to help others gain maximum benefit from their HPC use. Pascal particularly focuses on the radio astronomy and bioinformatics communities, and is now deeply embedded in their activities at Pawsey.

"Using HPC systems is not always a simple task, and tailoring software to make the most efficient use of Pawsey's new systems requires expertise in software development and computing, along with an understanding of the research being attempted," says Pascal. "The key goal of my work is to enable new science by allowing researchers to perform new, technically-challenging simulations and analysis – their critical computations – on our supercomputing infrastructure."

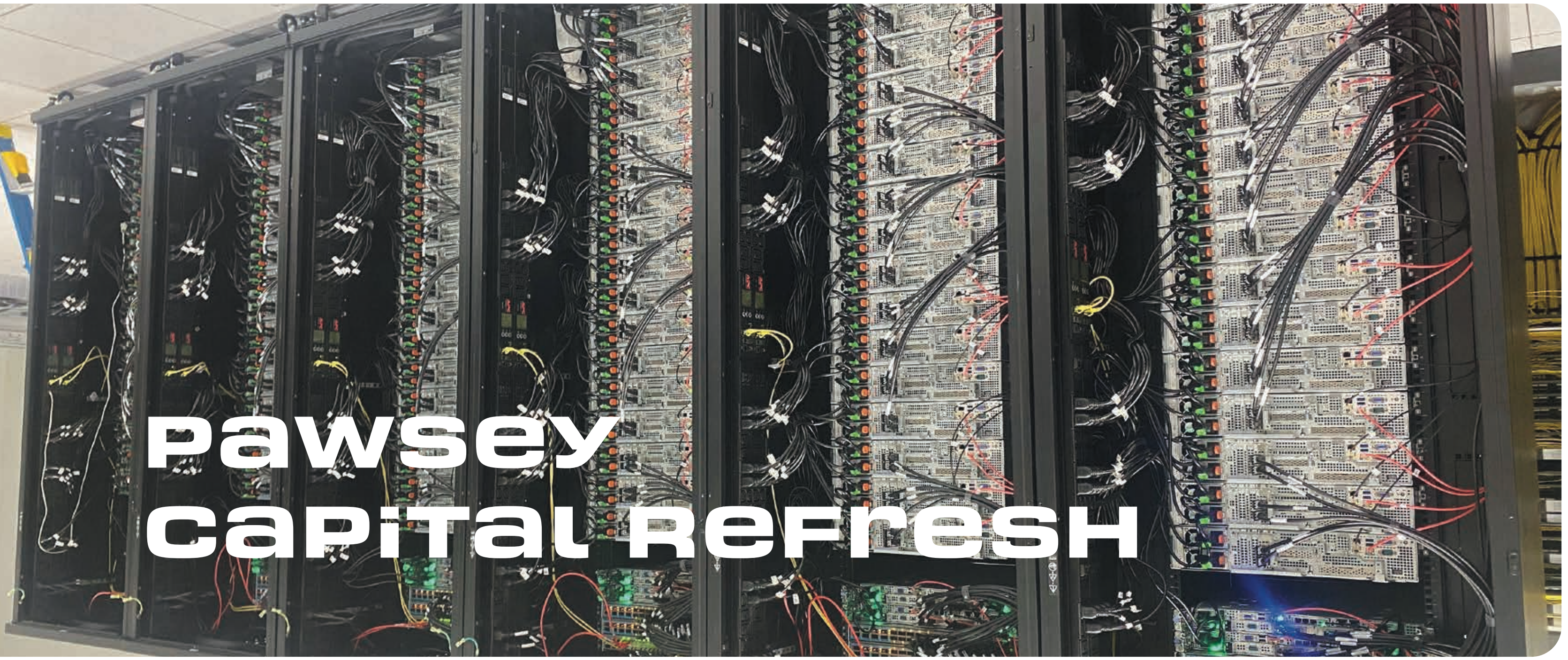
Pascal particularly appreciates the diversity of science it enables, from small research projects to mega-science projects like the SKA. As an example, Pascal has been helping the radio astronomy community using the MWA to migrate onto the new compute cluster Garrawarla, procured as part of the Capital Refresh, and improve their analysis pipeline to make the most of the latest-generation Central Processing Unit (CPU) and Graphics Processing Unit (GPU) processing now available. He also helped the CSIRO's ASKAP operations team commission their new compute cluster, designed to ingest and process data coming directly from the telescope. Extensive collaboration with the ASKAP team was required to get the new system running the complex software needed to process ASKAP data on schedule.

"HPC is moving towards more specialised accelerators making up a significant portion of the computing power provided, specifically GPUs as we're seeing here at Pawsey, and even eventually QPUs (Quantum Processing Units)," notes Pascal. "So we will always need software development and computing specialists to enable researchers to make the best use of the computational power available." And he's not complaining – the diversity of his working day, from meeting with researchers, training users, documenting computational processes and troubleshooting problems – is one of his favourite parts of working at Pawsey.

Pascal's future at Pawsey looks to be even more varied: "Our research users themselves are now joining us in pushing to transition to more 'green' computing, such as our new more energy-efficient GPU clusters. The carbon footprint of computing is significant, but by staying deeply embedded with the research communities we support, we're well positioned to keep Australian researchers at the forefront of energy-efficient computational ability as it continues to improve."

WORKING AT THE FOREFRONT OF TECHNOLOGY AND SCIENCE ON OUR PATH TOWARDS INNOVATION AND DISCOVERIES

A NEW ERA OF RESEARCH



Pawsey Capital Refresh

To ensure Pawsey continues to provide internationally-competitive supercomputing and data services, a $70 million capital refresh project funded by the Federal Government began in 2018 to secure the next generation of supercomputers, data and supporting infrastructure for Australia's research communities.

This upgrade is allowing Pawsey to keep pace with global advances in supercomputing technology, and provides significantly increased compute power and speed, expanded data storage capacity, and more efficient high-performance computing services for Pawsey users.

The upgrade is being undertaken as a staged process. In the previous financial year Pawsey's Nimbus cloud infrastructure was expanded, providing five times more memory and five times more storage with improved computational flexibility, accessibility and speed.

Astronomy's new high-speed storage for both the Murchison Widefield Array (MWA) and Australian Square Kilometre Array Pathfinder (ASKAP) was upgraded in the previous year, and procurement began for a new **MWA compute cluster.** The new 78-node cluster delivered by HPE was installed and available for MWA users from September 2020. The new cluster, 'Garrawarla' provides a dedicated system for astronomers to process in excess of 35 PB of MWA telescope data. It provides users with enhanced Graphics Processing Unit (GPU) capabilities to power Artificial Intelligence (AI), computational work, machine learning workflows and data analytics.

ASKAP's new ingest nodes were delivered by HPE and ASKAP commenced using the new ingest nodes in December 2020. The sixteen original ingest nodes were replaced with nodes containing the latest AMD processors designed for I/O. They have twice as much data bandwidth as the previous generation and more memory channels, ensuring they can keep up with the data produced by ASKAP. Along with three dedicated nodes for providing ancillary services, they have dedicated storage in the form of the ClusterStor E1000. Approximately half a petabyte of NVMe storage is dedicated to the ingest process, capable of speeds in excess of 150 GB/s.

The tender for the new **Pawsey Supercomputing System** was released in November 2019, with evaluation of tenders occurring in 2020 and the contract awarded to HPE in October 2020. Named 'Setonix' after the species name for the Quokka, native to Western Australia, it is being built using the HPE Cray EX architecture, with emphasis on accelerators with future-generation AMD EPYC™ Central Processing Units (CPUs) and AMD Instinct™ GPUs. It will deliver up to 50 petaFLOPs, or a 30-fold increase in raw compute

power when compared to its predecessor systems Magnus and Galaxy, and will be at least 10 times more energy efficient.

Setonix Phase 0 (test and development system) arrived at Pawsey in May 2021. Phase 1 hardware is due to arrive in September 2021, with Phase 2 (the full system) to be commissioned in Q2 2022.

Contracts were awarded in January 2021 for the **multi-tier Storage** procurements. Dell was awarded the contract for 'Acacia', the new online storage system that will provide over 60 PB of object storage for long-term archiving of researcher data. The system will be divided into two zones, one designed for data that needs to be accessed faster than the other, which is designed for energy-efficient long-term storage.

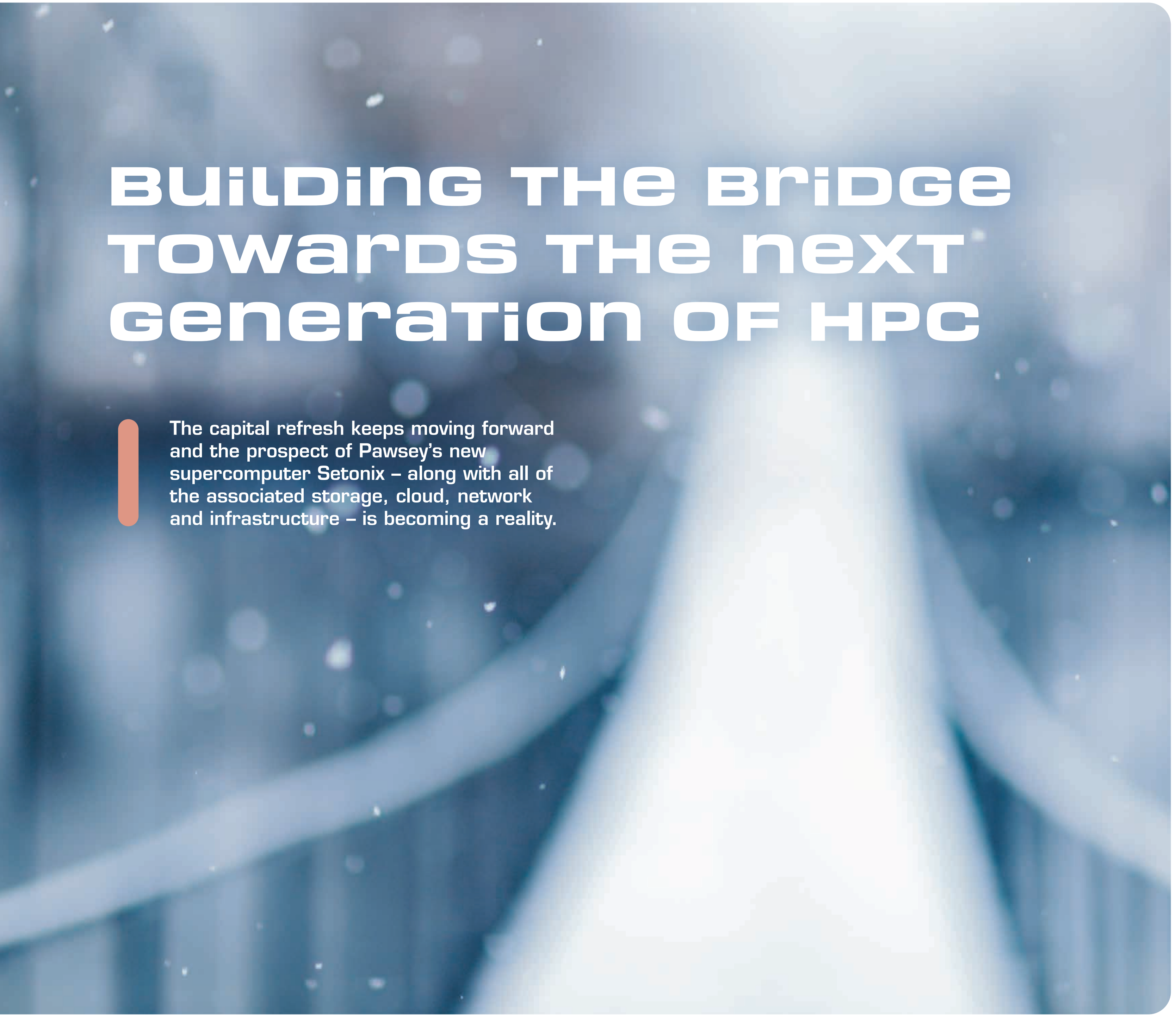
Xenon was awarded the contact for the new offline storage 'Banksia', and will replace the storage management software with an open system that will provide an expandable platform to build on Pawsey's investment in object storage. It will also re-use Pawsey's existing tape libraries and utilise a new 5 PB cache to take full advantage of the new 100 Gbe network infrastructure. With deployment of the new storage impacted by COVID-19 travel restrictions, the new storage is planned to go live in Q4 2021.

Building works were contracted to Shape Australia in April 2021, to perform the required building modifications to support the installation and running of Setonix, Acacia and Banksia.

The final stages of the upgrade were procured in February 2021. Palo Alto is providing the new **Firewall** needed to support a throughput of 100 Gbps into AARNet, and hence into other Research and Education networks and into the National Computational Infrastructure (NCI). The final configuration consists of a high availability pair of PA-7050 chassis, each with two 100 Gbps Network Processing Cards and a Data Processing Card to achieve the desired 100 Gbps throughput.

The **Network Upgrade** was procured from CISCO. Pawsey has moved away from a monolithic single core router to a spine-leaf architecture with a 400 Gbps backbone and 100 Gbps links to host endpoints. This allows the Nimbus cluster, which was upgraded to 100 Gbps internally, to connect at that speed to the Pawsey network. The network has been designed to be easily expandable to support Acacia's Ceph-based object storage platform as well as integrate with Setonix. This will allow all network endpoints (login nodes, visualisation servers, data mover nodes, etc.) to realise a ten-fold increase in bandwidth, moving from 10 Gbps to 100 Gbps ethernet.

With the capital refresh upgrades, Pawsey will allow Australia's researchers to remain globally competitive, upscale their research ambitions and accelerate scientific discovery.



Building the Bridge Towards the Next Generation of HPC

The capital refresh keeps moving forward and the prospect of Pawsey's new supercomputer Setonix – along with all of the associated storage, cloud, network and infrastructure – is becoming a reality.

Alongside the procurement, installation and testing activities, we have been working to ensure the journey onto this cutting-edge technology will be as smooth as possible for our researchers – from migrating code and data to best practices on the new architecture and scaling up Australian computational research to new heights.

During this year as we prepared for the arrival of Setonix, Acacia and Banksia, our primary focus has been on readying staff. Pawsey staff participated in over 4,000 hours of training sessions and workshops, building deeper skills in core technologies aligned with the incoming infrastructure and systems. These included sessions on:

- Directive-based GPU programming
- ARM Forge profiling and debugging
- Kubernetes
- CMake
- Python
- Spack
- CEPH
- Parallel I/O.

The Pawsey team collaborated with AMD to explore the ROCm software stack and HIP programming model, by testing various features and providing feedback to AMD's developers. This work will be crucial to the successful implementation of migration from current GPU-based systems to Setonix. The work of Pawsey's team was recognised in the official launch of AMD MI100 GPUs.

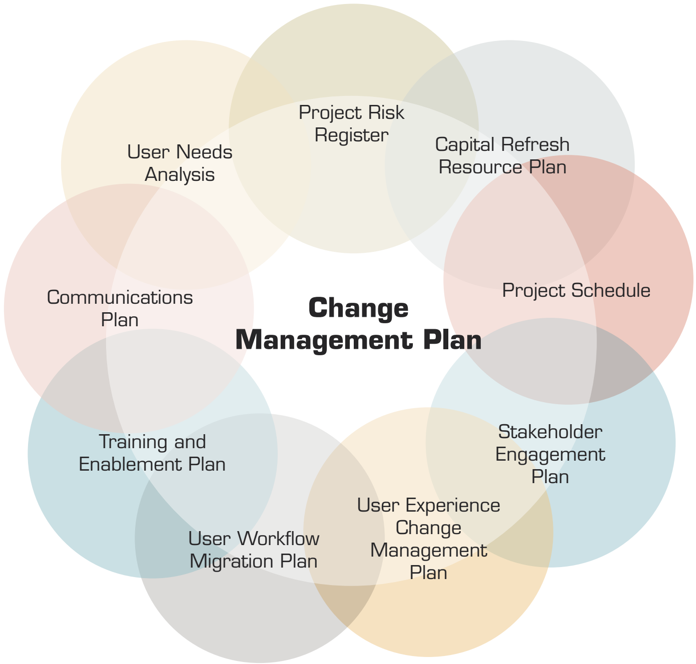
In preparation for the upcoming AMD GPU technology in Setonix Phase 2, Pawsey's supercomputing applications team developed and tested several mini-apps using a range of applicable programming options, such as HIP and OpenMP. These applications were initially developed on Topaz (Pawsey's existing GPU Cluster) using a CUDA backend for GPU acceleration, before being tested on both AMD MI50 and MI100 GPUs in collaboration with AMD. Team members also engaged in developer presentations and discussion sessions within the team to exchange knowledge and lessons learnt from working with the new technology.

The team also reviewed current policies and practices around configuration and management of the scientific software stack on our supercomputers. As a result, a new document was produced, defining the policies and protocols to be adopted on Setonix. In this context, the choice was also made to move from our current in-house tool for package installation, Maali, to a publicly available tool developed at Lawrence Livermore National Laboratory USA, Spack. This change will enable collaboration with other Centres, as well as reduce the amount of work required to maintain the scientific stack.

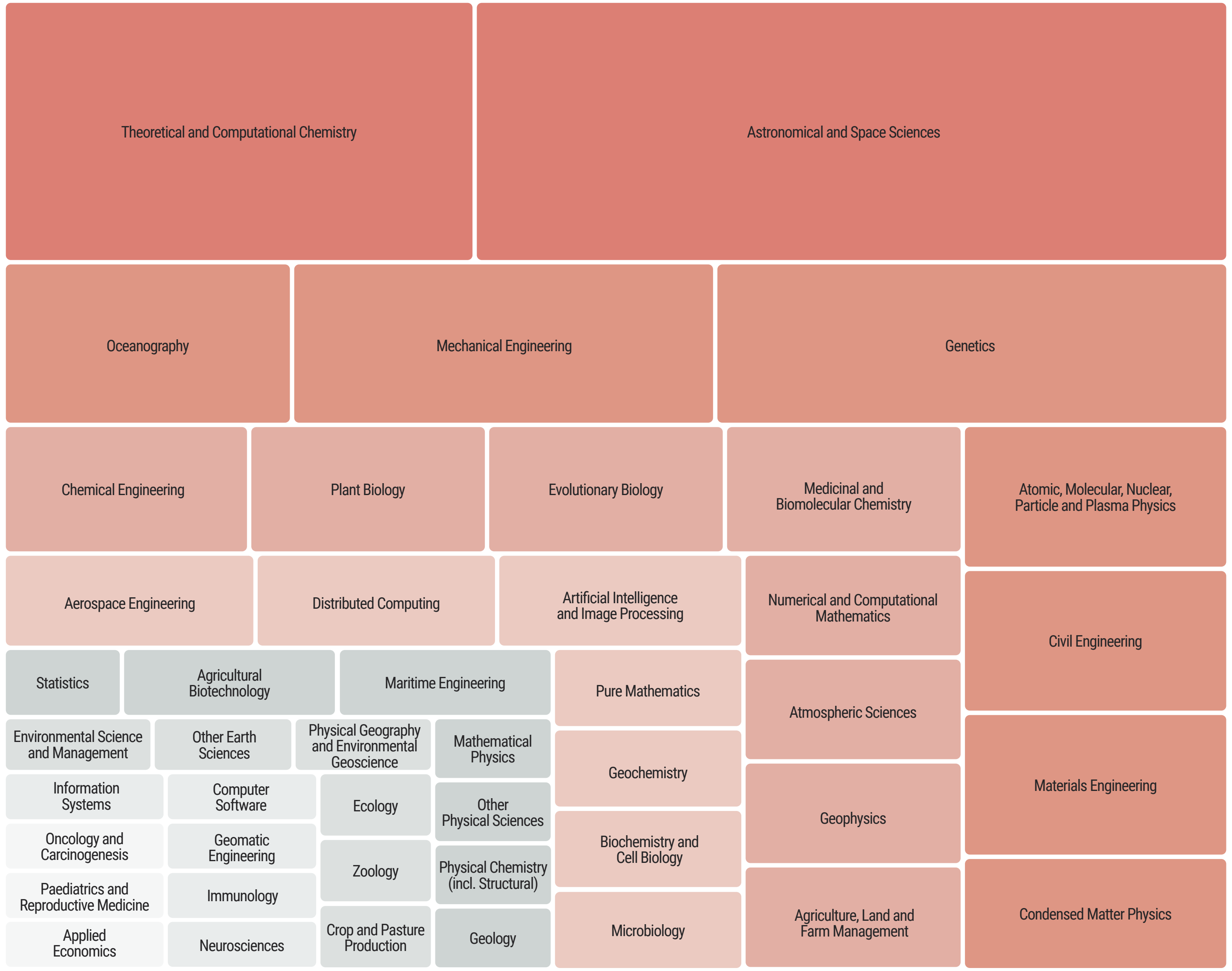
To enable continuous testing of all HPC systems, the team has used the ReFrame testing framework developed by the Swiss National Computing Centre CSCS and integrated it with the Jenkins automation server. With this setup, Jenkins will automatically launch ReFrame tests daily on all current production systems and failed results are written to the Confluence portal. It will also raise a Jira ticket for the Supercomputing Application team to investigate the failed tests. The applications team have identified major user HPC applications and libraries covering over 80 per cent of the user workflows and written scripts to incorporate them in the ReFrame test suite.

Pawsey staff also participated in change management, creating a plan on how Pawsey proposes to take its key stakeholders on a transition journey to enable successful adoption and usage of the new infrastructure. The goal of the overarching Change Management Plan is to capture the ethos and approach to successfully transition stakeholders from existing to new infrastructure – not in a technical sense but in user/human terms.

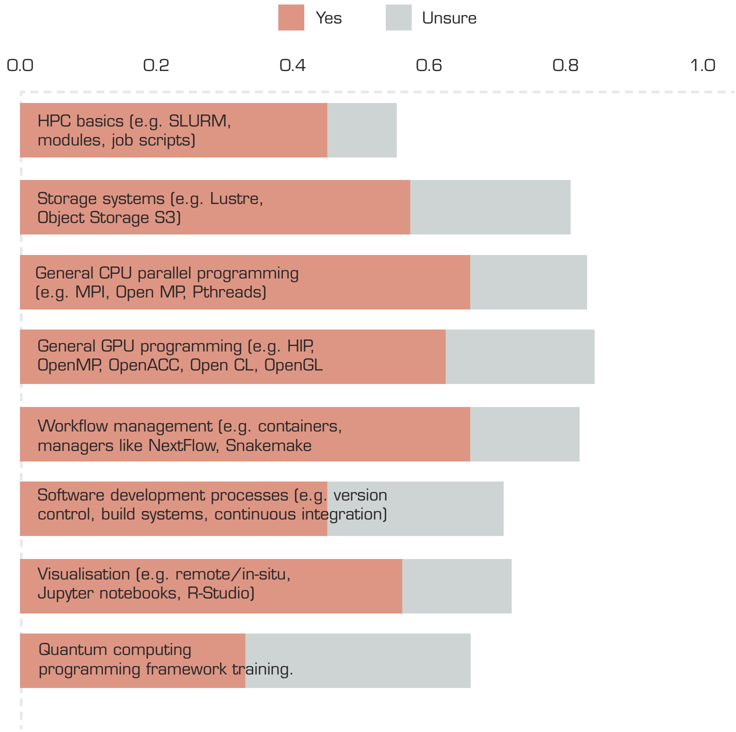
Migration success ultimately depends on individuals and teams of researchers not only transitioning to potentially new ways of working but also those ways becoming "business as usual". Integral to this Plan are practical sub-Plans, targeting key areas:

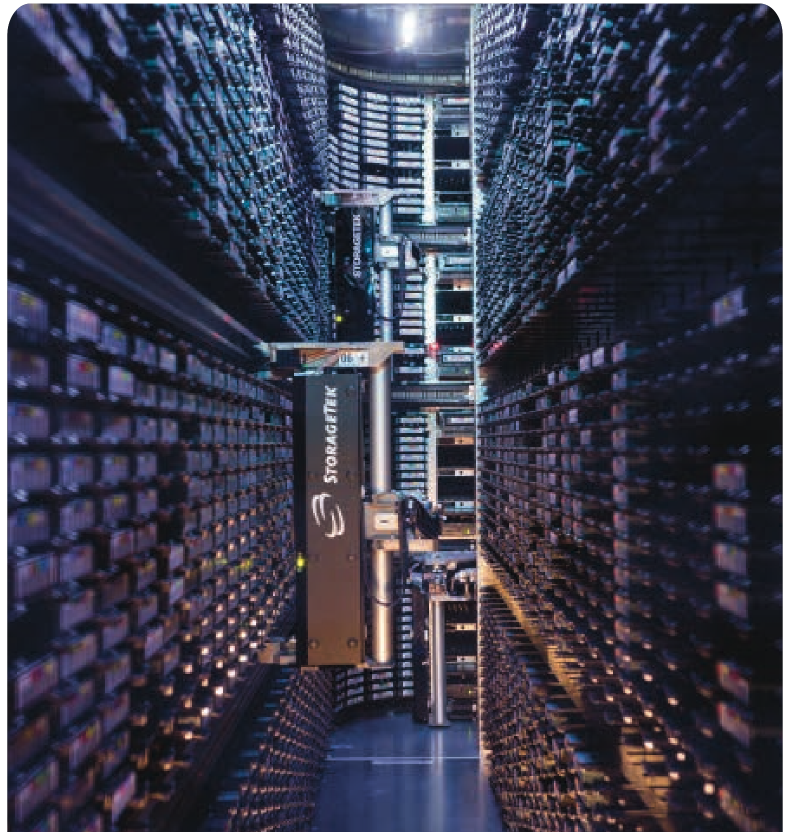


PROFILE OF THOSE WHO COMPLETED THE SURVEY BY FIELD OF RESEARCH



Supporting these Plans was an in-depth analysis of users based on an extensive user survey. One hundred and sixty-seven participants from a pool of 1,874 Pawsey users completed the 43-question survey, providing Pawsey with extremely valuable and detailed data. There was representation from all scientific research domains that Pawsey supports. Capturing data both across and within domains enables us to develop a multi-dimensional user profile, looking for patterns across multiple domains as well as within single domains. Work is now underway to create a series of user roadmaps tailoring training resources to users' needs to enable a smooth migration to Setonix.





They survey showed that most users are interested in workflows (like containers, nextflow), Storage, CPU and GPU, and large fraction in Visualisation services. More than half of the respondents were interested in advance training.

Systems and services

Allocation in numbers for the 2021 calendar year

Over 512 million

hours were requested across all allocation schemes.

276 million

core hours were allocated.

60

projects were granted as NCMAS allocation.

48

projects accessed Magnus via Partner allocation.

35

projects received an allocation via the Energy and Resources scheme.

ACCESS TO PAWSEY SUPERCOMPUTERS

Pawsey provides supercomputing access to researchers predominantly via competitive merit allocation schemes. Through these merit allocation schemes, the Centre strives to:

- maximise the overall research impact of Pawsey supercomputing resources
- promote scientific advantage in priority domains such as radio astronomy, energy and resources
- provide leading-edge supercomputing resources for researchers in Pawsey partner institutions
- enable wider adoption of, and benefit from supercomputing across Australia.

The Galaxy and Garrawarla supercomputers are only available for radio astronomy-focused operations. They are specifically used to support ASKAP and MWA operations, two of the Square Kilometre Array (SKA) precursor projects currently under way in the north-west of Western Australia. For ASKAP, Galaxy acts as a real-time computer and is a part of the telescope, allowing direct processing of data delivered to the Pawsey Centre from the Murchison Radio Observatory. Garrawarla is a dedicated system to process in excess of 35 PB of telescope data from the MWA, providing users with enhanced GPU capabilities to power AI, computational work, machine learning workflows and data analytics.

RESOURCE ALLOCATIONS

Over 512 million core hours were requested on the supercomputer Magnus during the period, almost twice as many hours as were available. 276 core hours were allocated.

The competitive merit allocation process in 2020 continued to underpin Pawsey's commitment to high-end science. As part of the 2021 allocation, released in January 2021, 143 projects received an allocation on Pawsey supercomputers.

Supercomputer Allocation Scheme

THE NATIONAL COMPUTATIONAL MERIT ALLOCATION SCHEME

Twenty-four per cent of the supercomputer time available at Pawsey is allocated through the National Computational Merit Allocation Scheme (NCMAS), along with compute-time allocations from other Australian High Performance Computing (HPC) centres including the NCI, the University of Queensland, and Monash University. The scheme operates annual allocation calls, which are open to the Australian research community and which provide substantial amounts of compute time for meritorious, computational-research projects.

THE PAWSEY PARTNER MERIT ALLOCATION SCHEME

Twenty-six per cent of the whole supercomputer time available at the Centre is allocated through this scheme. Calls are open to researchers in Pawsey Partner institutions – CSIRO, Curtin University, Edith Cowan University, Murdoch University and The University of Western Australia – and provide significant amounts of compute time for meritorious computational research projects. This scheme operates annual calls, with an out-of-session application process for newly eligible project leaders.

THE PAWSEY ENERGY AND RESOURCES MERIT ALLOCATION SCHEME

Thirteen per cent of Pawsey's total supercomputing resource is allocated through this scheme. Open to the Australian energy and resources research community, it provides significant amounts of compute time for meritorious research projects in these domains. Energy examples include generation, storage and distribution, while resources examples include exploration, minerals extraction and processing, groundwater, and waste management. This scheme operates annual calls, with an out-of-session application process for newly eligible project leaders.

THE RADIO ASTRONOMY SCHEME

The Galaxy and Garrawarla supercomputers (32 per cent of Pawsey's total HPC resources) are fully allocated to support ASKAP and MWA operations.

A National Research Facility

The Pawsey Supercomputing Centre supports researchers at universities across the nation. As part of the NCMAS, the two national Tier 1 centres, Pawsey and NCI, with Monash University and the University of Queensland contribute a number of core hours from their systems to service the national research community via meritorious allocations. Pawsey provides 100 million core hours to NCMAS for allocation to projects every year.

During the reporting period the top three institutions benefiting from those resources were Monash University, followed by University of Melbourne, and Curtin University.

National access is also provided via the Energy and Resources Merit Allocation Scheme and the Pawsey Director's Allocation Scheme.

NCMAS ALLOCATION BY INSTITUTION

INSTITUTION	ALLOCATION (CORE HOURS)
Monash University Total	22,840,000
University of Melbourne Total	19,500,000
Curtin University Total	14,750,000
University of New South Wales Total	11,430,000
University of Adelaide Total	7,700,000
RMIT University Total	7,400,000
University of Western Australia Total	5,600,000
University of Queensland Total	5,300,000
Queensland University of Technology Total	4,690,000
University of Sydney Total	2,400,000
Australian National University Total	1,050,000
CSIRO Total	1,050,000
Griffith University Total	810,000
University of Technology Sydney Total	480,000
Deakin University Total	400,000
Swinburne University Total	400,000
Murdoch University Total	320,000
Charles Darwin University Total	250,000

SUPPORTING AUSTRALIAN SCIENCE PRIORITIES

The Centre supports the National and State Science Priorities. Among the 143 projects that Pawsey supercomputers supported for the 2021 calendar year, Magnus uptake from engineering projects remained steady, accounting for 37.2 per cent of its use. Physical science projects increased with 31.0 per cent of the allocation, while chemical and Earth science allocations decreased slightly to 10.8 per cent and 11.7 per cent respectively.

Galaxy is Pawsey's real-time radio astronomy supercomputer, connected to the telescopes in the Murchison Radioastronomy Observatory, and its usage is 100 per cent consumed by radio astronomy projects. As reported as part of the Capital Refresh project, Garrawarla, a new GPU-based supercomputer was delivered to expand radio astronomy support for the MWA, to provide a better tuned system for that community while also allowing researchers using ASKAP telescopes to take full advantage of the Galaxy supercomputer, in preparation of ASKAP approaching its full surveys of the sky milestone.

SUPERCOMPUTER SPECIFICATIONS

GALAXY

- Peak performance in excess of 200 TeraFLOPS
- Real-time system for the SKA pathfinders and radio astronomy projects
- Cray XC30 system with over 9,000 Intel processor cores
- Over 30 TB of memory
- 64 NVIDIA Kepler K20X GPUs.

MAGNUS

- World-class supercomputer in excess of 1 PetaFLOP Linpack
- Cray XC40 featuring the 72 Gb/s Cray Aries interconnect
- Over 35,000 Intel 'Haswell' processor cores
- 3 PB of scratch file storage
- Over 95 TB of memory.

GARRAWARLA

- Peak performance in excess of 750 TeraFLOPS
- 78 compute nodes with a total of 3120 Intel 'Cascade Lake' cores, 78 NVIDIA Tesla V100 Tensor Core 32 GB GPUs and almost 30 TB of memory
- Fat Tree network topology with InfiniBand HDR100 100 Gb ConnectX-6
- Local storage: 240 GB SSD boot drive and 960 GB NVMe drive per node.

ZEUS

- Manufactured by HPE/SGI (USA)
- 92 nodes with 28 'Broadwell' Intel Xeon cores with a 100 Gb/s Omni-Path interconnect and 128 GB of memory per compute node
- 80 nodes with 64 Intel Xeon Phi 7210 cores with a 100 Gb/s Omni-Path interconnect and 96 GB of memory per compute node
- 11 nodes with four NVIDIA Tesla P100 SXM2 GPUs with a 100 Gb/s InfiniBand interconnect and 256 GB of memory per compute node
- Six nodes with 16 'Broadwell' Intel Xeon cores and 1 TB of memory per compute node
- SUSE Linux (SLES) 12 operating system.

TOPAZ (ZEUS EXPANSION)

- Peak performance in excess of 330 TeraFLOPS
- 22 compute nodes with a total of 320 Intel 'Cascade Lake' processor cores, 40 NVIDIA Tesla V100 16 GB GPUs and 4.6 TB of memory
- 20 visualisation nodes each with NVIDIA QUADRO RTX5000 16 GB GPUs
- Interconnect based on Mellanox ConnectX-6 HDR100 (100Gb/s) InfiniBand.

The Topaz GPU cluster is available for GPU migration purposes. Researchers with allocations on Pawsey HPC systems can experiment with porting their workflows to GPUs and benchmark performance of their simulations against preinstalled applications and libraries. Topaz also allows researchers to work with machine learning and AI workflows as well as prepare their codes for migration to AMD GPU-based Setonix resources.

OTHER SYSTEM SPECIFICATIONS

NIMBUS CLOUD

- 3,700 cores and 58 TB of RAM across 58 compute nodes
- 12 Nvidia V100 GPU nodes
- 9 PB of volume storage
- 100 Gb Ethernet networking.

Nimbus Cloud allows researchers to process and analyse larger volumes of data through additional storage and opportunities to work with Kubernetes container orchestration, building on Pawsey's existing container technology for its HPC systems. Researchers with memory-heavy workloads will find the upgraded infrastructure better suited to their research needs, with the capacity to support more computationally challenging projects.

Nimbus, Pawsey's cloud service, has seen significantly increased uptake during the reporting period with 113 different projects benefiting from its upgrades.

The cloud is often used as a stepping stone for researchers who find their computing capacity is not enough but have not identified a need for supercomputers or don't know how to use them. In some cases, Nimbus has been used by researchers as a resource additional to Magnus. Nimbus uptake from bioinformaticians has increased as the GPU technology included as part of the system has been demonstrated as extremely valuable to running their workflows as well as the container orchestration available.

VISUALISATION

- 20 remote visualisation Topaz nodes, dual Intel Xeon 2.5GHz 8 core processors.
- 9 high-end Windows nodes available via remote visualisation service for early adopters.

Pawsey's data visualisation services help researchers gather real-time insights from their data and allow the data to be more easily interpreted by translating it into more meaningful or recognisable forms. Seventy-eight projects benefited from Pawsey's visualisation services during the reporting period.

DATA STORAGE INFRASTRUCTURE

- Up to 100 PB of storage
- Two duplicate libraries for added resilience
- Connected at up to 40 GB/s.

As part of the data services provided at the Centre, Pawsey hosts a series of significant data research collections from national universities and government departments across Australia. In addition to housing in excess of 50 Petabytes of data storage resources, Pawsey's research collection service provides data sharing facilities to support scientific collaboration among researchers nationally and internationally.

During the financial year, 71 projects had access to Pawsey long-term storage. Although many projects are radio astronomy related, Pawsey also supports eight universities and CSIRO; Landgate and the Government Departments of Mines and Petroleum, Parks and Wildlife, Health, and Transport; and the Telethon Kids Institute.

NETWORK

50 devices supporting in excess of 6,000 ports

- ~1,400 non-blocking 10 Gbps ports
- ~70 non-blocking 100 Gbps ports
- The remainder are 1 Gbps ports, mainly used for management and monitoring purposes, and staff access.

Pawsey's network provides high-speed connectivity both within Pawsey itself, and with Pawsey's research partners in Western Australia through peerings with CSIRO, the University of Western Australia and Curtin University. Pawsey also peers with AARNet, providing access to the NCI and other national research institutions, and international research and education networks. Pawsey has multiple high-speed connections to the Murchison Radioastronomy Observatory, allowing the ingest, processing and storage of huge volumes of radio astronomy data. It also has a direct 100 Gbps link to the Curtin Institute of Radio Astronomy (CIRA).

Building the foundations of a greater leadership role in HPC and research for Australia

A World of Difference

A WORLD OF DIFFERENCE

At a time when science is demonstrating its importance daily, the infrastructure that supports and accelerates research should be widely accessible. Pawsey is dedicated to generating awareness of the supercomputing, large-scale data and scientific visualisation capabilities available within Western Australia and highlighting the possibilities available for researchers and industry to harness Pawsey resources and expertise.

Prime Minister Scott Morrison during his visit to the Pawsey Centre.

FIGHTING VIRUSES

The Pawsey Supercomputing Research Centre and the National Computational Infrastructure (NCI) Australia have joined the US-led COVID-19 High-Performance Computing (HPC) Consortium, with the announcement made during the Joint Commission Meeting on Science and Frontier Technologies Dialogue between the US and Australia to further strengthen our research relationship. The Honourable Karen Andrews, Minister for Innovation, Science and Technology, led the Australian delegation.

HARNESSING QUANTUM COMPUTING

Following our announcement to develop Australia's first quantum supercomputing hub for innovation in partnership with Quantum Brilliance, our 'Quantum Pioneers' were given access to a quantum emulator at Pawsey. This is in preparation for the world's first market-ready diamond-based quantum accelerator coming to Pawsey in 2022.

Pawsey hosted an online quantum computing event in August 2020, focusing on quantum computing for industry growth. Over 130 attendees heard from Australian leaders in the field to understand how quantum computing can benefit their research, business and help grow their industries. The expert line-up included Dr Cathy Foley, CSIRO's Chief Scientist; Mr Shaun Wilson, co-founder of quantum and Artificial Intelligence (AI) companies QxBranch and Shoal Group; Dr Andrew Horsley, co-founder and CEO of Quantum Brilliance; and Mr Bill Bartee, Partner at Main Sequence Ventures. The event is available on Pawsey's YouTube channel.

Two researchers and two SMEs granted early access to the system were then showcased in March 2021, allowing 122 community and industry representatives to find out more about their projects.

Head of the Australian Space Agency, Mr Enrico Palermo, with ASDAF and Pawsey representatives

LOOKING TO EARTH FROM SPACE

The Australian Space Data Analysis Facility (ASDAF) was established at Pawsey to support the growth of the Australian space industry, and unlock the value of space observational data for industries such as agriculture, mining, emergency services and maritime surveillance. Several videos were made showcasing case studies of potential opportunities available for SMEs and explaining how ASDAF will improve access to space data products and enable the exploratory use of space data to develop new products and services.

In May 2021 the ASDAF Expression of Interest call was announced with the support of Enrico Palermo, the Head of the Australian Space Agency and the Hon Christian Porter MP, Minister for Industry, Innovation and Science.

LOOKING TO SPACE FROM EARTH

Pawsey supported the NASA International Space App Challenge in Perth in October 2020. This global virtual hackathon takes a collaborative approach to problem-solving where teams aim to produce solutions to NASA-defined challenges we currently face here on Earth and in space. For its first virtual edition, Pawsey Executive Director Mark Stickells was a judge and Data Specialist Luke Edwards was a mentor.

Support for the Square Kilometre Array (SKA) project is also continuing, and in April 2021 Prime Minister Scott Morrison and Minister for Industry, Innovation and Science Christian Porter visited Pawsey and announced Australia's ongoing commitment with an $387 million investment in the SKA and the Australian SKA Regional Centre. Following this visit, other high-profile representatives from the Federal Government toured the Centre, including the Deputy Leader of the Opposition the Hon. Richard Marles MP and Mrs Madeline King MP and the Secretary of the Department of the Prime Minister and Cabinet, Mr Phil Gaetiens.

DIVERSITY AND INCLUSION

As part of the Centre focus on diversity and inclusion, Pawsey together with New Zealand eScience Infrastructure (NeSI), NCI Australia, Monash University, and Australasian eResearch Organisations (AeRO) launched the Women in HPC Australasian chapter during the eResearch New Zealand conference. The chapter was announced during a Birds of Feather (BoF) session *Challenge Accepted: Responding to community feedback for supporting diversity in HPC and eResearch.* The BoF saw 19 diversity and inclusion allies gather to discuss the next steps for the Chapter as priority activities in 2021.

Quantum Brilliance team member building the Diamond Quantum Computer

INNOVATIVE TECHNOLOGIES AT PAWSEY: JOURNEY TO EXASCALE AND BEYOND

Pawsey has historically had a focus on investigating and testing new information technologies as they become available, searching for opportunities to accelerate the computational power available to the wider research community. This focus solidified when the Centre revised its structure in 2020 to include a permanent team, variably staffed by Pawsey members with appropriate expertise suited to the technology opportunities as they arise, to continue exploring developing technologies and their potential relevance.

As part of the innovative technologies focus, Pawsey first gave our broader research community access to Graphics Processing Units (GPUs) in 2017, with the 'Advanced Technology Cluster' facilitating the evaluation of these then 'new' technologies. Eventually renamed Athena, the cluster provided access to two technologies: Intel Xeon Phi many-core processors, and NVIDIA Pascal GPU accelerators. It was the beginning of Pawsey's GPU journey, supported by activities to explore and expand their uptake, such as training and running GPU hackathons for our researchers.

New releases of the same technology have since been incorporated via our upgraded cloud Nimbus, Topaz, and most recently Garrawarla, the new GPU cluster supporting one of the SKA precursor telescopes, the Murchison Widefield Array (MWA).

The new supercomputer Setonix will give researchers access to a state-of-the-art HPE Cray EX supercomputer using over 200,000 next-generation AMD Central Processing Units (CPUs) and over 750 next-generation AMD GPUs.

The arrival of Pawsey's next-generation supercomputing system is being supported through Pawsey's involvement in collaborative work on programming models. Pawsey's adoption of the open-source AMD ROCm ecosystem and HPE Cray programming environment will also pave the way to exascale readiness for the Australian research community.

Object storage is now being deployed as part of the capital refresh, as an opportunity to improve data availability, data transfer speed, and overall storage capacity for researchers. The object storage 'Acacia', named after Australia's national floral emblem the Golden Wattle Acacia pycnantha, will provide 60 PB of high-speed object storage for hosting research data online. To maximise value, Pawsey will deploy community-edition Ceph software for building storage systems out of generic hardware, and has built the online storage infrastructure around Ceph in-house. As the service is expanded, the online object storage will become more stable, resilient, and even faster.

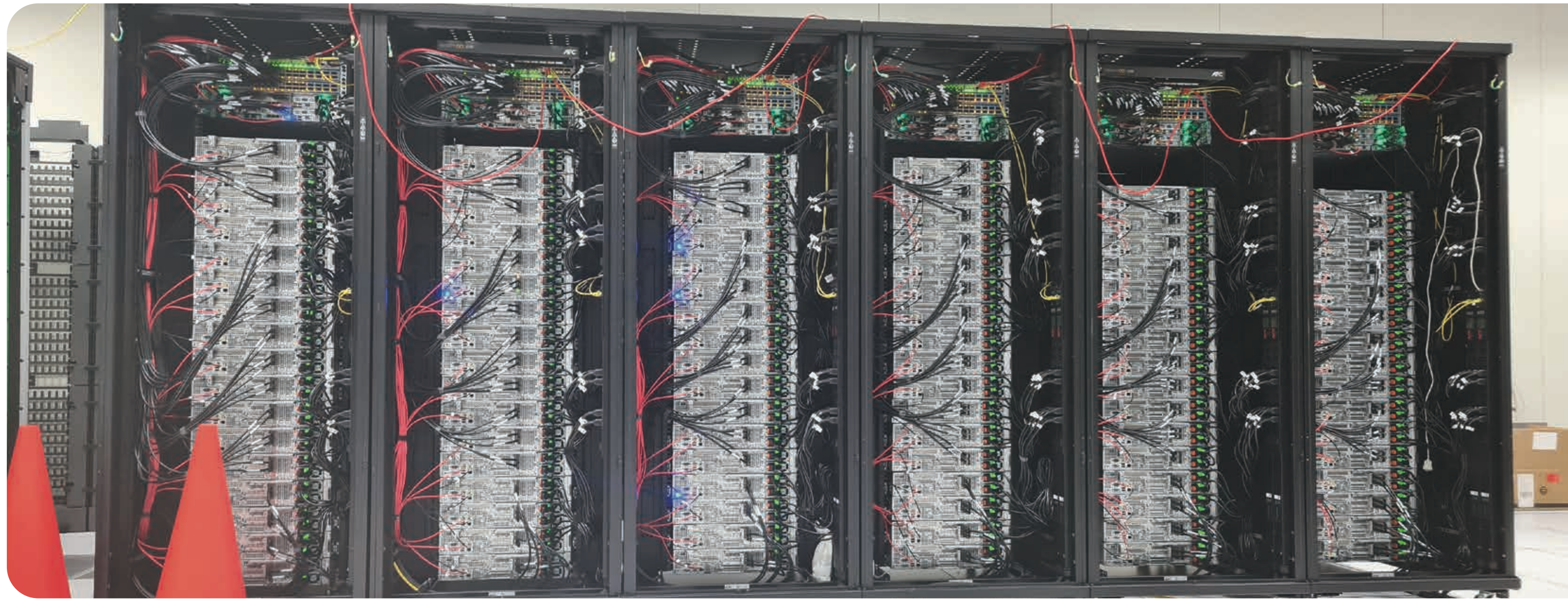
In June 2021 Pawsey announced that it will become the world's first quantum–supercomputing hub, hosting a diamond-based quantum accelerator as part of a new technologies testbed strategy to provide researchers with opportunities to test quantum computers and potentially help bring these new technologies to market.

Quantum Brilliance uses synthetic diamonds to build quantum accelerators that do not require near absolute zero temperatures or complex laser systems to operate like existing mainframe quantum computers. It is one of very few companies worldwide able to deliver quantum computing systems for customers to operate on-site.

To support hosting a quantum computer on-site, Pawsey and Quantum Brilliance have joined forces with other Australian industry leaders and researchers to establish Pawsey's Quantum Pioneer Program. As part of the program, four groups have been granted early access to a quantum emulator installed in Pawsey's systems by Quantum Brilliance to develop cutting-edge quantum applications in machine learning, logistics, defence, aerospace, quantum finance and quantum research.

Interest in quantum activities at Pawsey has been strong from not only the research community, but also industry and government representatives. Several information events have been hosted as part of the implementation, including one focusing on industry opportunities around this technology.

The Pawsey–Quantum Brilliance collaboration may yet bring us one step closer to small, affordable, 'plug-and-play' quantum computing.



The backplane of Acacia, Pawsey's new object storage system for research, exposes the cabling and high speed networking essential to building this new, state of the art infrastructure


National Youth Science Forum during their first visit to the Centre

EDUCATION AND TRAINING

Pawsey continues to offer advanced computing training, making supercomputing and high performance computing more accessible to researchers across Australia. This year's training has targeted the building of skills for the new systems and services coming through the Capital Refresh.

RESEARCHER DEVELOPMENT

Pawsey training programs are raising awareness and building core skills for Setonix and Pawsey's other refreshed resources, such as Acacia and Banksia.

All training activities are now offered online. Pawsey's previous in-person, two-day roadshow has been refreshed and is now offered as a nine-part modular webinar series. That series provides foundational knowledge on Pawsey's cloud, super compute, visualisation and storage resources. These and other webinars are available as recordings on Pawsey's YouTube channel.

Pawsey has also supported and/or provided training targeted to specific technologies and/or domains, including:

- GPU computing with Python
- Number crunching with Python
- Using containers in radio astronomy workshop series
- Using containers in bioinformatics webinar/workshop series
- MWA migration awareness raising.

Significantly more learners are reached through 'live' online events and subsequent views compared to previous in-person events.

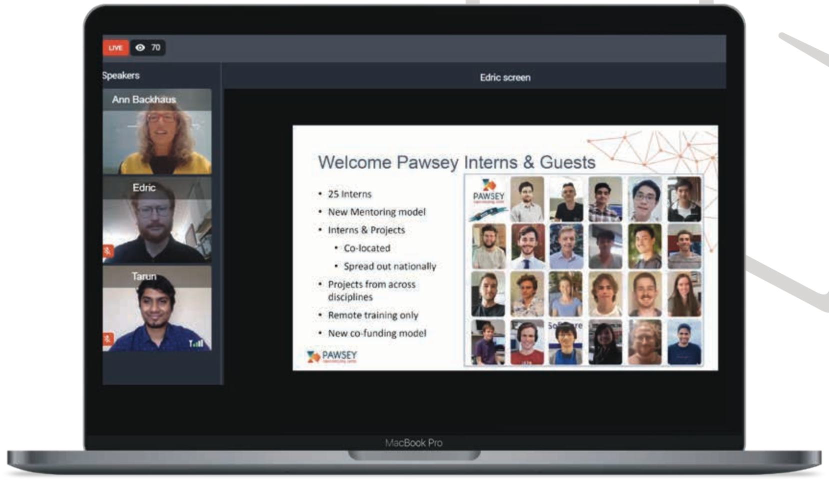
In this year, Pawsey-supported and provided training that reached over **5,500 users**, totalling some **5,300 hours** (live and recording hours).

PAWSEY INTERNSHIPS

The annual Pawsey Summer Internship Program reached new heights and 'changed it up' this year.

- We received a record-breaking 300 national applicants to the Program.
- Our Intern intake doubled, to 25 students.
- We trialled a new Intern Mentorship model, in which two previous Pawsey Interns returned to mentor and coach the new cohort.
- The Program was run entirely online – from an initial immersive week of training through to the final poster and presentation Pawsey Friday event.
- We expanded our training outreach, extending invitations to Perth-based International Centre for Radio Astronomy Research (ICRAR) interns, Australian Mathematical Sciences Institute (AMSI) award recipients, interns from Data61 and beyond.
- We trialled a new interactive, networking platform for our final event, which attracted more than double the attendance at previous, in-person events.

Intern videos about their projects are available for viewing on Pawsey YouTube.



PAWSEY TEAM DEVELOPMENT

Like Pawsey users, Pawsey staff focused on learning more about the technologies behind the infrastructure being secured through the Capital Refresh, including the new Pawsey Supercomputing System, Setonix, as well as object storage Acacia.

During the year, staff spent more than 4,500 hours in training and learning, equalling an average of 90 hours of training per person. Technical training was far and away the focus, including:

- ARM
- OpenMP
- CMAKE
- CEPH
- Parallel I/O
- Parallelware tools
- Spack
- Cray ClusterStor E1000
- Advanced C++
- Fortran modernisation
- Python.

Staff also maintained their focus on soft/essential skills, for example:

- Advanced Inclusive Leadership
- 'Science in Public' online media and communication masterclass
- Emotional Intelligence
- Change Management
- Inclusive teaching (Carpentries Instructor certification training)
- Lunch 'n Learns for staff.

LEADERS IN TRAINING AND EDUCATION OUTREACH

DReSA. Partnering with the Australian Research Data Commons (ARDC) and a national Working Group, Pawsey co-invested in the development of Digital Research Skills Australasia (DReSA). This platform provides a central location for learners and trainers to browse, discover, and organise training events and training resources, collected from Australasian training providers. This one-stop shop is scheduled for release later this year.

Learn@Pawsey. A new education outreach portal for learning debuted this year, coming out of a previous Pawsey Internship Project. The Sleep Survey features as the first activity on the site, and encourages students, classes, and individuals to try their hand at data collection and analysis. Using the site, data can be generated, collected, and visualised. Teacher resources in Maths, Science, and Digital Technologies are aligned with Federal/State curricula, making it easy to take up and use.

Conferences. Pawsey participated in numerous virtual conferences this year, including:

- The largest international supercomputing conference, SC20:
  - Presentation with an associated paper published in the Journal of Computational Science Education: Volume 12 Issue 2: Pawsey Training Goes Remote: Experiences and Best Practices.
  - Tutorial on Containers with leading global partners
  - Presentation and participation in a jointly conducted Teaching/Education BoF
- IEEE 2020 High Performance Extreme Computing virtual conference, contributing to the education presentation and panel discussion
- PEARC 20 Program education workshop facilitation
- ISC21:
  - Tutorial on Containers
  - Presentations on Scalable and Effective Online HPC Education and Training
- eResearch Australasia, featuring Pawsey presentations on moving online, and a Totally Nailed It BoF
- ARDC Skills Summit, co-leading sessions
- SCAsia21:
  - Pawsey Executive Director Mark Stickells facilitated a HPC Centre Leaders' Forum
  - Presentations included Australia's First Quantum–Supercomputing Hub and Next Generation Supercomputing at the Pawsey Supercomputing Research Centre
  - Participating in a 'live' virtual event for teachers and students, The Many Faces of HPC.

STEM AND EDUCATION OUTREACH

Education outreach activities and events were predominantly virtual, reaching 535 students and teachers with some 1,600 learning hours. These activities included:

- an eResearch workshop and a HPC course lecture at UWA
- a Science Café networking event for STEM students
- a HPC lecture series at Curtin, as well as a school-based How to teach online session
- ICRAR/UWA/Pawsey Work Experience Days held quarterly
- Pawsey's first live-streamed tour coincided with the launch of a Pawsey impact video for National Science Week in August 2020, reaching more than 900 people in total
- an incursion to Pawsey from the National Youth Science Forum
- ResBaz Perth 2021 sessions: Getting Started with Nimbus, Scaling your Research, Teaching Online
- BHP Foundation Science and Engineering Award activities for current national awardees and alumni, following the popularity of Pawsey's Virtual Tours
- Australia NextGen Program, providing a session on big data as well as judging final student projects
- International Women's Day activities, including a Women Using Technology: Data@Pawsey session.

SCAsia 2021 virtual HPC Centre Leaders' Forum

PACER

The Pawsey Centre for Extreme scale Readiness (PaCER) program was established in late 2020 to ready the research community to achieve extreme performance on Pawsey's next-generation supercomputer Setonix. This long-term, $1.42 million partnership between researchers and Pawsey's supercomputing specialists provides an opportunity for researchers to become exascale-ready, working on new algorithm designs and optimising codes, workflows and data movement pipelines.

Prof Igor Bray, Curtin University, one of the PaCER project researchers

Ten successful research projects were announced in March 2021. These research teams are gaining early access to supercomputing tools and infrastructure, training and exclusive hackathons focused on HPC performance at scale. Each three-year project is supported by a Pawsey co-funded doctoral or postdoctoral position and Pawsey expertise working closely with the research team.

A major focus of PaCER is to develop software for exascale computing. All of the projects are focused on developing computer codes for their research problems, with Pawsey working closely with them on software development and optimisation.

Each project also has a 'grand challenge' – a scalable problem tackling a major scientific question that is not possible to address using the current Magnus and Galaxy supercomputers. The focus is to determine the best ways to scale the research for exascale computing.

In 2022 these projects will be the first to exploit the scalability offered by Pawsey's new 50 petaFLOPS supercomputer Setonix to solve these 'grand challenge' problems, demonstrating pathways to what is currently unachievable in Australian research.

The PaCER program to date involves over 60 researchers, provides co-funding for 15 doctoral and postdoctoral research positions, and involves ten supercomputing application specialists at Pawsey.

The successful projects represent collaborations between ten Australian universities, CSIRO, ICRAR, and internationally with the SKA, US National Laboratories (Argonne National Labs, Ames National Labs, and the Oak Ridge National Labs), University of Toronto, and industry partner General Electric through GE Global Research.

The ten PaCER projects also currently use or have recently had allocations on supercomputing facilities worldwide including in the US (Summit at Oak Ridge National Laboratory, Argonne National Laboratory, XSEDE, Texas Advanced Computing Centre (TACC)), Europe (Partnership for Advanced Computing in Europe (PRACE), ARCHER, CINECA, Jülich Supercomputing Centre, HLRN, DiRAC), China (Tianhe-2), and Australia (NCI, Bracewell at CSIRO). The work done with Pawsey through their PaCER project is expected to translate to further research efficiencies in their work with these supercomputing centres worldwide.

PROJECT	PRINCIPAL INVESTIGATOR	SCIENCE DOMAIN	PARTNERS	IMPACT AND OUTCOMES
EXESS: The EXtreme-scale Electronic Structure System, predicting the chemistry of nanomaterial interfaces	Dr Giuseppe Barca	Quantum chemistry	ANU, Argonne National Laboratories, Ames National Laboratories, Monash University, Flinders University, Deakin University	Provide a framework to develop next-generation molecular modelling capabilities that support a broad spectrum of chemistry, biology, and material science research on computing systems ranging from petascale to exascale and beyond. Leveraging software development already underway in the US GAMESS Exascale Computing Project.
VISCOUS: Towards a molecular level understanding of flow-induced physical and chemical reactions	Professor Debra Bernhardt	Statistical mechanics, rheology	University of Queensland	Develop codes for nonequilibrium molecular dynamics atop the extensively-used LAMMPS molecular dynamics software for performance at exascale. Simulate flow-induced chemical and physical reactions in liquids important to a wide range of industrial and biomedical applications including advanced materials manufacturing, nanoscale filtration and drug delivery technology.
MCCC: Calculation of collisions with molecular targets using the convergent close-coupling method	Professor Igor Bray	Atomic and molecular physics	Curtin University	Produce high quality and comprehensive data describing the collision of electrons and positrons with molecules, which are needed in a range of applications. Exploit the capabilities of next-generation supercomputers, with a particular emphasis on accelerating calculations using GPU accelerators. Developments will allow the first-ever large-scale collision calculations to be performed for molecules more complex than H_2. Resulting data sets will drastically improve plasma models in a wide range of scientific and industrial applications.
PIGI: Parallel interferometric GPU imaging	Professor Melanie Johnston-Hollitt	Radio astronomy	Curtin University, University of Toronto	Develop new algorithms required to achieve science goals ranging from detecting the first stars and probing cosmic magnetic fields, to mapping the large-scale structures of our Universe. Combine the distributed nature of interferometric reconstruction algorithms with fast instrumental modelling using GPU accelerators, to accurately reconstruct images from extremely large data sets for future instruments such as the SKA.
EmPRiSM: Emergent phenomena revealed in subatomic matter	Dr Waseem Kamleh	Nuclear physics	University of Adelaide	Develop novel algorithms for Lattice QCD calculations at extreme scale to unlock previously unachievable calculations of quantum fluctuations in the space–time vacuum. Explore how the foundation of matter evolves to regimes relevant to the interior of neutron stars and the early universe.
MaPMoPS: Massively parallel models of particle suspensions	Dr Christopher Leonardi	Computational fluid dynamics, geoscience, petroleum	University of Queensland	Develop computational models of complex particle suspensions at previously intractable scales to investigate novel reservoir simulation techniques. Significantly improve the management of subsurface resources. Rearchitect in-house open source CFD code to leverage the next generation supercomputers.
HiVIS: Delivery of a next-generation data storage approach to unlock deep SKA and pathfinder observations	Dr Martin Meyer	Radio astronomy, cosmology	ICRAR, University of Western Australia, CSIRO, SKA, Oak Ridge National Laboratories, Australian SKA Regional Centre	Address one of the most significant grand challenge problems for the SKA – how to optimally image multi-day deep data sets. Aim to reduce the visibility storage requirements for these projects by an order of magnitude by developing a sparse data storage and processing pipeline based on UV-grids. Enable ground-breaking new studies of the role atomic hydrogen has played in the ongoing evolution of galaxies and its connection to their dark matter halos.
GTx: Towards exascale simulations for efficient, low-emissions gas turbines	Professor Richard Sandberg	Computational fluid dynamics, turbulence, engineering	University of Melbourne, University of New South Wales, General Electric Global Research	Upscaling in-house codes for high-fidelity CFD simulations of components of gas turbines at engine-relevant conditions on next-generation supercomputers, and beyond. Investigate phenomena contributing to energy efficiency, pollutant emissions, and operability to improve predictive tools used in industry.
EXA-GAMBIT: Searching for new particles from the attoscale to the exascale with GAMBIT	Dr Pat Scott	Particle physics	University of Queensland, Monash University, University of Adelaide	Model new particles towards explaining the identity of dark matter and dark energy, why neutrinos have mass, why the Higgs boson is as light as it is, and why we are surrounded by so much more matter than antimatter. Make it possible to combine the results of all relevant experiments and bring them to bear on all of the leading theories for new particles, by using exascale computing hardware to simulate billions of possible experimental signatures simultaneously.
BLINK and you'll miss it: blazingly-fast all-sky radio astronomy pipelines	Dr Marcin Sokolowski	Radio astronomy	ICRAR, Curtin University, Australian SKA Regional Centre	Combine the data processing technology offered by next-generation supercomputers and novel data processing algorithms that have been optimised for both speed and sensitivity to transient signals. Enable real-time image-based transient searches (pulsars, gamma ray bursts and fast radio bursts). This will allow MWA and SKA-low to detect fast radio transients like spontaneously emitting neutron stars in real time.

Image credit: Rottnest Island Authority

ENGAGEMENT AND OUTREACH

PAWSEY FOR THE FUTURE AND THE NAMING OF SETONIX

Pawsey undertook a major rebrand in January 2021 to reinforce our connection with science and research, provide a nod to our origins, but also allow us to be seen by and engage with new audiences. This has helped set the tone both internally and externally for the step-change in performance expected of Pawsey in coming years as a result of the Capital Refresh. Known everywhere simply as Pawsey, you will still see us officially as the Pawsey Supercomputing Research Centre, to reinforce our role to serve Australian scientific advancement.

The name of our new supercomputer, Setonix, was unveiled in February 2021. Setonix is the Latin genus name of the Quokka, one of the world's happiest animals, a cultural icon and an ambassador of Western Australia on the world stage. It is also an animal that we have been able to assist, through our support of research mapping its unique genome for conservational purposes.

International coverage of our new supercomputing system has been broad, with more than 1,280 total news references to Setonix and Pawsey this year, from Germany to Japan to Korea. Among our key stakeholders, the name and potential of Setonix is now understood and recognised, even as the supercomputer is being installed. Our user population has also been galvanised and is enthused for the Setonix installation and the migration of our services from our existing supercomputers Magnus and Galaxy.

VIRTUAL EVENTS

As part of the vendor announcement for the new supercomputer Setonix, an online event for researchers was held to explore the opportunities that the new architecture will provide for them and the broader scientific community. The opportunities for extreme-scale computing showcased three researchers from diverse scientific backgrounds discussing their research projects and how infrastructure like Setonix could change the way their research is done for good. Almost 100 people joined the session, including representatives from government.

Following the same format, the Supercomputing Series panels were created to enable and engage communities of practice from scientific fields identified at Pawsey as key users of the current HPC architecture, or key groups that could benefit from the new architecture. The series is intended to address two fundamental questions: what does it mean to scale? And what does it take to scale? Participants share the challenges and new possibilities enabled by next-generation supercomputers and exascale in relation to their domains of interest.

The first event focused on the Computational Fluid Dynamics (CFD) community using Pawsey. As a result of this event, a Slack channel about CFD with HPC was created with the support of the Australian Fluid Mechanics Society to continue the participants' connection. The event was followed by Bioinformatics at scale and Australia's generation of supercomputers with 357 participants in February. Facilitated by Pawsey Executive Director Mark Stickells, the event panellists were renowned Australian researchers in bioinformatics and Dr Dan Jacobson from Oak Ridge National Laboratory, a 2018 Gordon Bell prize winner for achievements in high performance computing in the field.

Pawsey astronomy month took place in April. The online event Astronomy and Australia's next generation of supercomputers was hosted by Professor Alan Duffy, astrophysicist and Director of Space Technology and Industry Institute at the Swinburne University of Technology. The panel discussion was led by a stellar line-up of researchers from Curtin University, Leibniz Supercomputing Centre and CSIRO. Two hundred people registered for the event, with participants from nine different federal and state government departments, other NCRIS facilities and industry.

Pawsey continues to lead the biannual International Best Practice Newsletter, gathering contributions from HPC centres around the world on a chosen topic to release prior to the two major global supercomputing conferences, SC and ISC. Newsletters focused on COVID19 computing activities around the world and building communities in HPC. With a changed format from text to video, audio-visual contributions including CSCS, EPCC, NCI, NERSC, NeSI, PRACE and Pawsey elaborated on their roles in those two areas. Pawsey also hosted a joint event with NCI to showcase the outcomes of researchers who had accelerated access to allocations at Pawsey and NCI as part of our joint COVID19 initiative.

After a hiatus in 2020, Data Science Week returned in full force in 2021. With the support of the WA Data Science Innovation Hub and NCI, Pawsey saw an unprecedented 25 events during the week with over 350 attendees celebrating everything data science through demonstrations, workshops, tours, training, presentations and networking sessions across a range of disciplines including AI/machine learning, the resources sector, career journeys in data science, bioinformatics, industrial data science and high performance data.

The HPC Wire Readers' and Editors' Choice Awards were announced at SC20. This year four Pawsey projects were selected as finalists for best use of HPC in life sciences, in response to societal plights and data analytics and artificial intelligence.

DIGITAL LANDSCAPE

Over the year, Pawsey has seen drastic growth in event engagement online as a result of the evolving digital landscape. The Centre's accessible content has seen a global audience with global experts coming together around various subjects, resulting in unprecedented visitor numbers and visibility of the Centre. This digital shift has been supported by CSIRO and universities across Australia, who have increased the reach of Pawsey communications through a strong network of HPC champions.

This shift towards digital events has also resulted in an expectation that events are still available to access after they have concluded. These expectations are being met, and this availability is reflected in the growth of Pawsey's YouTube channel and the surge of training and supercomputing showcase video views.

WEBSITE

Pawsey's website is the primary portal for its engagement, and has been the foundation of sharing information, engagement and reporting at the Centre. New and returning website users have shown over 75 per cent growth in the financial year compared to the previous period. Peaks in engagement with the website were recorded during the call for students to apply for a Pawsey internship (resulting in an unprecedented number of applications), and quantum computing activities and announcements from March onwards.

YOUTUBE

With digital media trending towards videos for information-seeking, Pawsey's YouTube channel has side-stepped from its traditional researcher-focussed content, and seen significant growth in its training and engagement videos. The most engaged videos during the reporting period have been Pawsey's training playlists, followed by its scientific showcases. In a world where instant, visual information is in increasing demand, Pawsey is delivering with both its impressions and watch time increasing ten-fold.

USERS
76.99%
41,355 vs 23,366

NEW USERS
76.04%
40,381 vs 22,939

IMPRESSIONS AND HOW THEY LED TO WATCH TIME

Data available Jul 1 2019 – Jun 30, 2020 (366 days)

IMPRESSIONS	VIEWS FROM IMPRESSIONS	WATCH TIME FROM IMPRESSIONS (HOURS)
25.6K	**850**	**39.73**
3.3% click-through rate	2:48 average view duration	

IMPRESSIONS AND HOW THEY LED TO WATCH TIME

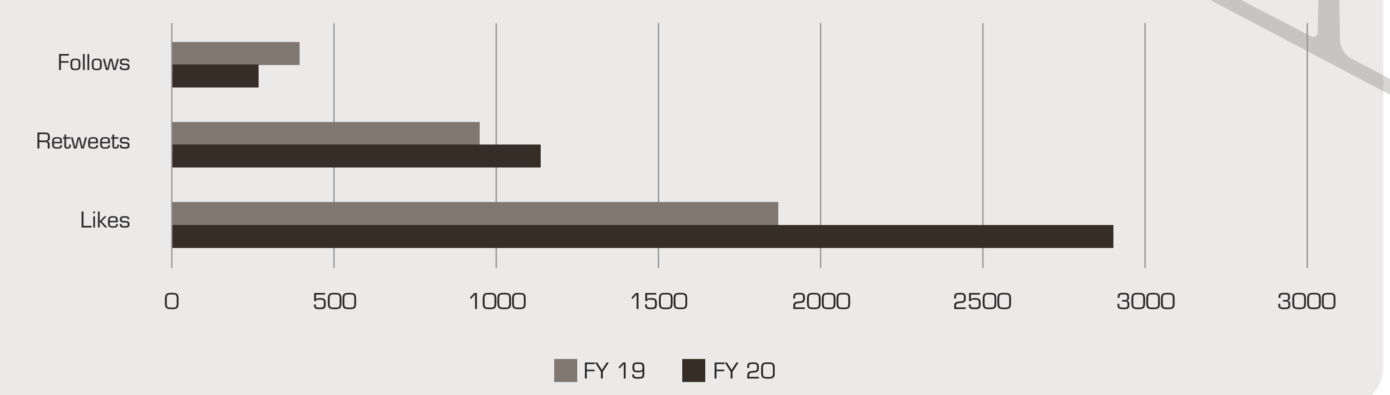
Data available Jul 1 2020 – Jun 30, 2021 (365 days)

IMPRESSIONS	VIEWS FROM IMPRESSIONS	WATCH TIME FROM IMPRESSIONS (HOURS)
247.8K	**7.5K**	**613.36**
3.0% click-through rate	4:53 average view duration	

TWITTER

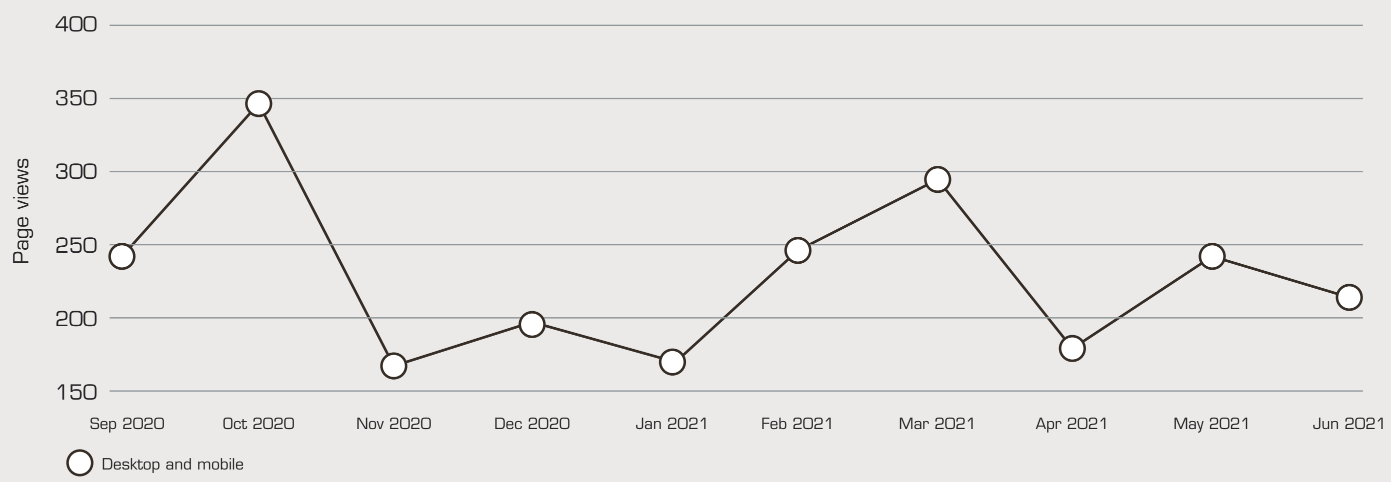
Twitter continues to grow as Pawsey's primary social medium to engage with technical and general audiences and communities in specialised scientific domains. Content such as Pawsey's impact video released during National Science Week, the announcement of Setonix, space data videos, and the bioinformatics at scale and Australia's next generation of supercomputers events garnered steady growth for the channel.

TWITTER GROWTH



LINKEDIN

Pawsey established its LinkedIn page in 2020. This is a priority platform given its professional reach and impact, and leverages the reach of the Executive Director's profile to showcase Pawsey milestones. In the nine months since its establishment it has had over 1,800 views, with highlights including the announcement of HPE to deliver Pawsey's new HPE Cray EX supercomputer Setonix, and receiving acknowledgment from the international SKA Observatory.



MAP OF PROJECTS ACROSS THE GLOBE

INTERNATIONAL COLLABORATIONS

SKAO

NERSC

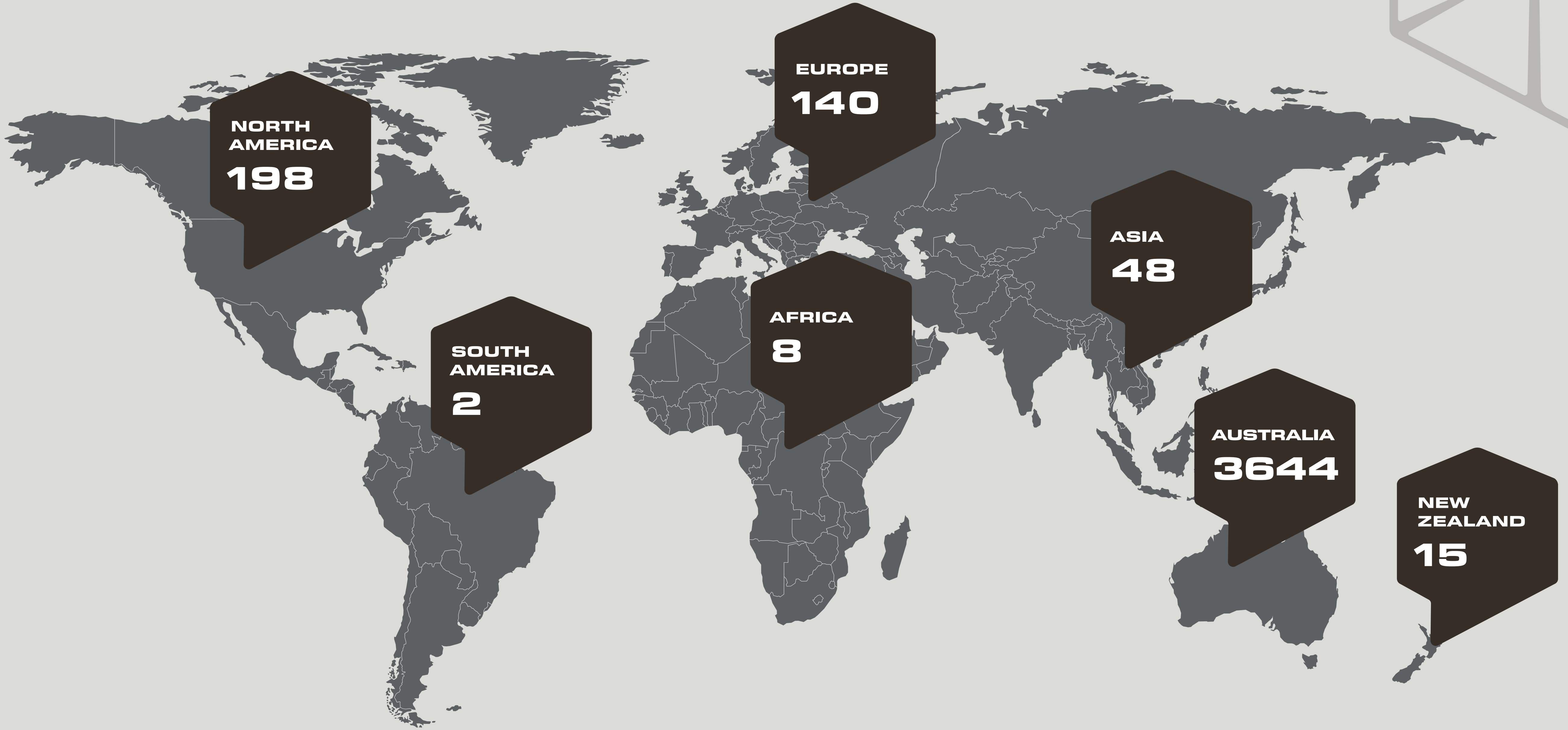
HPC-AI ADVISORY COUNCIL
NETWORK OF EXPERTISE

PRACE

epcc

CERN openlab

NSCC SINGAPORE National Supercomputing Centre



Financials

FINANCIAL REPORT 2020 – 2021

The Pawsey Supercomputing Research Centre is an unincorporated joint venture (UJV) between five institutions: The Commonwealth Scientific and Industrial Research Organisation (CSIRO), Curtin University of Technology (CUT), Murdoch University (MU), Edith Cowan University (ECU) and The University of Western Australia (UWA). CSIRO, the appointed Centre Agent of Pawsey, holds and manages its assets and finances. With this arrangement, the Pawsey Centre is required to adhere to CSIRO's governing legislation, constitution and adopted policies including its reporting, budgeting, and auditing framework.

In 2009, the Commonwealth Government as part of its Super Science Initiative, provided initial capital investment of $80 million for the construction of the Pawsey Supercomputing Research Centre's building and High-Performance Computing infrastructure. In April 2018, the Centre through Curtin University, received a $70 million Capital Refresh Grant from the Australian Federal Government to secure the next generation of supercomputers, data infrastructure and service upgrades. This is financially managed and reported by Curtin University, one of the UJV member of Pawsey. From the start of the project a total of $9.2 million of asset procurements through this capital refresh program have been transferred to Pawsey's Fixed Assets Register as at end of June 2021.

The operational requirements of the Pawsey Supercomputing Research Centre are continuously supported by both the Commonwealth and the State Government of Western Australia. The Commonwealth Government made available its funding through the National Collaborative Research Infrastructure Strategy (NCRIS) program run by the Department of Education, Skills and Employment. The State Government of Western Australia's funding through the Department of Jobs, Tourism, Science, and Innovation supports the Centre with its Financial Assistance Agreement. The UJV partners contribute to the Centre and share the site operations costs through the agreed cost-sharing model.

At the start of June 2020, Pawsey was able to attract additional funding from the Federal Government for the Australian Space Data Analysis Facility (ASDAF) in Western Australia, co-funded by the Western Australian Government. This project is a collaboration between Pawsey and WA Data Science Innovation Hub (WADSIH) which is managed by Curtin University. The project aims to harness the space data that will enable Australian SMEs and researchers to access, understand, and use space data to make important decision and develop new products and services.

Pawsey, through CSIRO, employs professionals and experts in High Performance Computing to run operations of the facility and provide exceptional services on supercomputing access, cloud, data storage and visualisation, training, and consulting. As a Tier-1 High Performance Computing facility in Australia, Pawsey Supercomputing Research Centre's function is to accelerate scientific research for the benefit of the nation by enabling research in areas of astronomy, bioinformatics, space, and other national and state science priorities.

In this reporting period, Pawsey ventured into Quantum Computing which makes the centre to be one of the first supercomputing centres globally to host a universal quantum computer onsite. Pawsey's Quantum Pioneer Program is geared towards developing cutting-edge quantum applications in machine learning, logistics, defence, quantum finance and quantum research which will further accelerate scientific workflows using the potential of quantum computing within the advanced infrastructure of a national research supercomputing centre. The continuance of these funding support is essential for the Centre's existence.

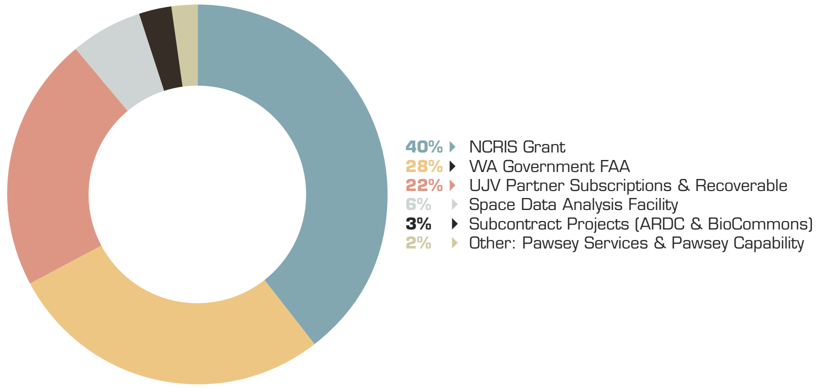
The funding model for Pawsey aims to reflect the proportionate usage of machine on research projects of which funding streams are mainly coming from the Federal, Western Australia State Government and the UJV partners. The Federal Government through its NCRIS program provided around 40% of the total funding source, followed by 28% from the State Government of Western Australian whilst 22% is provided by the UJV partners. The ASDAF project contributed 6% and subcontract projects from Australian Research Data Commons (ARDC) and University of Melbourne contributed 3% whilst 2% comes from other sources such as Pawsey Services and Pawsey Capital Refresh.

CSIRO, as the Centre agent also contributed to the running of the facility by providing administrative support services and facility maintenance. This contribution is considered in-kind as it is not paid through Pawsey's external funding sources and is part of CSIRO's appropriation budget. For the year, CSIRO's Enterprise Support Distribution totalled $6.03 million to the Pawsey Supercomputing Research Centre. This is not reflected in the above Statement of Income and Expenditure.

The major cost driver for the Centre is labour which is 50% of the total expenditure. This represents the staffing costs for technical experts and allied support services which are essential in running the facility. Pawsey's diverse team of Technical experts are recruited nationally and internationally. Machine Maintenance and Other Operating expenses are each at 14% of the total costs whilst utility cost is at 11% which includes electricity, gas, and water usage. The Centre receives electricity credits from the solar panels installed at the building. Asset procurement for minor hardware and network upgrades comes next at 11% funded from NCRIS grant. These are essential minor upgrades to the existing infrastructure to continue usage when capital refresh is still ongoing. Part of the surplus indicated in the report are asset procurements which are work-in-progress as at end of June.

FUNDING %

Figure 1: Funding Source



EXPENDITURE %

Figure 2: Expenditure Details

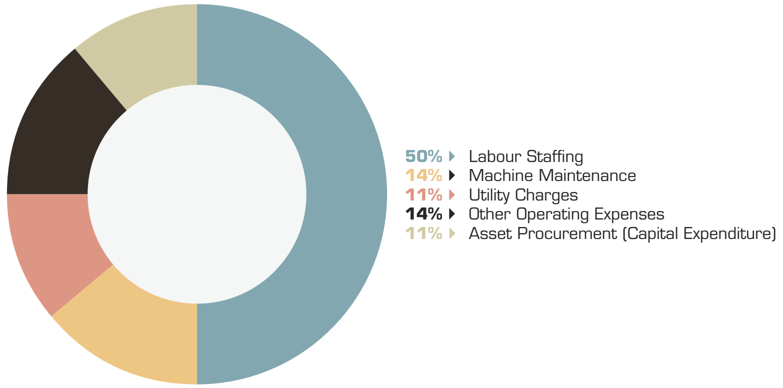


Table 1: Financial Report for FY20/21
Pawsey Supercomputing Research Centre
Statement of Income and Expenditure
01 July 2020 to 30 June 2021

	$'000
Income	
External Project Funding	12,647
Internal Joint Venture Partner Subscriptions and Recoverable	3,524
Total Income	**16,171**
Expenditure	
Labour staffing	7,688
Machine Maintenance Contracts	2,073
Utility charges	1,655
Other Operating expenses	2,149
Asset Procurement (Capital Expenditure)	1,730
Total Expenditure	**15,295**
Surplus / (Deficit)	**876**

*Note: Surplus refers to activities which are work-in-progress as at year end and will be carried forward to FY21/22.

The Pawsey Supercomputing Research Centre is supported by the Australian Government through a $70 million grant made under the Industry Research and Development Act and administered by the Department of Industry, Innovation and Science. Pawsey is also supported by the Australian Government under the National Collaborative Research Infrastructure Strategy (NCRIS) and related programs through the Department of Education. The Centre would also like to acknowledge the support provided by the Western Australian Government and its Partner organisations.

NCRIS National Research Infrastructure for Australia An Australian Government Initiative

GOVERNMENT OF WESTERN AUSTRALIA

CSIRO

Curtin University

AUSTRALIA ECU EDITH COWAN UNIVERSITY

Murdoch UNIVERSITY

THE UNIVERSITY OF WESTERN AUSTRALIA

pawsey.org.au