making a difference through data-driven research

Australian Research Data Commons
CEO’s message

The Australian Research Data Commons (ARDC) was formed on 1 July 2018. In May 2019, the ARDC became a company limited by guarantee.

The ARDC’s history is relatively short. However, through its legacy initiatives such as the Australian National Data Service (ANDS), National eResearch Collaboration Tools and Resources (Nectar), and Research Data Services (RDS), the ARDC brings to the research sector over 10 years’ experience in research data infrastructure and services.

The purpose of the ARDC is to provide Australian researchers with competitive advantage through data, providing access to leading edge data intensive infrastructure, tools, services and collections of high-quality data.

The mission of the ARDC is to accelerate research and innovation by driving excellence in the creation, analysis and retention of high-quality data assets.

This brochure highlights the various ways the ARDC has supported Australian researchers and their research, and in turn, how this has translated into environmental, economic and social benefits for Australian and international governments, a range of industries and the broader community.

Rosie Hicks
Chief Executive Officer, Australian Research Data Commons
Global data sharing is helping Australian kids battle cancer

THE OUTCOME
Improving the lives of children with rare cancer

Eleven-month-old Ellie was admitted to the hospital with a tumour in her chest and was enrolled in the program. A team rapidly identified the likely cause of her cancer and compared her genetic information with global data. They identified a new US anti-cancer drug that specifically targeted her case. After several weeks, Ellie was able to leave intensive care.

“The innovative technology of CAVATICA enables researchers from Australia and the United States to seamlessly share data and novel analysis methods with ease, driving improved outcomes and novel research,” said Associate Professor Mark Cowley, Computational Biology Group Leader at Children’s Cancer Institute.

Support for the program continues to build. The Australian Government and the Minderoo Foundation recently announced an additional $67 million to expand the Zero program. This enables all Australian children and young adults who have been diagnosed with cancer to access targeted treatment through the collaborative research and clinical program.

THE APPROACH
Creating a seamless, secure way of accessing and sharing global data about rare cancers

The ARDC, along with the Australian BioCommons and other groups, are providing funds, facilities and expertise to harmonise different computational systems, software and data which enable timely genomic information processing.

One result is a cloud-based platform, CAVATICA, which provides a coordinated analysis of Australian cancer information with data from other countries. It allows scientists worldwide to rapidly access large amounts of genomic childhood cancer data and workflows where they can share, process, integrate and analyse data.

The platform also harmonises information without physically moving any data, alleviating issues related to the security of genetic data.

The ARDC and partners have worked together to expand and strengthen access and data sharing between Australian initiatives like Zero Childhood Cancer with international organisations, including CAVATICA.

THE PROBLEM
Insufficient information to treat Australian children with cancer effectively

Every year around 200 young Australians are diagnosed with a rare, high risk type of cancer with low survival rates. It’s 200 too many, but too small a number to accurately analyse data and identify effective treatment methods.

However, a child is diagnosed with cancer every two minutes somewhere around the world. Finding other patients with a similar mutation means a better chance of finding a successful treatment for each Australian case.

THE OUTCOME
Improving the lives of children with rare cancer

Sydney Children’s Hospital, Randwick recently implemented a new clinical trial program called the Zero Childhood Cancer Program. It uses rapid data access and sharing through the CAVATICA platform, allowing medical staff to quickly access global data related to new cancer cases.

As soon as a child is diagnosed and enrolled in the trial, researchers analyse the genetic and molecular make-up of their cancer cells and test hundreds of options to identify which drugs are most likely to be effective. Once a promising drug is identified, it is tested in living laboratory models to see how it performs. The results are used to develop a personalised treatment plan.

The trial is showing promising results. During the first 12 months, nearly 70 per cent of participating children were able to benefit from these rapidly produced treatment plans. For some, the trial has been life saving.

Global data sharing is helping Australian kids battle cancer

THE OUTCOME
Improving the lives of children with rare cancer

Eleven-month-old Ellie was admitted to the hospital with a tumour in her chest and was enrolled in the program. A team rapidly identified the likely cause of her cancer and compared her genetic information with global data. They identified a new US anti-cancer drug that specifically targeted her case. After several weeks, Ellie was able to leave intensive care.

“The innovative technology of CAVATICA enables researchers from Australia and the United States to seamlessly share data and novel analysis methods with ease, driving improved outcomes and novel research,” said Associate Professor Mark Cowley, Computational Biology Group Leader at Children’s Cancer Institute.

Support for the program continues to build. The Australian Government and the Minderoo Foundation recently announced an additional $67 million to expand the Zero program. This enables all Australian children and young adults who have been diagnosed with cancer to access targeted treatment through the collaborative research and clinical program.

THE APPROACH
Creating a seamless, secure way of accessing and sharing global data about rare cancers

The ARDC, along with the Australian BioCommons and other groups, are providing funds, facilities and expertise to harmonise different computational systems, software and data which enable timely genomic information processing.

One result is a cloud-based platform, CAVATICA, which provides a coordinated analysis of Australian cancer information with data from other countries. It allows scientists worldwide to rapidly access large amounts of genomic childhood cancer data and workflows where they can share, process, integrate and analyse data.

The platform also harmonises information without physically moving any data, alleviating issues related to the security of genetic data.

The ARDC and partners have worked together to expand and strengthen access and data sharing between Australian initiatives like Zero Childhood Cancer with international organisations, including CAVATICA.
The bleak outlook for reef sharks is now in the spotlight, thanks to accurate data

**The Problem**
A lack of data on global shark populations made it impossible to assess their future prospects.

Decades of exploitation have devastated shark populations, leaving doubt about their long-term survival.

Sharks play an important ecological role, especially around coral reefs, where they help promote a healthy biodiversity. Reef sharks also enhance the economies of small island nations via tourism.

However, there was little factual data about the status of shark populations, especially around coral reefs. Concerns about the impact of climate change and depletion from fishing signaled a need for accurate information.

**The Approach**
Using platforms and data services to manage and interpret vast amounts of data.

A global research project set out to assess the status of the global shark population. It involved the cooperation and collaboration of hundreds of scientists, researchers, and conservationists who captured and analysed more than 15,000 hours of video from surveys of 371 reefs in 58 countries, states, and territories over four years.

Translating the video footage into materials that could be shared and analysed across the project required extensive data services and analysis tools.

Australian project participants relied on the ARDC supported GlobalArchive, hosted on the ARDC’s Nectar Research Cloud, to manage the massive quantities of data and make them accessible to all project members.

It also allowed researchers to gauge Australia’s shark conservation efforts on a global scale.

**The Outcome**
The world now has its first factual database documenting the global demise of reef sharks.

The study revealed the profound impact fishing has had on reef shark populations. Sharks were almost completely absent from reefs in several nations; 20 per cent had none.

It also found that the depletion was strongly related to socio-economic conditions, including the size and proximity to the nearest market, quality of governance and population density.

“Fortunately, Australia is a country where shark populations on coral reefs are still largely intact,” Dr Tim Langlois, from University of Western Australia’s School of Biological Sciences and Oceans Institute, explained.

Australia is one of a few nations where conservation efforts, which include either banning or sharply limiting shark fishing, have been effective.

The study pointed the way for improvements.
THE PROBLEM
Finding and accessing Indigenous data

A number of organisations, including research groups, galleries, libraries, museums and service providers hold a wealth of data generated by and about Indigenous Australians, their ancestors and the communities.

However, there was no system of knowing how these groups were collecting, storing and distributing this information.

THE APPROACH
Creating a framework for Indigenous data collections

The Indigenous Data Network (IDN), which works with Indigenous communities to develop technical capability and resources for managing their data, wanted to fix this conundrum.

The group developed a project, which was supported by the ARDC, to discover how existing widespread data about Indigenous Australians could be linked together. They also wanted to create a national technical and governance framework for data collections, including improvements in data integrity, accuracy and scale.

The project team conducted a national survey to document the extent, substance, integrity and security of this data and identify what was required to turn the isolated collections into FAIR (Findable, Accessible, Interoperable and Reusable) data assets.

THE OUTCOME
Providing Indigenous Australians with better access to their cultural heritage and improving local community services

The project led to a series of new research database platforms that give individuals and communities more effective use and oversight of their cultural heritage.

Having easy access to this data allows the community to better address their specific needs, rather than relying on outdated government statistics.

The project also prompted a nationwide recognition of the importance of Indigenous data collections.

In January 2020, the Prime Minister Scott Morrison announced that the IDN, in collaboration with the Coalition of Peaks, would be implementing a new priority reform for the updated National Agreement on Closing the Gap, to ensure that Aboriginal and Torres Strait Islander people have access to locally relevant data and information.

This included the government providing IDN with $1.5 million to build a pilot program, including a new platform to provide individual portal access to the widespread organisations.

“The ARDC project is testimony to what can be achieved with a small amount of funding,” said Dr James Rose, the project leader and Senior Research Fellow at the University of Melbourne.
THE APPROACH
Using data to create an accurate way of estimating crop yield and the best time to harvest

Researchers from Central Queensland University’s Institute for Future Farming Systems were supported by the ARDC in the development of an innovative product called FruitMaps, which translates data from sensors located in mango farms onto maps and into tables. The data includes an array of information, including temperature, fruit assessment and processed tree images to help farmers decide when to harvest their mangoes.

THE OUTCOME
The new product enables growers to pinpoint the best time to pick their crops

The data-driven system has been widely adopted within Queensland’s mango industry, providing farmers with a more accurate, efficient and cost-effective way of managing their mango production.

Farmers can now better estimate the size of their crops and the best time for harvesting, allowing them to employ the optimal number of pickers and packers at the ideal time.

Consumers benefit from having better mangoes. Farmers benefit from reduced costs and improved performance — up to 40 per cent improvement for some mango farmers.

The project’s success laid the foundation for further improvements for mango farmers. The research team has now developed an automatic fruit harvester, based on the data services and technology running FruitMaps.

The prototype takes only five seconds to harvest a mango.

Fruit growers in Queensland have described the new data-driven technology as a game changer.

“This end goal is to save costs and improve productivity on the farm, while driving consumer demand by ensuring a top-quality eating experience every time,” a satisfied mango farmer noted.

“Knowing how much fruit is in that block, when it’s going to be mature and the size of the fruit, means we can schedule our workforce, order the right number of cartons and the size of the carton inserts — this could be a real game changer, not only for our farm but for the entire industry.”

THE PROBLEM
Predicting the optimum harvest time and estimated yield for their fruit

Timing is everything to Australian mango farmers. They regularly face the challenge of deciding when to harvest their fruit, which affects how many fruit pickers and packers they need and when to market their fruit.

This could be a real game changer, not only for our farm but for the entire industry.
A virtual lab helps premature infants have healthier lives

THE PROBLEM
Doctors are reluctant to use clinically proven drugs on premature babies due to potential side effects.

Premature babies often have breathing problems and chronic lung disease, a result of having immature lungs and prolonged ventilation.

To alleviate these serious health issues, neonatal providers are now looking closely at the use of non-invasive respiratory treatments. A group of drugs are known to promote lung development in these infants, but doctors have been reluctant to use them due to concerns about neurological harm.

THE APPROACH
Investigating the real impact of these drugs with the help of high-powered data processing and management.

A team of researchers led by Nathanael Yates, an Australian brain specialist and research fellow at the Queensland Brain Institute, used high powered magnetic resonance imaging (MRI) to investigate the impact of different drug therapies on brains.

The study, which required extensive computer capacity and imaging, was accomplished with the help of an innovative tool called the Characterisation Virtual Laboratory (CVL).

The CVL, which is supported by the ARDC, streamlines the integration of imaging instruments with specialised high-speed computing and tools. As a result, researchers like Nathanael can access the data quickly, allowing them to run and modify experiments in real time, without having any special computing skills.

Having access to this high speed, integrated tool allowed Nathanael and his team to quickly and accurately assess the neurological safety of these drugs.

THE OUTCOME
Doctors and parents now have more accurate information about the safety of crucial, life-saving drugs.

The research provided physicians with evidence that the treatment had no obvious negative effects on brain structure within certain timeframes.

The results are helping alleviate concerns about these vital drugs and allowing them to be used to improve the health of premature infants.

Additionally, this innovative use of data and technology is continuing to uncover further insights into neuropathology.

The results are helping alleviate concerns about these vital drugs.
THE PROBLEM
A collection of irreplaceable geological information was inaccessible to those who needed it and under threat of disintegration.

Geoscience Australia (GA) has accumulated more than 250,000 slides, including handwritten entries from extensive mapping programs and scientific expeditions. The information is useful for geological surveying and also plays a valuable role in calibrating remote sensing data from satellites, drones and other surveying methods. Because the slides were created and stored in the pre-digital era, this valuable data was both invisible and inaccessible. And the clock was ticking on their future due to threats from physical deterioration. However, the cost of making the information available and safeguarding its future was prohibitive—an estimated $100 million.

THE APPROACH
Using a crowdsourcing platform to allow volunteers to retrieve, document and make the information easily available.

An innovative crowdsourcing platform called DigiVol, which was developed by the Australian Museum and is hosted by the Atlas of Living Australia, was called into play to help solve the problem. DigiVol allows an army of volunteer citizen scientists who don’t have special research or data skills to input an array of information, from field notebooks to photographs, into digitally accessible data. A research team, which included an ARDC data management expert, used the platform to develop a user-friendly geoscience project. This allowed volunteers to help transcribe thousands of GA’s handwritten entries into usable data.

THE OUTCOME
Nearly 100,000 slides have now been digitised and are freely available.

Access to the information only takes a few minutes. The data is designed around Findable, Accessible, Interoperable and Reusable (FAIR) principles, so it can be used by other external platforms and systems.

Industry and academic researchers now have easy access to essential geological information, including the presence of minerals and their formation, without incurring expensive field time. Developing data-rich, user-friendly avenues that allow people without special training to help secure important information for current and future generations is a powerful tool for the Australian community.

CASE STUDY / SOLVING A PROBLEM WITH A NEW PLATFORM
Citizen scientists are providing essential geological information through a cost saving platform

This valuable data was invisible and inaccessible.
The Australian Research Data Commons is enabled by NCRIS.

**CONTACT**
Visit ardc.edu.au
Call +61 3 9902 0585
Email contact@ardc.edu.au

**SUBSCRIBE**
To our newsletter and discover collaboration opportunities

**FOLLOW US**
@ARDC_AU
linkedin.com/company/australian-research-data-commons

Providing Australian researchers with competitive advantage through data.

The Australian Research Data Commons is enabled by NCRIS.

**ARDC MEMBERS**

This product is made from 100% recycled materials, 100% recyclable.