THE UNIVERSITY OF AUCKLAND
Te Whare Wānanga o Tāmaki Makaurau
NEW ZEALAND

Centre for eResearch
Annual Report 2017

ACHIEVE THE AMAZING

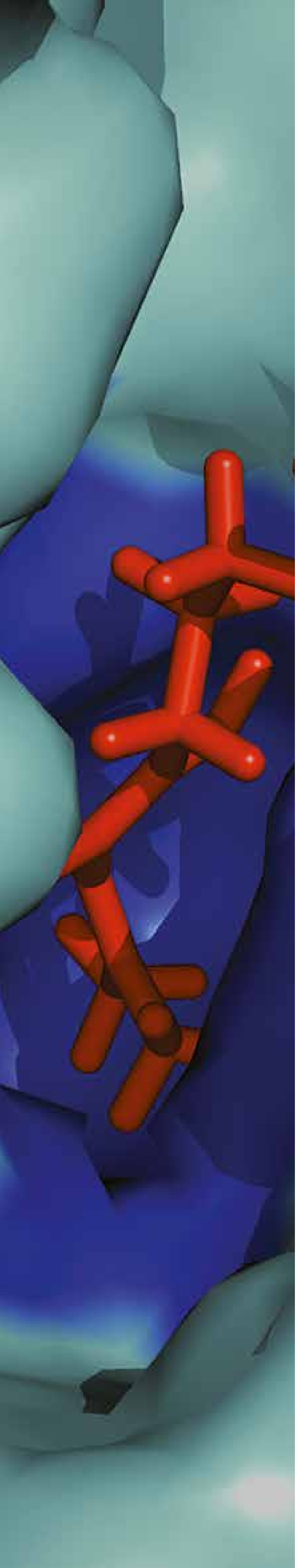


TABLE OF CONTENTS

MESSAGE FROM THE DIRECTOR

In this report, you will find evidence of the impact of the Centre for eResearch on scholarly activity across the University. It describes how our core services, consultancy and skills development initiatives have helped to increase both the effectiveness and the efficiency of our researchers, and perhaps also in a small way, our research culture. The many positive outcomes over the past year are, I believe, a testament to our hard work and skills, and also to the ability and creativity of our research community. Helping to stimulate and nurture these shared capabilities is hugely satisfying, and I am grateful for such an excellent team to work with and for the trust shown us by the research community. Specific highlights to mention include the successful launch of the *Research Hub*, the *Figshare data publishing service* going into production (with 10,000 downloads so far), the creation of a self-service *Nectar research cloud*, the *Jupyter Notebook* for supporting research-led teaching and learning, and interactive visualisation of bio-chemical models for advanced tutorials using personal VR headsets.

Of all the highlights achieved this year, the one that encourages me the most is the week-long Winter Boot Camp we offered in July, with over 26 workshops and learning opportunities, from Software Carpentry to how to increase your impact as a researcher. We had over 400 course participants in these workshops from all corners of the University, including senior academic staff as well as research students. Almost everyone in the eResearch team played a role in delivering this event, along with colleagues from the Library and from NeSI. More details can be found at:
https://uoa-eresearch.github.io/winterbootcamp2017/

We took time earlier in the year to relocate to our lovely new office space in the New Science Centre (302), rooms 585, 590 and 520 on the 5th floor is where you will find us now. We now also have a dedicated Visualisation Suite (room 580) that is well worth a visit. We are now settled in and have expanded to fill the available space.

During the course of 2018, we will be transitioning from temporary to ongoing funding for the team delivering our core services. That’s 11.5 FTE – a huge commitment from the University for which we are truly grateful. The stability this change will provide for both our staff and for the researchers they support is most welcome.

Professor Mark Gahegan
Director, Centre for eResearch

ASPIRATION

Recognising that computational tools and IT play an increasingly important role in scholarship and research of all kinds, we must support our researchers to utilise this technology better, in order to provide our community with a research advantage.

The following Research IT Core Values drive our objectives:

1. Computing power is easily accessible to researchers, in the right form for them to use;
2. research data is a valuable, long-term asset that requires carefully managing its value;
3. researchers require new technology skills and competencies to embrace fully the digital transformation occurring in research;
4. advanced computational methods provide a research advantage so it is important that the University is ready to adopt emerging tools and methods that improve research effectiveness;
5. good research support services are easy to find, to request, and to use.

OBJECTIVES

From these aspirations, we set the following long term objectives:

1. Provide a comprehensive set of research computing platforms and services to address established and emerging needs across the University with enough flexibility and scalability to address complex research problems effectively. Partner with NeSI (nationally) and NeCTAR (in Australia) to ensure the available compute platforms are efficient and meet our needs.
2. Offer researchers timely and relevant IT advice and support that they need to plan and execute their research well.
3. Provide a comprehensive set of research data services to manage, archive, share and publish research data created by the University research community. Offer training to enable researchers to manage their research data more effectively and encourage the creation and operation of a suitable Data Management Plan (DMP).
4. Enable researchers to communicate, collaborate and share information with colleagues locally, nationally and internationally.
5. Establish a culture of skills development and peer support within the research community via appropriate training, education and community-building opportunities so that our researchers are better equipped to leverage advances in computational technology.
6. Provide an information visualisation and analytics service, to foster discovery, and to improve the impact and communication of our research. Help researchers to leverage these capabilities in the pursuit and communication of their research.
7. Lead the development of a Research Hub that forms a central repository of research computing service descriptions, a means by which researchers can learn about and request these services and the ability to link services to research outcomes enabled by their use, as well as encourage better cross institutional collaboration in delivery of research support.

RESEARCH IT SERVICES

The University Image bank - Science student

The University image bank - Engineering wind tunnel

In order to support the University's ambitious strategic research goals, the Research IT services aim to provide support for researchers with access to tools, support and education that enhance the broad range of research undertaken at the University. These include aspirational computational facilities, collaboration sites and tools, data protection and management, ability to publish and to catalog research profiles and skills, and to support collaborative research activities across institutions nationally and internationally.

It is important to recognise that there is a broad range of research conducted at the University which requires similarly a broad range of computational services to support it.

The Research IT services offer compelling advantages to researchers: from scientific computing to deep learning, to computational workflow technology, to data mining and knowledge discovery methods, to information visualisation methods and equipment, to supporting open research along with sharing, and long-term archiving of data. In addition, we regularly scan for new and emerging technologies and services; evaluate them and bring those which offer strong promise into production.

The new Wanhal catalogue - School of Music

The Research IT services are planned and provided by a variety of operational groups across the University and nationally, including the Centre for eResearch (CeR), New Zealand eScience Infrastructure (NeSI), the Faculty IS (FIS), the Library, Improve Research Support Service (IRiS) group, the Research Office, Central IT (ITS), and New Zealand Research and Education Network (REANNZ).

In response to a growing skills gap between what researchers know and what they need to know in order to operate effectively as 21st century scholars, we offer leadership and training in digital literacy and scholarship, and in aspirational methods to enable and improve the research process in the following areas:

1. Training in Software and Data Carpentry,
2. assisting researchers with Open Data and open research initiatives,
3. providing deep learning and information analytics,
4. supporting computational workflow technologies and research reproducibility, and
5. using best-practice approaches to sharing data, methods, code and outcomes,
6. continue to expand the institutional capability and knowledge via educational and outreach offerings to research staff and students, including short courses, annual Boot Camp and Research Bazaar events.

OUR AIMS FOR 2017-2018

The Centre for eResearch has been working with the ITS, FIS, the Library, and the IRiS to develop a delivery model for Research IT service with clear roles and responsibilities, and increased intra-institutional collaboration and service maturity. The following are the key elements of the service area strategy:

Research IT advisory / Research Hub

Develop and maintain a Research advisory and Research Hub that directs research staff and students to appropriate services, resources and information, all described clearly, with a well-defined support structure, and agreed entitlements.

Ensure the Hub supports improved orchestration of technologies and research capability, and enable self-service and automation where feasible.

Train support staff across CeR, Faculty research support staff, FIS, ITS and the Library to help connecting people, processes and infrastructure via the Hub.

Support researchers in jointly addressing their needs, and augment existing research management systems to include capturing research project details, service utilisation, and linkage to research outcomes.

Research computing platforms

Create standard services for researchers which include clearly articulated IT service and infrastructure entitlements.

Provide a range of computing platforms and services to researchers, including Infrastructure as a service (IaaS), NeSI High Performance Computing (HPC) facilities, a research virtual machine (RVM) farm, Docker containers (e.g. Jupyter Notebook), an OpenStack cloud service (Nectar) and a deep machine learning (via GPU) service.

During 2017, we deployed the Nectar virtualised research computing infrastructure in collaboration with Nectar in Australia, to leverage international progress in sharing entire research code “stacks” and virtual laboratories.

Continue to work with researchers to optimise strategically important HPC codes and applications.

Ensure that all the HPC and storage facilities continue to meet institutional needs and demand, and proactively expanding facilities to meet growth. Prioritise needs where demand cannot be met.

Visualisation & analytics

Expand the visualisation service to include virtual and augmented reality display capabilities that can be used to support research-based teaching. In order to meet the growing demand for visualisation support services, a diversity of visualisation and analytics libraries and tools are to be added to address problems beyond STEM subjects.

Professor Stuart McCutcheon, Vice-Chancellor put his research data in Figshare

Increase communication and outreach on the use of this service. Develop workflows for image and 3D data capture, to foster easier digitisation of research artefacts and creation of virtual spaces.

Research data management

With limited storage available in 2017, our focus has been on managing research data across the University. This needs to be broaden to include all digital research artifacts including text, numerical files, databases, computer code, images, videos, workflows, and virtualised environments.

The goal is to provide all researchers with services that meet their storage and research management needs. This includes sharing and using 'working objects', assuring confidentiality for archiving research objects, and 'publishing' finished objects for others to find and use. Promote the use of research object lifecycle services, data management plans (DMPs), code repositories and related productivity tools.

Digital skills development and community building

Prioritise the development of digital research skills among research staff and students, through short courses (Software & Data Carpentry), community outreach and annual learning events such as Research Bazaar (ResBaz) and Boot Camps.

Develop digital scholarship materials that can be delivered from within existing (or new) university courses, building on the work of Research IT Literacy project through the Vice-Chancellor Strategic Development Fund (VCSDF).

Create a dedicated team for digital skills development and community building, to provide learning opportunities in order to keep up with demand. Develop innovative ways to engage beyond traditional STEM subjects, and to ensure that our researchers can maximise their effectiveness in the fast-changing landscape of research IT.

Collaboration with researchers

Provide services to researchers that enable collaboration and the sharing of research ideas within the University and across other institutions nationally and internationally. This includes web-based collaboration services such as WordPress sites and similar information portals and virtual laboratories as well as immersive and collaborative visualisation environments. Foster better collaborative culture via new services and coordinated activates (such as the operation of the Research Hub, service design activities, hack days etc). Act to build more trust and cooperation between different organisational units that involved in the delivery of research IT services and support.

High speed networking

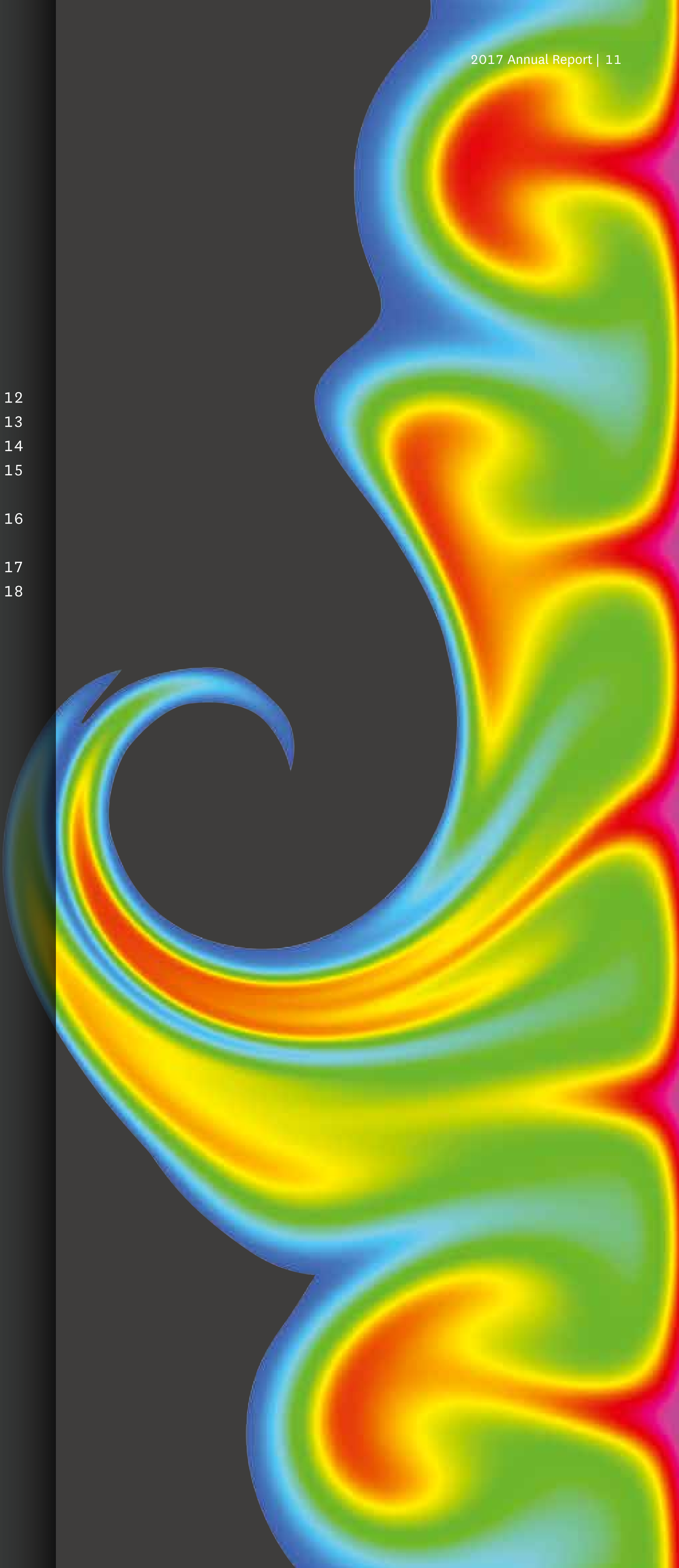
The Centre is working with ITS to identify researchers and instruments with high bandwidth needs to enable faster connectivity by prioritising the upgrading of edge switches to these sites, providing a research data overlay across the UoA campuses.

Plan, develop and deliver a research data network backbone and demilitarised zone (DMZ) in order to enable very fast connectivity between our compute resources, including the new NeSI facilities at Greta Pt in Wellington, and the RVM farm at the Tamaki Data Centre, the Nectar cloud, the research instruments, and research data storage.

Alignment with the IT Strategic Plan 2016-2018

1. Provide a research IT advisory service that addresses all aspects of IT needs encountered by researchers and connects them to the spectrum of appropriate services.
2. Significantly enhance IT services and infrastructure to support research data management, collaboration, research computing, and visualisation.
3. Optimise publications and research outputs, data, projects, citations, and grants both internally and externally.
4. Enhance research management systems and processes to cater the needs of researchers and administrators.
5. Provide training and education opportunities that enable researchers to develop an appropriate level of IT skills and practices.
6. Partner with faculties to meet their specific technology needs.

WHAT WE DELIVERED IN 2017



Research IT advisory and Research Hub

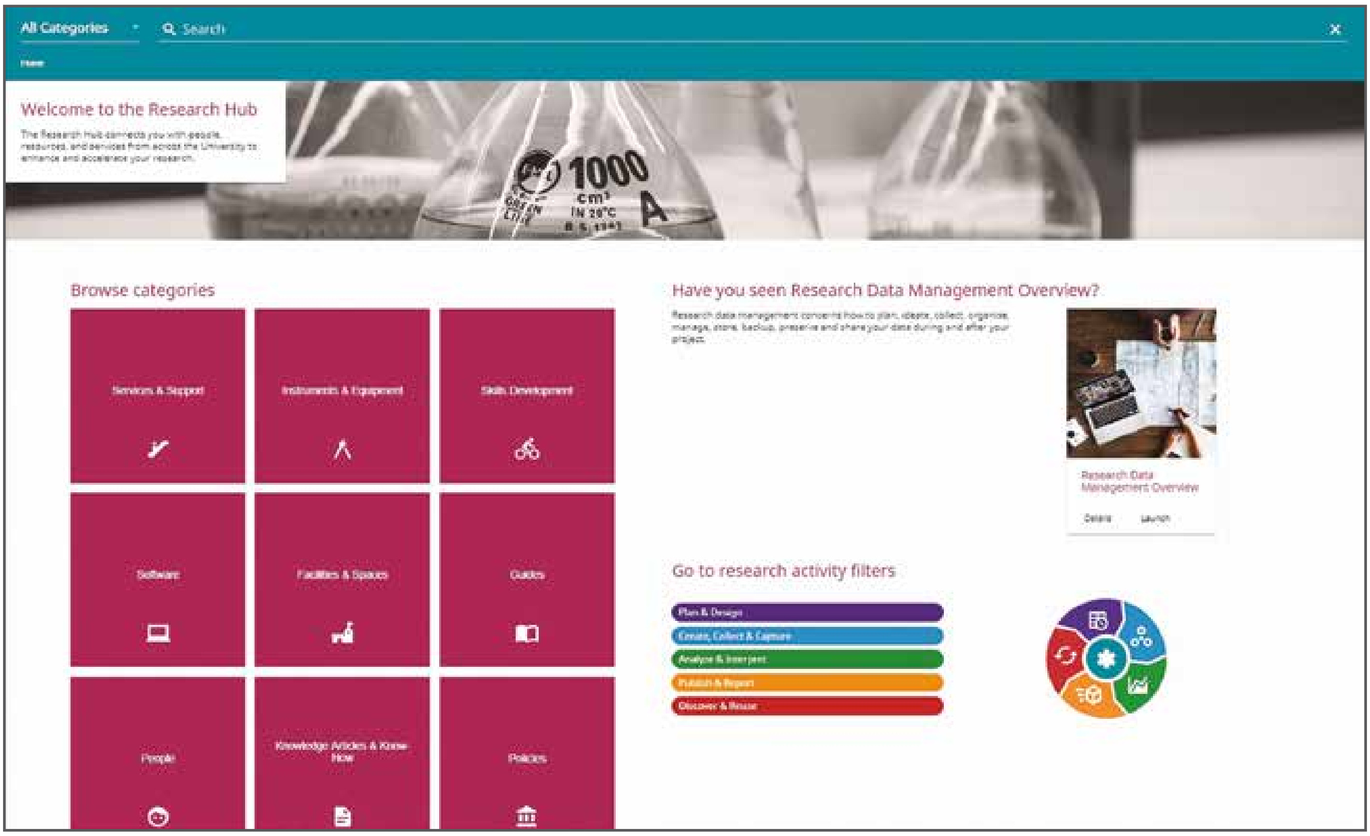
3 service platform workflows have been digitally transformed.

30 new service requests have been captured and projectised.

70 research services are fully described.

700 visits to the Research Hub since August 2017.

1 fully functioning and accessible Research Hub



https://research-hub.cer.auckland.ac.nz

Research computing platforms[1]

100+ new projects across 10 departments.
300+ total researchers supported to-date.

750+ Jupyter Notebook users supported.

Workflow processing time reduced by 50%.
94% average user satisfaction.

1 self-service Nectar cloud to leveage virtual lab and facilitate international collaboration.

The Jupyter logo is attributed to Jupyter Project https://jupyter.org
The Nectar logo is attributed to The National eResearch Collaboration.
Tools and Resources project https://nectar.org.au

[1] *These figures do not include NeSI HPC platform.*

Visualisation & analytics

22 different departments across the University have used the visualisation and analytics services.

31 research projects assisted with visualisation and analytics services.

70% utilisation rate for the new Visualisation facility since it's been relocated to the new Science Precinct.

1 advanced wearable technology (ODG R-7 AR smartglasses).

6x4K tiled-wall displays.

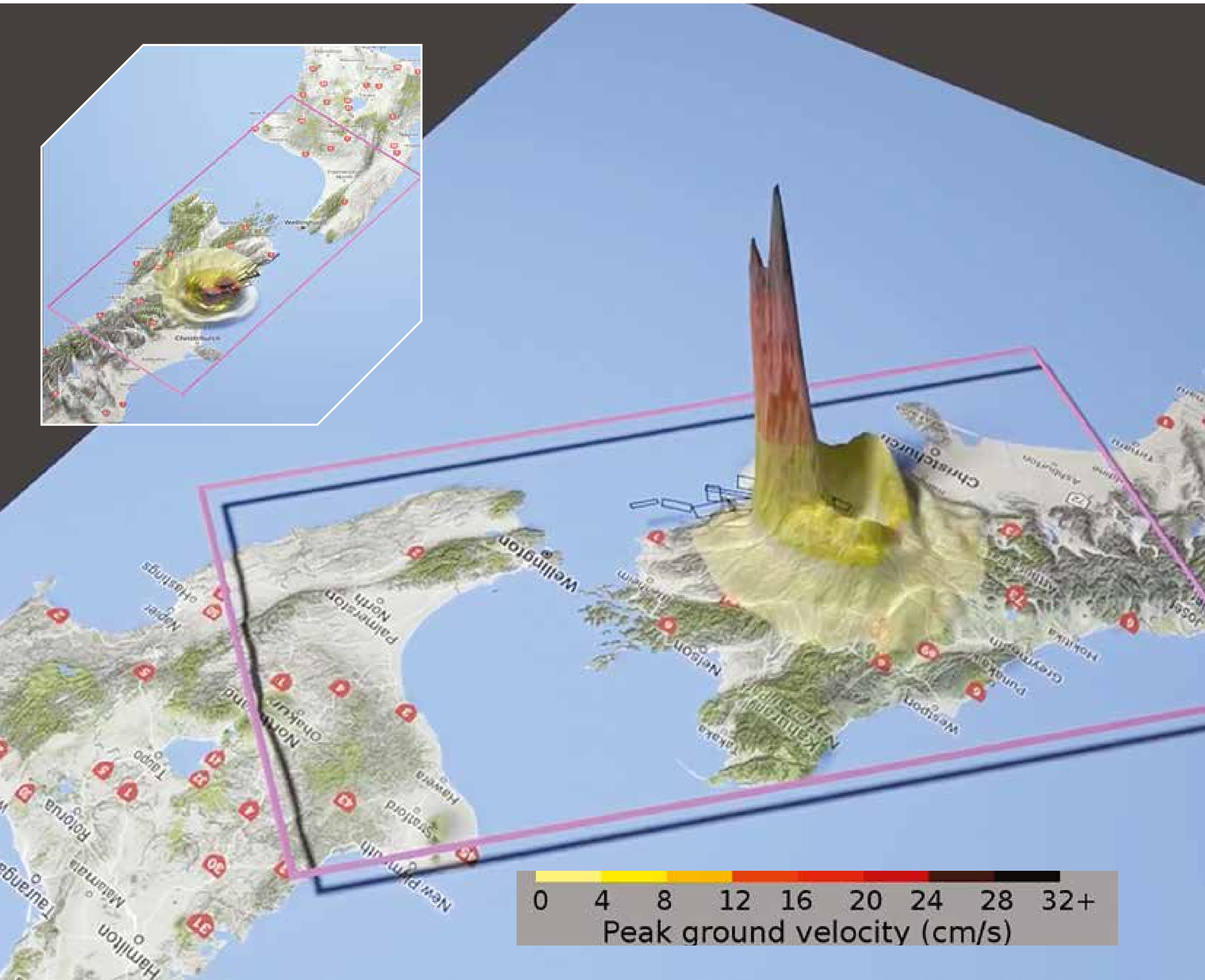
1x86" 3D immersive portable display.

1 stereoscopic 3D zSpace.

6 Microsoft Hololenses.

4 HTC vives.

A screen shot from the earthquake of 2016 Kaikoura.



https://auckland.figshare.com

Research data management

76K approximate views for institutional data publishing in Figshare since it started.

13K hits and downloads of UoA materials via Figshare.

53% of all identified existing datasets are assessed for migration to appropriate storage.

58 in-depth research data consultations including deep learning and artificial intelligence projects.

55 research data storage projects are projectised and supported to provide collaboration, sharing and discovery of data.

12 instrument research data repositories are now properly managed.


Sir Peter Gluckman,
Government Chief Science Advisor

Digital skills development and community building

22 departments / research groups with over **250** researchers and students have been trained to use visualisation facility and tools.

100+ individual consultations or training in using RVM and Nectar computing platforms.
80+ research computing workshops (Doctoral Skills Programme and HPC).

Visualisation & analytics

Research computing platforms

Digital skills development / community building

Research data management

150+ attendees in ResBaz.
26 workshops with **400+** course participants in Winter Boot Camp.
45 sessions of weekly HackyHour.
100+ Software Carpentry attendees.

29 workshops with over **350** participants university-wide have been trained in research data management and planning.

Approximately **1,600** counts of participation in CeR's digital skills development and community building throughout the year.

Collaboration with researchers: special projects

The following are some examples of research collaboration we have undertaken this year.

Index of Multiple Deprivation (IMD)

Create a dynamic, accessible, interactive web interface to enable users to extract and combine data; aim to improve understanding of the impact of deprivation and health inequalities in NZ.
- *Epidemiology and Biostatistics*.

Systems Obesity Prevention under 5's

Assist the development of the group model building, quantitative systems models and workshops support, and provide feedback to the intervention of childhood obesity. - *Population Health*.

eDNA Virtual Hub

Produce a working example of a visualisation system used to geographically display eDNA data from different regions of New Zealand. - *School of Biological Sciences*.

The Writer's Diet

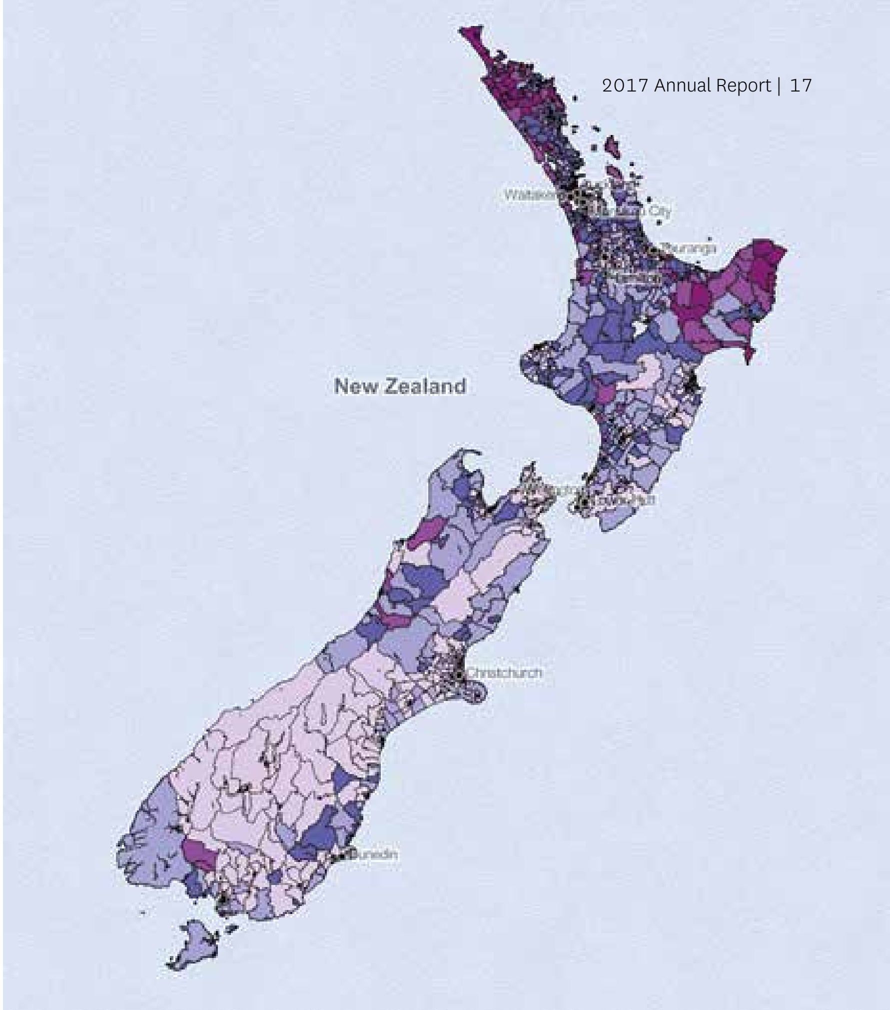
Develop and test a prototype Word add-in for the Writer's Diet website to enable writers to diagnose their writing for flab. - *Centre for Learning and Research in Higher Education (CLeaR)*.

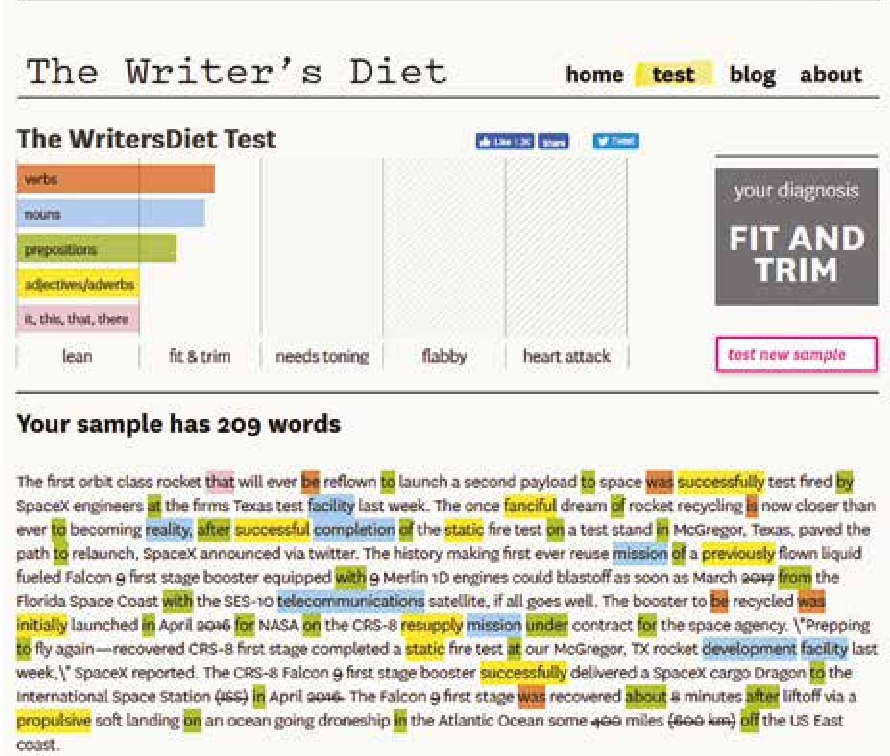
Ad Health

Measure exposure of adolescents to unhealthy food advertising through Facebook. *https://adhealth.azurewebsites.net/info/* - *Population Health*.

INFORMAS

Develop an interactive web presence for International Network for Food and Obesity / non-communicable Diseases Research, Monitoring and Action Support (INFORMAS). - *Population Health*.





Research IT skills programme

Examples of programme and events that the CeR either organised or provided in 2017 include the following:

1. Software Carpentry: Includes program design, version control, data management and ask automation.
2. Introduction to Tidy Data and Data Cleaning with Open Refine: A free tool for working with messy data.
3. Mobile App Development for Research: Covers both Android and iPhone App development fundamentals.
4. Introduction and Advanced Message Passing Interface (MPI): Take parallel code from desktop to an HPC system; tuning and performance analysis.
5. LaTeX Intro and Advance: Enables high-quality typesetting of complex mathematical formulas, multi-lingual symbols, creation of graphics and wider flexibility in editing text.
6. Maximise your Academic Success - Guidance for Publishing with Springer.
7. Open Access Licences: Making choices about publication and manage copyright.
8. Introduction to HPC: How to access and use NZ eScience Infrastructure (NeSI) High Performance Computing cluster.
9. Presenting Research with Academic Posters: Practical tips for designing a poster for competition and conference.
10. Introduction to Deep Learning: Keras - a powerful tool to wrap the efficient numerical computation libraries that allows you to define and train neural network models in a few short lines of code.
11. Introduction to Open MP: Provide an API for parallelising your research on your desktop.
12. Social Media for Communicating Research and Building Impact.
13. Digital Publishing on the Semantic Web: Understand semantic web technologies and the copyright issues on open web.
14. Prestigious or Predatory Publishers? How to navigate the shark-infested waters of modern publishing.
15. Intermediate R: Data Manipulation and Visualisation: Centre on the conceptual underpinnings and implementation of the dplyr and ggplot2 packages which form the core of the tidyverse suite of tools.
16. Commercialisation and Open Science: Can commercialisation and open science be undertaken at the same time or are they mutually exclusive?
17. Research Data Management: Develop strategies for capturing and organising research data, explore options around publishing, sharing and reusing data, and have an opportunity to draft a Data Management Plan.
18. Online Presence: Develop and manage your digital / online presence, and share ideas through social media.
19. Free Up your Desktop by using Virtual Machines for Research: Learn to offload long-running simulations to a different computer, how to request access and work with a virtual machines.
20. Referencing for Research: Explore reference management tools and resources available at the University.
21. Three Things to Know Before Signing Away your Rights (before entering into an agreement with a publisher).
22. Intellectual Property Rights: Introduction to IP, the patent process, and touch on related collaborative research and commercial issues.
23. Introduction to User Experience Design: Introduce what are user experience research and design, and how to improve online learning experience.
24. Research Project Management: Introduce some commonly used processes for project management in an agile manner.

CASE STUDIES

Research computing platforms

1. New analytics tools for workload planning for the 2018 New Zealand Census
2. Skin-omics: exploring the volatile organic compounds on human skin
3. Mobile Click Fraud Attack (MCFA)
4. Distributed and cloud-based control at field-level for systems interacting with soft bodies
5. Improving arrival time predictions for vehicles in a public transport network
6. Growing Up in New Zealand
7. Centre for eResearch machine learning service

Research data management

8. Taking a ‘Big Data’ approach to find new clinical-omic associations in cancer
9. MFT-ICR Mass Spectrometry data management and analysis workflow
10. Making stroke recovery prediction tools freely available
11. Improving diagnosis for schistosomiasis by using the ‘metabolic footprint’ of urine samples from an animal model of Schistosoma infection to identify possible biomarkers
12. Giving Pacific research greater reach

Visualisation & analytics and collaboration with researchers

13. Disposition of Microsoft HoloLenses for a Pop-Up Reality Shop to demonstrate the progress of a research project
14. ALTER: Between human and nonhuman - a VR art exhibition
15. Wandering around the molecular landscape: embracing virtual reality as a research showcasing outreach and teaching tool
16. Visualising the New Zealand Index of Multiple Deprivation (IMD)

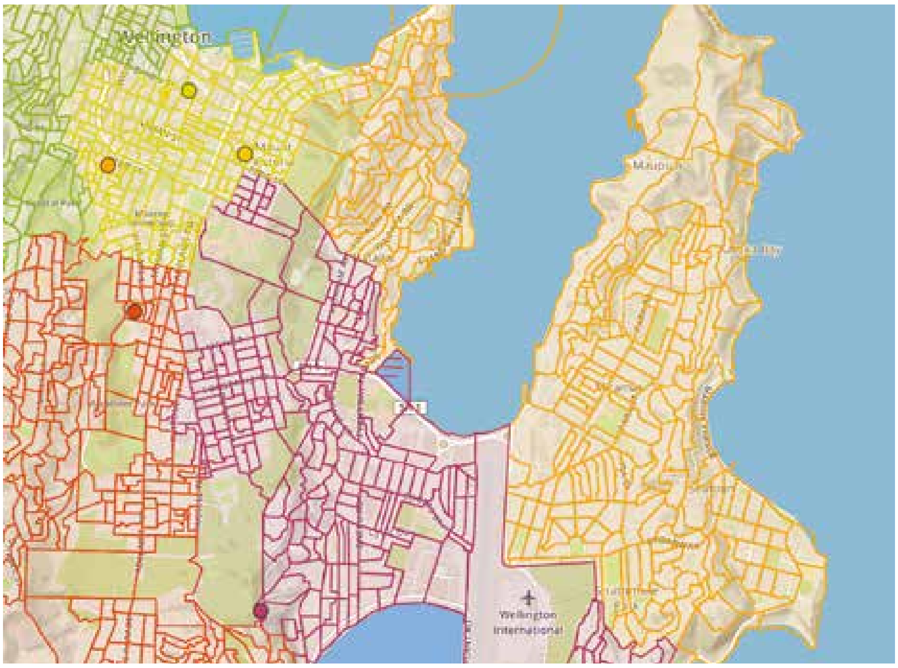
NeSI High Performance Computing platform

17. Modernising models to help diagnose or treat disease and injury
18. Aerodynamics modelling paves the way for improved yacht designs
19. Shedding new light on dark matter
20. Using GPUs to expand our understanding of the Solar System

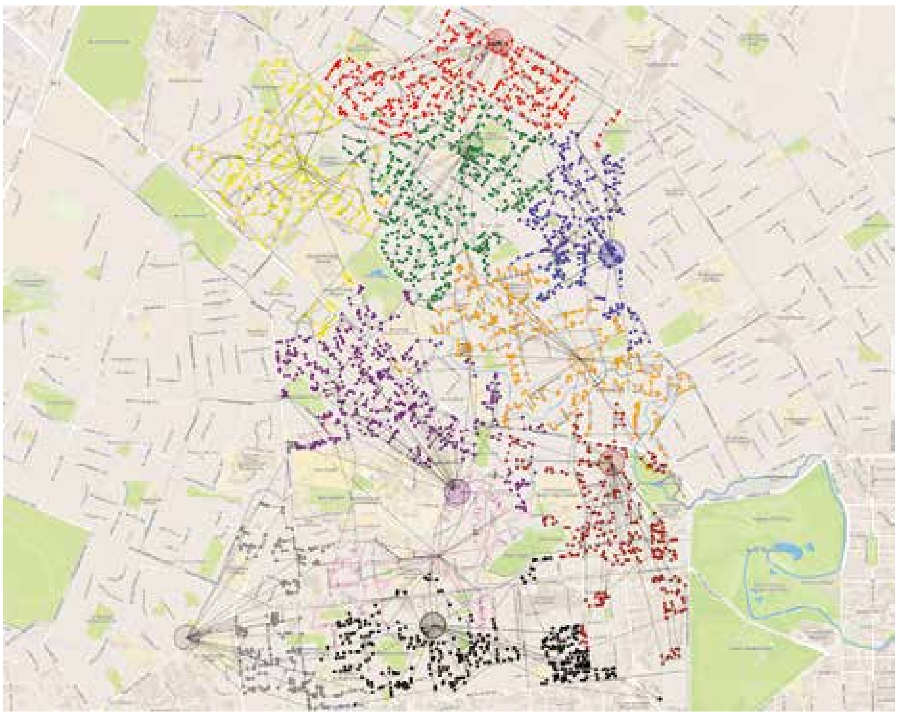
Case study 1 | Research computing platforms

NEW ANALYTICS TOOLS FOR WORKLOAD PLANNING FOR THE 2018 NEW ZEALAND CENSUS

Associate Professor Andrew Mason, Engineering Science; Geoff Leyland, Incremental Ltd.; Chris Hodgins, John Crequer, Craig Lange, Hayley Gargiulo, Philippa Sowman and Chris Deake, Statistics New Zealand.



Meshblock-based work in Wellington



Dwelling-based work in Christchurch

Background

New Zealand's next census will take place in March 2018. The census is an official count of all the people and dwellings in New Zealand, and will involve around two-and-a-half-thousand field officers and cover almost two million households. This is a large field operation with officers observing or visiting every dwelling in the country at least once. The operation confirms the locations of known dwellings and identifies new ones to ensure that Statistics New Zealand's address list is up-to-date.

New Zealand is divided into around 53,000 areas called "meshblocks". Each meshblock contains roughly the same number of dwellings. The census involves a field officer visiting every dwelling in each meshblock. Additional work may include following up with non-respondents after the census day.

For the first time in New Zealand, the census will use the internet as the primary mode of response. The operation will cover the same geographical area as in previous censuses, but with fewer field officers working across the country.

Fieldwork has already begun for the 2018 Census. Over 80% of dwellings were visited between June and August 2017 by a team of over two hundred field officers as part of an operation called "Address Canvassing". By the time the census gets underway in March 2018, every dwelling in the country will have been observed or visited by one of the teams of field officers, building the most up-to-date

Meshblocks in the Auckland Region

record of where people in New Zealand live. In the weeks following the census, households that have not responded will be visited by another team to encourage response in order to ensure the quality of the census findings.

The workload planning tools

Proper planning is critical for efficient census fieldwork. A team consisting of specialists from Statistics New Zealand, Incremental Ltd., and the University of Auckland is building fieldwork planning tools that enable the right officers to go to the right places at the right time. These planning tools form part of a case management system at Statistics NZ which is integrated with mobile devices being used by field officers. The planning tools communicate with these devices to receive information on work that has been completed each day, and to deliver the next day's plans to field officers. These plans are built using new optimization algorithms developed for this project. The algorithms are dynamic in that new plans are developed each night that respond to the actual work completion rates across field officers.

Planning largely relies on geographical information: the boundaries of the meshblocks; the underlying road network and how it interacts with the meshblocks; the locations of dwellings within the meshblocks; boundaries of area units, and population centres. These are all used to define what needs to happen as part of the census.

Application and methodology

Our tools introduce a range of improvements over previous census operations. One of the key innovations in our census tools is the use of a road-network model, and its integration with the meshblock and dwelling models, in order to take actual road distances into account when planning the work. The need to explicitly visit every road in NZ means that the OpenStreetMap data we are using is being tested in ways that are almost certainly new. We have made multiple improvements to the OpenStreetMap maps to fix data errors we have found.

Previous censuses relied on complex GIS processing to determine pairs of meshblocks that share a common boundary, but should not be grouped together because they were separated by geographic features such as rivers and mountain ranges. This processing has been improved and simplified by using new algorithms based on our road network. If a river or mountain range is impassable, there will be no road joining a meshblock pair. If there is a road, then the river or range is passable, and so the meshblocks should be the candidates for grouping together.

New tools require careful testing. All our algorithms have been embedded in detailed simulations, which are tightly coupled with a suite of new visualisation tools. These visualisations have proven invaluable in evaluating and improving the approaches we have developed, and in giving the team confidence that our new algorithms and tools will help lead to a successful 2018 Census Operation.

Case study 2 | Research computing platforms

SKIN-OMICS: EXPLORING THE VOLATILE ORGANIC COMPOUNDS ON HUMAN SKIN

Nita McKenzie, Student, Liggins Institute; Dr Chris Pook, Research Technician, Auckland University of Technology; Dr Elizabeth McKenzie, Postdoctoral Fellow, Liggins Institute (Project leader).

Figures above: Skin sampling, storage and analysis by student Nita McKenzie and Research Technician Dr Chris Pook.

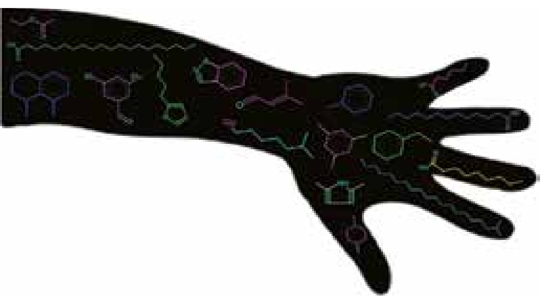
Overview

Skin is the largest organ of the human body. Its importance in health has only just begun to be investigated. There are hundreds of compounds on the skin surface, and they have the potential to answer many questions. These compounds can be used for diagnosis, monitoring pollutant exposure, and for understanding the skin microbiome. The ease with which skin can be sampled and analysed makes it an attractive choice for researchers. However, the major drawback of skin sampling is sample stability. Volatile samples quickly evaporate if left at room temperature, or stored in inappropriate containers. A method for stable bio-banking of human skin surface samples could open a new world of possibilities for the study of human health and the environment.

Effect of storage on the metabolite profile of human skin surface samples

The project tested the feasibility of long term storage of skin samples using miniature silicone-coated glass magnetic laboratory stir bars. Samples were collected by rolling the stir bars on the forearm. Half of the stir bars were analysed immediately using thermal desorption gas chromatography-mass spectrometry (GC-MS). The other half were stored for one month at -80°C in glass vials, then were analysed in the same manner.

We discovered 330 organic compounds including fatty acids, alkanes, alkenes, azoles, amines, sulfides, aldehydes, acetates, ketones, and terpenes. We also found many compounds, including plasticisers that originated from environmental pollution and personal care products. The majority (86%) of the compounds we discovered remained stable after one month storage, showing that the integrity of the metabolic profile was not significantly compromised by storage. Storage of skin samples for long periods therefore appears feasible under the right conditions.



This was the first project that trialled untargeted discovery metabolomics data extraction from start to finish on the metabolomics virtual machine, set up for metabolomics data extraction by the Centre for eResearch.

Previously, analysis of high-dimensional omics data of this type involved installing specialist software on a computer for the student, liaising with the ITS to find a computer powerful enough to process the data, or the student had to book to use the single computer available at the metabolomics instrument laboratory. The virtual machine enabled the summer student to access the data processing environment she needed from a standard university's computer. This represented a significant saving in setup cost and time, which allowed us to focus more time on the science.

The future of the project

The purpose of this study was to assess the feasibility of using human skin samples for detection of biomarkers of acute illness from infant skin in neonatal intensive care. The results from this study were presented to the Australia and New Zealand Society for Mass Spectrometry Conference in July 2017 and the data is now being used to support grant applications for research on biomarkers of acute illness from infant skin. This collaboration with the Centre for eResearch not only advanced science in the field of human skin diagnostics, it also provided support for the ideas and research of early career researchers from both the Liggins Institute and Auckland University of Technology.

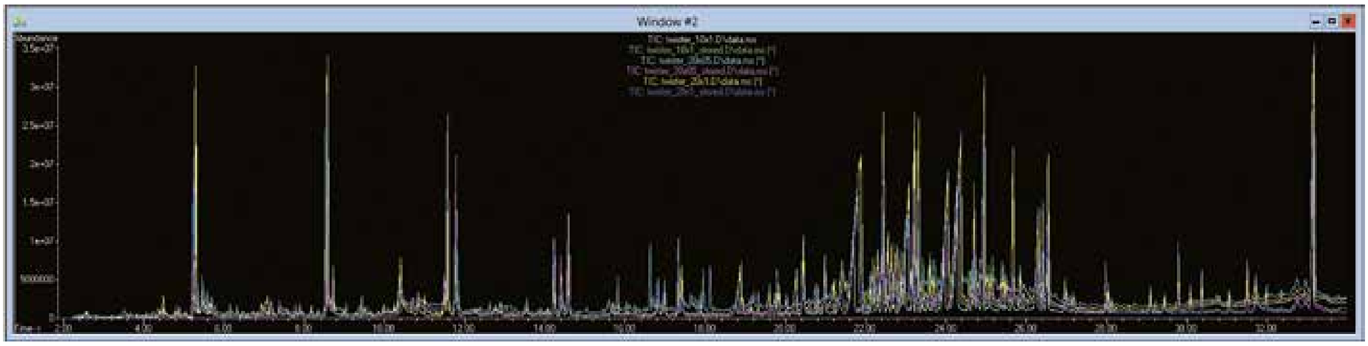


Figure 1. Overlaid chromatographic profiles showing compounds (peaks) detected from human skin samples.

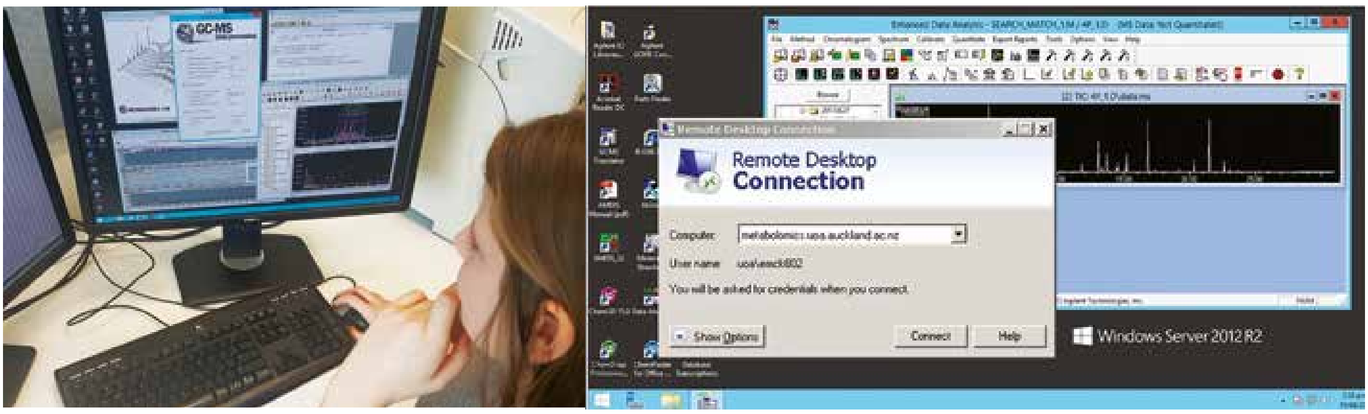


Figure 2. Data extraction and pre-processing by Nita McKenzie using the metabolomics virtual machine.

Case study 3 | Research computing platforms

MOBILE CLICK FRAUD ATTACK (MCFA)

Elliott Wen, Doctoral Candidate; Dr Gerald Weber, Senior Lecturer, Department of Computer Science.

A series threat: Mobile Click Fraud Attack

Nowadays, Mobile Click Fraud Attack (MCFA) is becoming a frequent topic within cyber security experts community. In an attack such as Click Fraud, malicious individuals repeatedly generate click events on mobile applications with the intention to increase revenues or exerting personal influence. Common examples include boosting product ratings or increasing the `like' number in social media pages. Shockingly, such attacks are known to have caused a substantial damage of US$16.7 billion on mobile application economy in 2017.

In general, most of these attacks are carried out by automating a massive number of physical devices. But, to purchase a large volume of devices would have incurred substantial costs. Therefore, a cheaper alternative to the physical devices is to use emulators. However, current existing emulators are inefficient and vastly blocked due to their immense resource demand and defective device signatures.

In this research, we propose a programme called Fraus - a cost efficient and scalable approach to conduct large-scale click fraud using device emulators.

Fraus maintains a low resource profile by circumventing graphics emulation and applying lazy-loading techniques on system components. In addition, Fraus provides a seemingly authentic device signature and disguises itself as a legitimate device by fully emulating the missing hardware components including WiFi interfaces and cellular modems. To facilitate the management of numerous emulator instances, Fraus also offers a distributed management system, which is scalable and fault-tolerant.

We evaluated the performance of Fraus by mocking attacks against the top 300 applications from the Google Play store. The results demonstrate that Fraus has high system stability and application compatibility.

Fraus also significantly reduces CPU usage and memory footprint by 90% and 60% respectively when compare it with the existing emulators.

By designing Fraus, we aim to raise public concerns about the simplicity of committing click fraud and to suggest countermeasures to mitigate such risks.

Support from the Centre for eResearch

This research involves extensive compiling tasks of the Android source codes. The whole process may take up to 3 hours in a normal desktop machine. The long-lasting waiting greatly impact the research progress.

Thanks to the staff at CeR who assisted our group gaining the access to the virtual machines with abundant CPU cores and terabytes of high-speed data storage, which significantly boost our research progress. Meanwhile, CeR created ready-to-go compiling environment for us, which is quite helpful as our research group can simply focus on our research targets without tuning the complicated settings of the virtual machines.

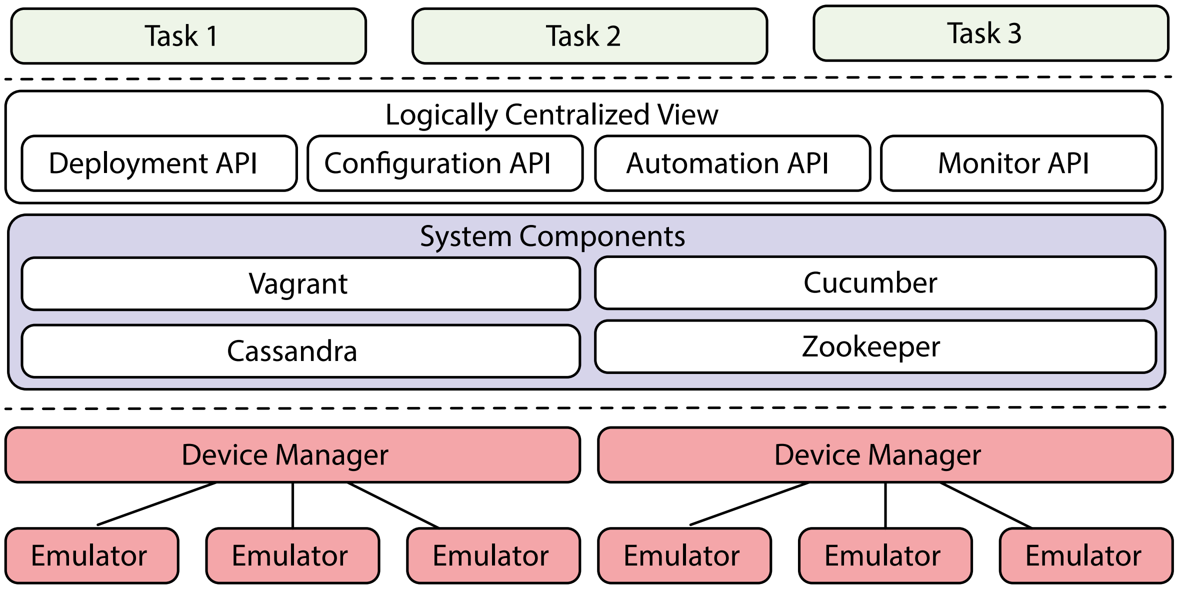


Figure 1. System Architecture of Fraus

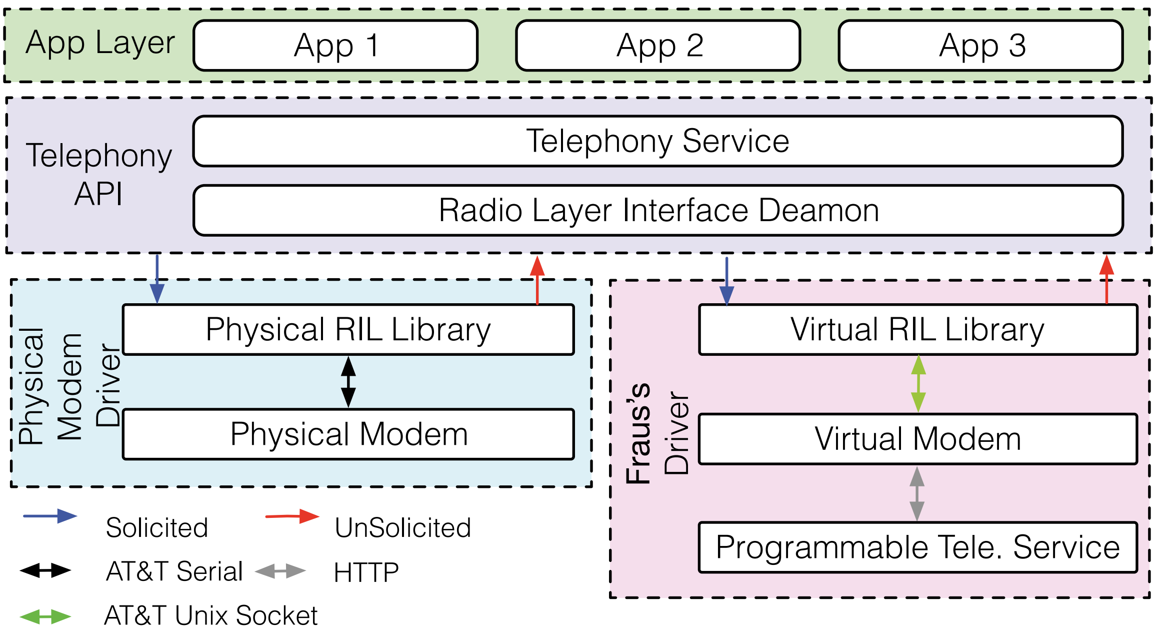


Figure 2. Overview of Android Telephony Architecture

Case study 4 | Research computing platforms

DISTRIBUTED AND CLOUD-BASED CONTROL AT FIELD-LEVEL FOR SYSTEMS INTERACTING WITH SOFT BODIES

David Tomzik, Doctoral Candidate, Department of Mechanical Engineering.

Addressing the need for increased flexibility in automation and research.

Overview

Systems that interact with soft bodies can face significant challenges. In the areas of agriculture (e.g. fruit harvesting, meat processing) and medical engineering, ill-defined, varying properties of objects and environment pose an obstacle for possible automation. A way to deal with the increased uncertainty is to use innovative sensors and actors. These sensors and actors produce varied and extensive amounts of data. A possible control system needs to provide increased computational resources and uncomplicated integration. Conventional control systems (often located next to the machine) lack these traits. We propose a new approach: individual components equipped with elementary computational resources, directly connected to an Ethernet-based network and are able to communicate with each other. Through the network, computationally expensive tasks can be off-loaded to components in the cloud. The components, both hardware and software, of such system and control can be easily combined and re-arranged.

Starting at the field level

To connect and communicate within a network, a component needs an interface and elementary computational resources. If it is not inherent, it needs to be added. Gateway devices that connect multiple sensors or components to Ethernet networks, so called IoT hubs, are already available. However, we wish to develop a similar solution by taking advantage of highly integrated and affordable Systems-on-a-Chip (SoC). Their price and size are small enough to bring connectivity directly to the individual component. The computational capability will also enable elementary control functionalities. Simple control loops can be realised by sensor and actor components communicating directly, eliminating the additional control instance in between.

Platform for a control network

Each hard- and software component can be seen as a node with clearly defined functionalities, outputs and inputs that are part of a control network. To allow easy and seamless communication, the nodes need to be implemented with a common interface. A good candidate is the industrial OPC UA standard. It is versatile enough to accommodate different platforms and use cases.

Additionally, the nodes will require internal flexibility to allow for changes of parameters and the control system arrangement. Crucial parts of this project will be the design of a node template and the implementation of methods to introduce nodes to each other and to set up the control connections. The platform will cover aspects of hardware, software and firmware alike.

Initial case study: retrofitting an 11-Year old robot into the Cloud

To assess the capabilities of possible hardware and to acquaint myself with network technology, I fitted an 11-year old industrial robot with a WiFi-enabled SoC, the ESP8266 from Espressif. The robot control system consists of a Programmable Logic Control (PLC) and a computer running Windows Embedded. For this operating system, I wrote a small program that takes the position information, available in the in the control system, and sends it to the ESP8266, which is connected via Serial Port. Even though the Serial Port has only limited bandwidth, it is easy to use, sufficient for this application and does not interfere with the existing setup of the robot within the laboratory.

The ESP8266 supports WiFi and is connected to the University network. It forwards the machine data, received at the Serial Port, to a MQTT message server running on a virtual machine. This message server can further distribute the data to various other components within the network. I chose to build a simple dashboard that visualizes the robot in 3D in real-time. The IIS webserver receives the machine data and hands it further down to the webclients. The website incorporates the three.js library that helps to visualize 3D content using HTML5. Possible additional features are the storage of the obtained machine data for further analysis and optimisation.

The use of a research virtual machine

The Centre for eResearch equipped me with a virtual machine that I use as a message and web server. They helped me to migrate my project from the Azure Cloud Platform into the University network and provided valuable advice on how to realise my plans. Additionally, I was invited as a beta tester for the Nectar platform. It enables researchers in Australia and New Zealand to easily deploy and connect virtual machines accessible from the internet - similar to solutions from commercial vendors like Microsoft or Google. I expect this service to be a very valuable tool during my research.

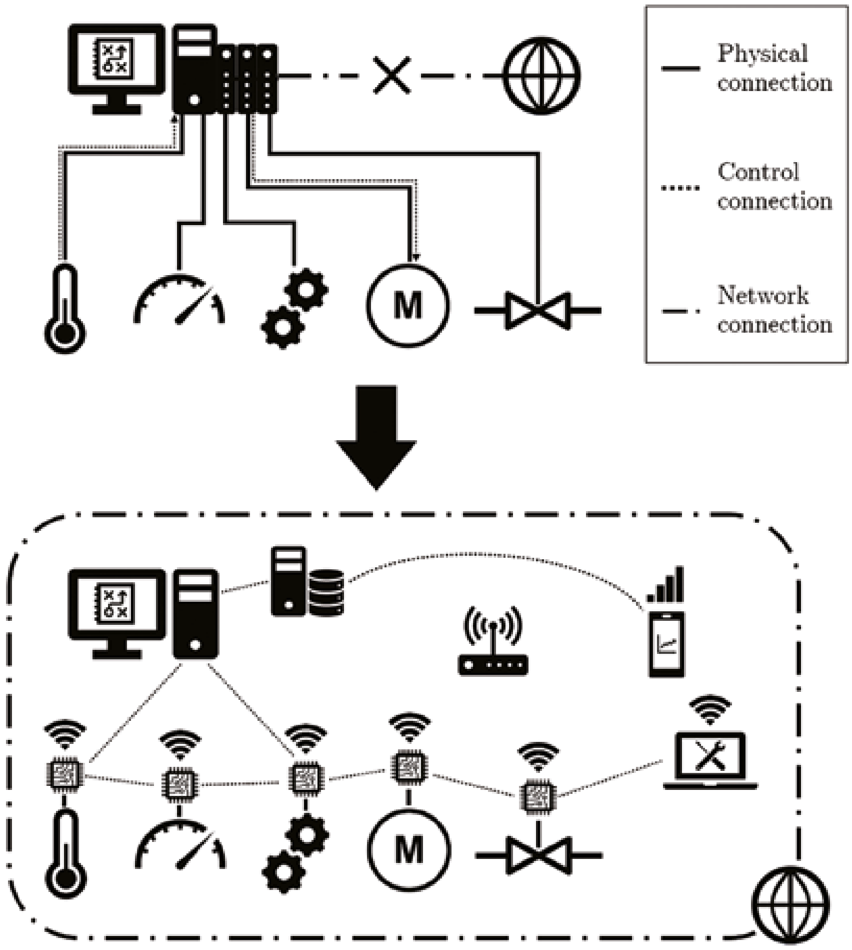


Figure 1. Transformation from central to a distributed control system based on an Ethernet network

Case study 5 | Research computing platforms

IMPROVING ARRIVAL TIME PREDICTIONS FOR VEHICLES IN A PUBLIC TRANSPORT NETWORK

Tom Elliott, Doctoral Candidate, Department of Statistics.

Predicting the arrival time of buses

In recent years, vehicle tracking technologies (GPS) have advanced along with communication technologies (internet, mobile apps) to the point where public transport users can see the real-time location of their buses on a map through their phones. However, the information on estimated arrival time (ETA) are still undependable. In place of statistically-founded predictive models that have been used historically, ETAs are often calculated based solely on the vehicle's current delay, as is the case here in Auckland. These predictions assume that the timetables are accurate (they often are not), and that the delays will not be affected by traffic conditions (they often are). Thus, the predictions are frequently inaccurate and ultimately unreliable.

Modelling network status using real-time data

We have been working on a predictive modeling framework that uses real-time vehicle locations to estimate network state (i.e., bus speeds along roads within the network), which is then used to generate ETAs that reflect current traffic conditions. This involves three stages:

1. a ***vehicle model*** that uses vehicle location data, freely available via Auckland Transport's public API, to estimate the vehicle's state: how far (in meters) it has traveled along the route, its speed, and travel times along the way.
2. a ***network state model*** that takes the estimated travel times of all vehicles, irrespective of route, and models the current average travel time of transit vehicles traveling along each road in the network; and
3. an ***arrival time prediction model*** that combines the estimated vehicle state from (1) with estimated travel times from (2) to predict the arrival time at all upcoming stops, taking into account uncertainty due to traffic lights and intermediate stops. (see Figure 1)

How the virtual machines make it possible

We are using a particle filter to implement the vehicle model, as it provides a robust, extensible framework; in particular, it handles multimodality (for example at bus stops), and allows us to use an intuitive likelihood function (based on the distance between two points on a map).

The downside of this particle filter is its high computational demand: at peak times, there are upwards of 700 buses operating, each modelled using 1000's of particles. These particles are also used in arrival time prediction, so predictions are required for every upcoming stops for all particles of each bus. In addition, this all needs to run efficiently in real-time to ensure ETAs are available to commuters as soon as possible, not 5 minutes after receiving the data.

A dedicated virtual machine is the ideal tool for the job. It can continually update ETAs in real-time as new data is received, and be supplied with enough resources so that the entire model runs in under 30 seconds. After contacting staff at the Centre for eResearch, they quickly assessed our needs and had a RVM up and running for our use. (see Figure 2)

Future work

Our goals for the future are to have the full model running on the VM in real-time, generating arrival time predictions for all future stops of active trips. We then hope to investigate using historical data, along with real-time, to estimate delays for trips that are yet to begin (currently, these are assumed on-time until the bus begins the trip). We will also look into journey planning applications that make use of our results and deciding on the "best route" to take: for example, which route is the fastest, or most likely to get me to work on-time.

Figure 1. Real-time vehicle locations are used to model travel times at are used to predict arrival time.

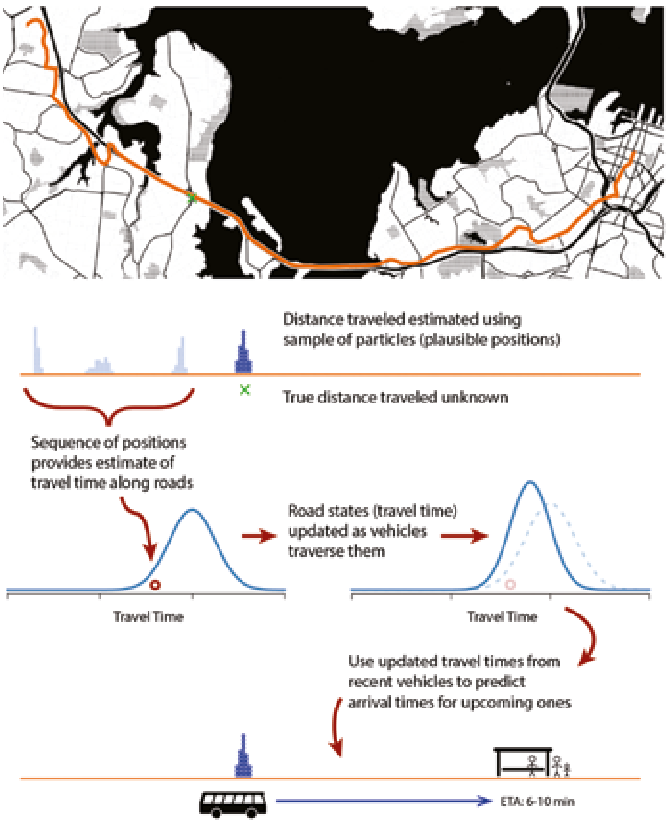
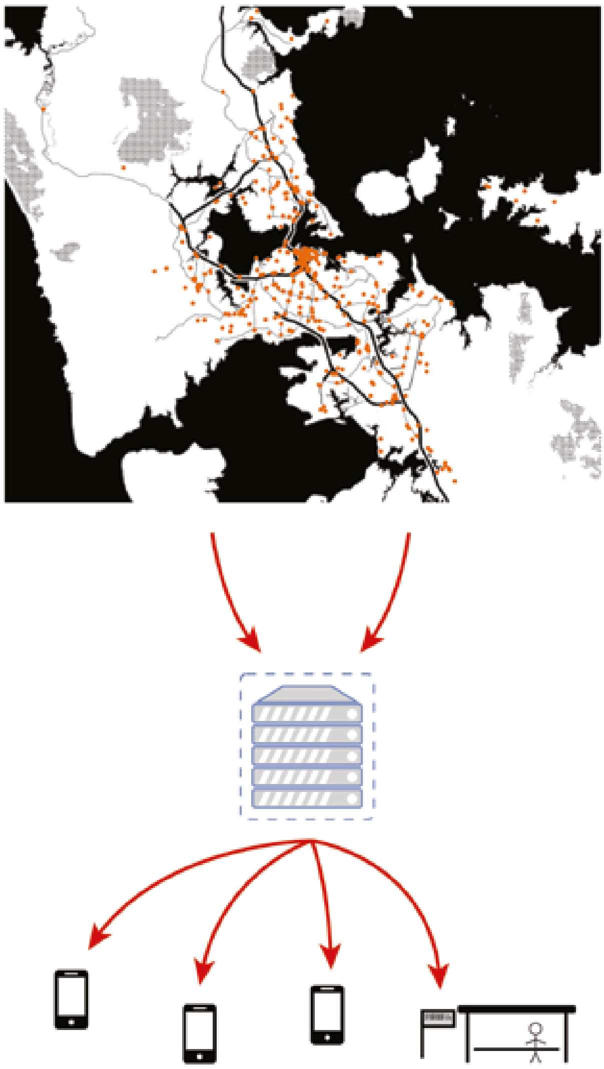


Figure 2. The VM processes real-time positions of all Auckland buses to provide ETAs to commuters.



Case study 6 | Research computing platforms

GROWING UP IN NEW ZEALAND

Peter Tricker, Data and System Manager, Growing up in New Zealand, UniServices.




Overview

The Growing Up in New Zealand (GUiNZ) longitudinal study was designed explicitly to understand what it is like to be a child growing up within the family, cultural, economic, societal and technological complexity of 21st century New Zealand. The study provides multi-disciplinary, scientifically robust, population-relevant research evidence which social policy and programme developers can use to decide what works best to improve social and developmental outcomes for children and families for children and families.

At the core of the GUiNZ study are the 7,000 children and their families who have been sharing their individual life stories with the study team since 2008. Specifically, the study recruited 6853 children before birth via their pregnant mothers (6823). The study includes 4401 partners of the children's mothers, the majority of whom are the children's biological father.

The cohort size and diversity are key strengths, ensuring that there is the capacity to deliver accurate information about all children growing up in New Zealand today. The main cohort is generalisable to all New Zealand births in terms of ethnic and socioeconomic diversity: one in four Māori, one in five Pacific, one in six Asian, two in three New Zealand European. Nearly half the children identify with more than one ethnic group, one in three have at least one parent born outside NZ, 40 percent of children and families live in the three highest deciles of area deprivation.

The commitment and contribution of each one of these families has been instrumental in establishing the study as a rich national resource. The study is set to continue until the children reach adulthood.

Drivers for change

By the time the children were five years old GUiNZ researchers had amassed more than 68 million pieces of data about the cohort. With each consecutive data collection, the eight year collection wave is currently underway, the volume and complexity of data and the analytical opportunities increase exponentially.

Data from each collection wave are cleaned, anonymised and prepared for release to external researchers. Through the early years of the study, these data sets were distributed to external users via a mix of dropbox-type technologies and by granting direct access to the GUiNZ network drives. It became obvious to the project team that not only was the system becoming increasingly difficult to maintain long-term, but that there was growing potential for security breaches of the data access protocol that governs principles of the study. Once in the public space, data would be difficult, if not impossible, to control and relied on the technical competence and goodwill of the user to maintain its security.

In 2016 GUiNZ was tasked with finding a means of increasing the utility of the data by making it more readily accessible to researchers outside the team and at the same time much more secure. The GUiNZ team decided it needed a datalab-type system that emulated systems currently

in use at Statistics New Zealand and The Ministry of Social Development.

System requirements

The requirements for the proposed system were quite straight forward:

1. Researchers must be able to access the system from outside of the University of Auckland network.
2. The system needed to be highly restrictive to protect both the anonymity of the study participants and the integrity of the researchers. Therefore, there should be no clipboard access, no print redirection, no internet connection and no LAN connection.
3. The system was to come pre-installed with software packages, such as SPSS, STATA and R, necessary to perform analyses on the data sets.

Solution

The Centre for eResearch (CeR) designed and built a remote desktop solution (RVM) that is accessible from any location in the world.

In addition, it is highly secure requiring a double login, and is provisioned with all the necessary statistical packages. Access is also platform agnostic, meaning the RVM can be accessed via Windows and MacOS platforms, for example.

The onboarding of users is controlled by GUiNZ and access to the RVM is only possible if CeR receives an authorising request from GUiNZ. As part of the value-added service, CeR also set up folders that can be shared between users on the same project so they can collaborate on their analyses.

CeR sat down with the GUiNZ team to review the project drivers and work out the details that would meet the team's requirements. Over the course of several consultative meetings, CeR team brought in experts and stakeholders, representing a range of fields, from networking to software licensing, from around the University. Throughout, they were thorough and professional in their approach and worked with GUiNZ to develop a highly secure, remote desktop RVM solution that would satisfy the project's requirements. Every step of the way Sean Matheny and his team kept GUiNZ informed of their progress on system development and raised flags where GUiNZ needed to make decisions. Now that the system is live, CeR provides ongoing support and maintenance for GUiNZ and its external users.

Future project aspirations

GUiNZ has recently initiated a data management improvement programme. The initial scoping exercise has highlighted several potential areas for improvement with regard to the internal data infrastructure and data management planning in general. CeR has indicated that it has the skills and expertise to assist GUiNZ in these areas and the team looks forward to engaging with CeR again through 2018.

Case study 7 | Research computing platforms

CENTRE FOR ERESEARCH MACHINE LEARNING SERVICE

Sina Masoud-Ansari, Research IT Specialist, Centre for eResearch.

Figure 1. Machine Learning as a Service (MLaaS) prototype website. Currently creates a Jupyter Notebook environment on request with all the tools required for popular deep learning workflows.

The initiative

There has been significant growth in interest around machine learning and its applications to research. This has largely been driven by advances in algorithms, particularly in neural networks where the use of Graphics Processing Units (GPUs) has reduced the time to train complex or 'deep' neural networks by orders of magnitude. With the rise of 'big data', researchers are increasingly finding opportunities to apply these methods to deliver new insights.

Machine learning is often useful when explicit formal models are difficult to develop, for example, a model to identify human faces. Whereas in the past, researchers would 'tell' the computer what features to look for, machine learning takes the opposite approach and attempts to 'learn' appropriate features based on successive attempts at guessing the correct answer and adjusting to reduce mistakes. Examples of research in this area that the Centre of eResearch is supporting include using machine learning to improve modelling of financial markets, developing intelligent question/answer systems for understanding text and developing predictors of heart disease based on the shape of coronary arteries.

In order to support these and future projects, CeR has initiated a Machine Learning Service to cater to a wide range of needs and experience levels at the University.

Powerful graphics cards are crucial to this type of research and can be difficult to get hold of in sufficient quantities. In addition, the software stack required to run these workflows can be barrier for researchers who are curious about using these tools but do not have the expertise or time to set up the required environment.

Self-service booking portal

To make it easier for researchers to experiment with machine learning and 'deep learning' in particular, we are developing a self-service booking portal where researchers can request time on computers with powerful hardware and have a fully capable Jupyter Notebook environment available to them with all the required tools (Figure 1). Our self-service notebooks are hosted in the University's

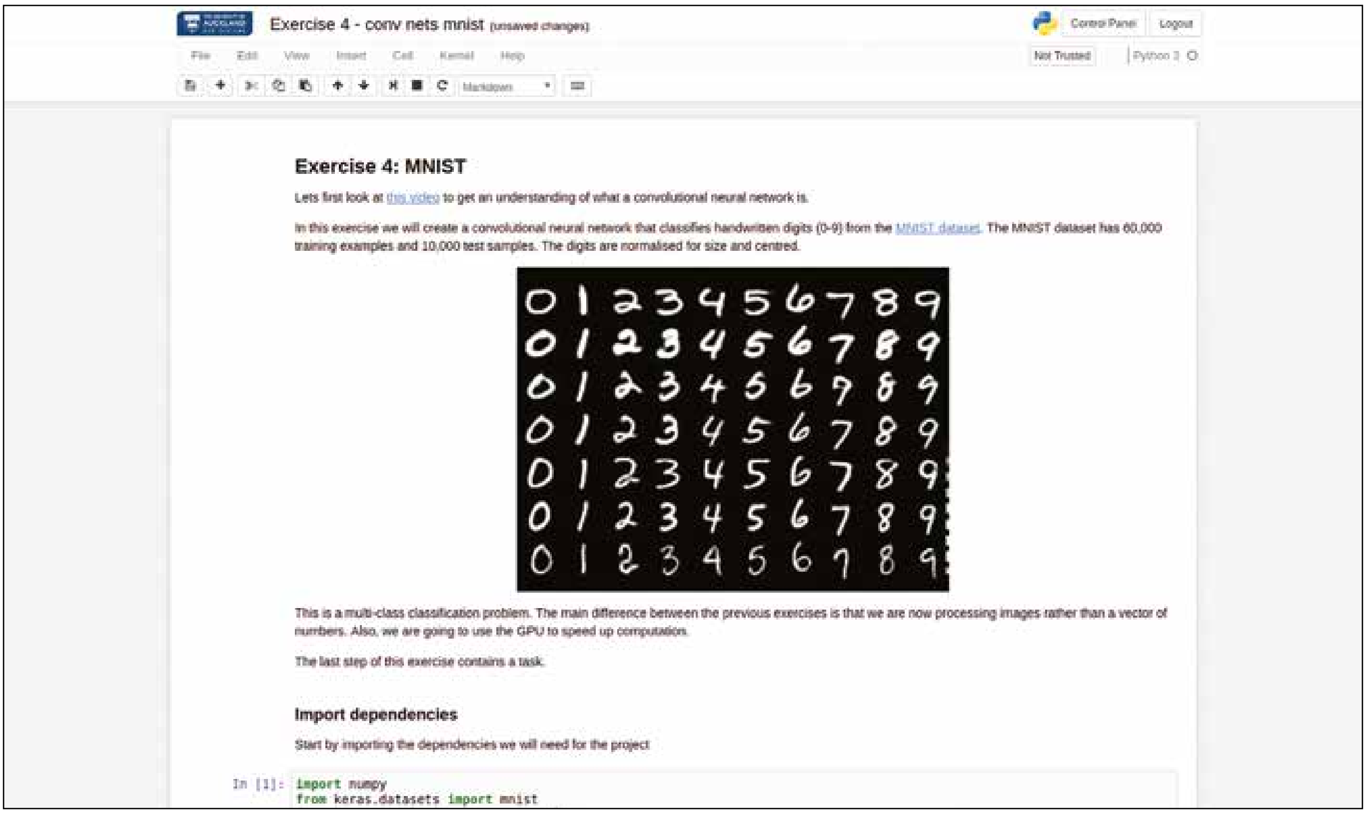


Figure 2. An example of a Jupyter Notebook running via the MLaaS prototype. The Notebook covers an introduction to neural networks for image recognition used as part of the Winter Bootcamp machine learning workshop in July this year.

Docker containers which allow each researcher to have an isolated computing environment on the shared system. This system was successfully trialled at the Winter Boot Camp in July where we ran an introduction to machine learning workshop with 40 participants working concurrently on the system (Figure 2). The benefit of using Docker containers in research is also in their reproducibility, where each instance has the same container operating system and libraries. This is helpful when researchers need to publish or replicate their workflows.

What more can CeR help with

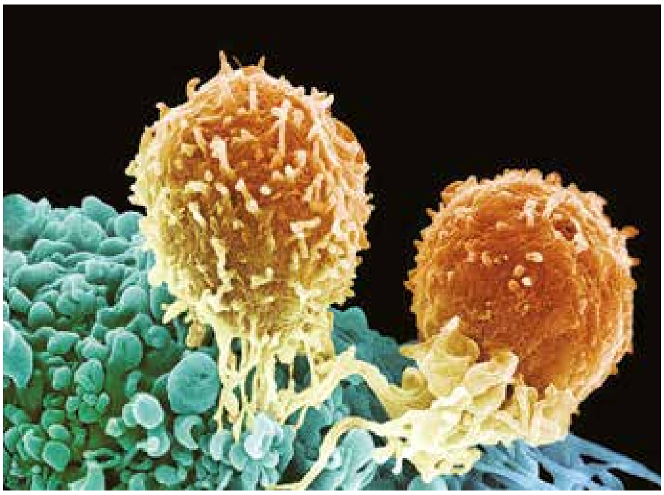
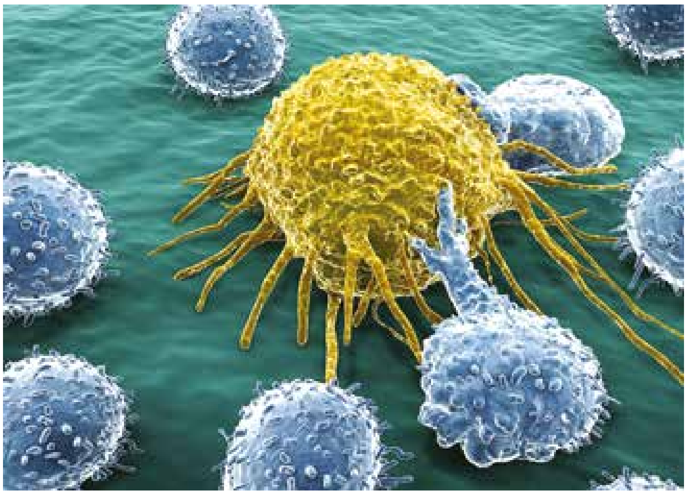
For more experienced researchers or for those who would like require more control over their computing environment, we offer remote access to a shared Linux server with powerful GPUs and fast SSD based storage for demanding workflows. Researchers can customise this environment to suit their needs, and the Centre provides additional support in installing libraries and additional software required for machine learning. For researchers looking to run large scale projects using multiple GPUs, we are planning to offer access to the Nectar research cloud where we have a large memory server with four GPUs. Our plan is to allow researchers to request virtual machines on the Nectar system where they can select the number of GPUs they would like to use. For most demanding workflows, we offer training and advice for researchers to help them make use of NeSI Pan cluster and its array of 20 GPUs ready machines.

The Centre also has the capacity to support researchers' development of machine learning approaches to solving research problems. For example, working with Dr. Susann Beier from the Faculty of Medical and Health Sciences, the Centre has started on a project exploring the applications of machine learning to improve segmentation of coronary arteries from medical images and exploring the potential of assessing coronary disease risk from the resultant 3D structures.

Case study 8 | Research data management and research computing platforms

TAKING A 'BIG DATA' APPROACH TO FIND NEW CLINICAL-OMIC ASSOCIATIONS IN CANCER

Nicholas Knowlton, Peter Tsai and Professor Cris Print, Molecular Medicine and Pathology.

Cancer cells. Source: https://www.technicalforweb.com/wp-content/uploads/2017/04

Clinical-omics

Cancer is the number two cause of mortality in the OECD behind heart disease. Up until the late 1990's, there was a concerted effort by drug companies to develop 'blockbuster therapies' for the treatment of cancer, i.e. cancer therapies developed with a one-size-fits-all approach.

This philosophy was upended by the improvements in computer hardware and software that now allow scientists to link multiple types of genomic data about a given cancer together to identify patient-specific therapies.

Finding genomic link

In this study, we are using new statistical methods to find genomic links between a patient's genomic immune signature and clinical information across multiple types of cancer. This will help identify the types of cancer patient who can benefit from the new 'Immune Checkpoint Inhibitor' drugs, which activate the patient's own immune system to kill cancer cells.

Currently, these drugs are only effective in a subset of patients – we are working towards identifying why this is and how to predict which patients will respond. We are accessing a 2.5 petabyte international genomic data archive of over 11 000 patients in an attempt to identify clinically useful but previously undiscovered information about cancers.

To identify potential antigens (altered proteins on the surface of cancer cells encoded by mutant genes), we download a patient's raw tumour gene sequence along with a normal tissue gene sequence and extract out the portion of the genome responsible for creating an immune response. Next, we use computer algorithms to compute the likely affinity between novel antigens presented by each patient's individual cancer and the individual patient's genomic immune signature. The initial results are promising, by simply counting the number of novel antigens predicted in 491 melanoma patients a survival difference is identified, where more mutations in individual tumours leads to a greater immune response by the patient's body against the tumour cells and longer patient survival (Figure 1).

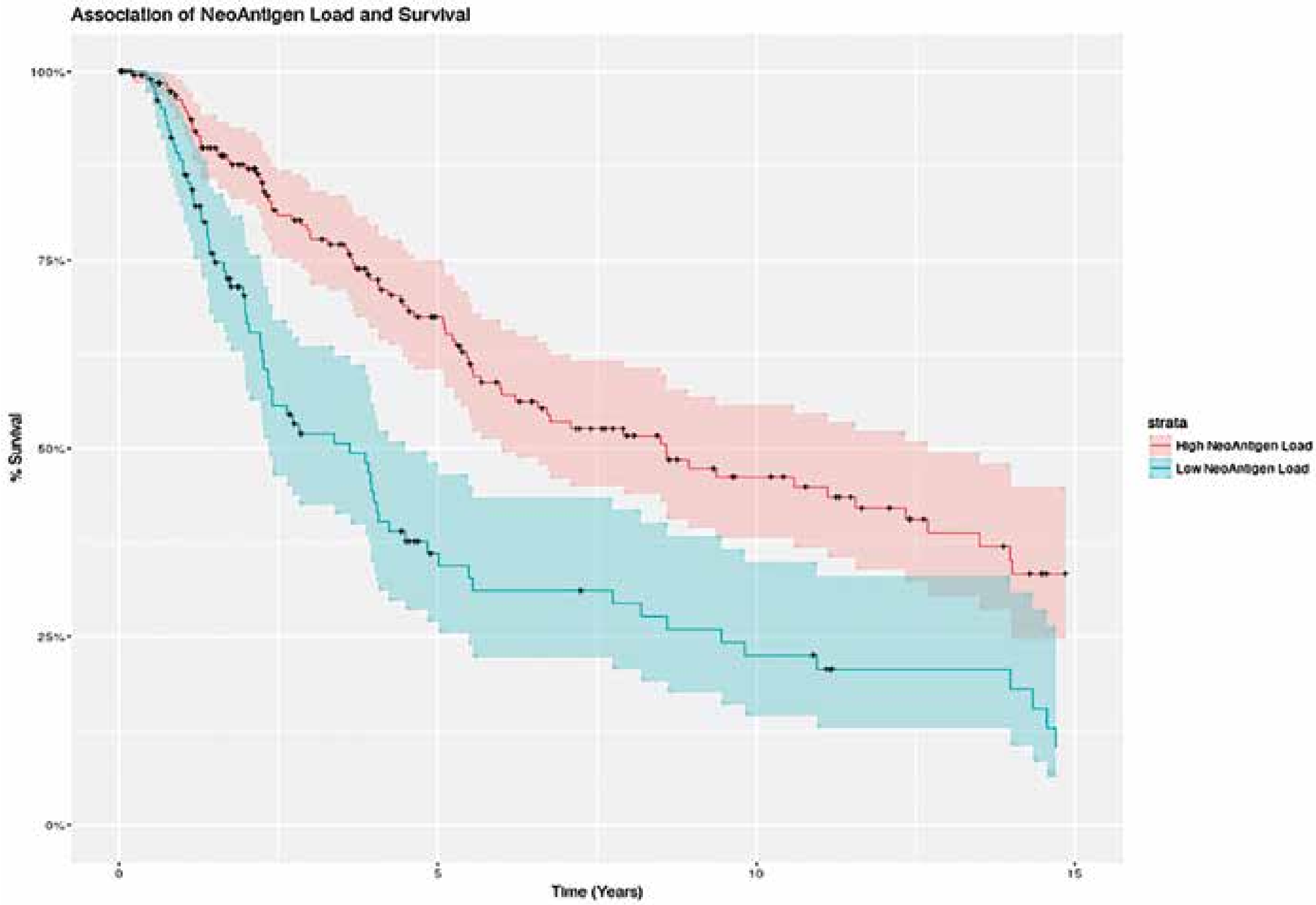


Figure 1. Association of NeoAntigen Load and Survival.

Data transfer, analysis and storage

We worked with the Centre for eResearch to develop and host the download of data from GDC secure servers, run analysis and access long term storage using a research virtual machine (RVM) running and research storage. This provided the essential compute, memory and network needs to process hundreds of terabytes of information down into digestible chunks. One advantage of the RVM 24/7 access to a dedicated 1 GB network link to download the necessary data. For example, processing a single tumour type it takes approximately 192 hours when including time to download and extract the relevant data. This is followed by an additional 60 hours to predict possible interactions of novel antigens in the tumours.

The future

The results of the initial analysis are promising and we will continue to process additional cancer types that will enable us to perform a pan-cancer analysis of computationally predicted novel antigens. This work is supported by the Maurice Wilkins Centre of Research Excellence.

Case study 9 | Research data management

MFT-ICR MASS SPECTROMETRY DATA MANAGEMENT AND ANALYSIS WORKFLOW

Dr Nicholas Demarais, Research Fellow, School of Biological Sciences.

Introduction

Mass spectrometry (MS) is an analytical technique that ionizes chemical species and sorts the ions based on their mass-to-charge ratio (m/z). In addition to the m/z value, the natural abundance of atomic isotopes provides a unique fingerprint for each ion. Fourier transform-ion cyclotron resonance mass spectrometry currently offers the highest resolution, mass accuracy, and sensitivity of all available mass analysers. This allows not only the assignment of an m/z value with high precision and accuracy, but also detection of the fingerprint of each ion even at low abundances. This is achieved by trapping ions in a static magnetic field. The ions travel in a circular motion and the sum of all frequencies of all ions is detected. Finally, this frequency sum is converted to a mass spectrum through a Fourier transform. A Bruker SolarixXR Fourier transform-ion cyclotron resonance (FT-ICR) mass spectrometer is located in the School of Biological Sciences mass spectrometry facility (Science Centre 303-B45). In addition to the advanced mass analyser, the utility of multiple ionization techniques make this instrument amenable to many different types of research in the fields of biology, chemistry, materials, health, and fundamental science. Existing users span both the Faculty of Science and Faculty of Medical and Health Sciences, and the capacity is expected to expand to other researchers within the University of Auckland and across New Zealand. A broad range of users creates issues due to their diverse locations, and the ultra-high resolution experiments can result in large and varied data sets. To overcome these issues, Dr Nicholas Demarais has worked closely with CeR to develop a robust data management plan, and an efficient analysis workflow that incorporates the use of RVM.

Data management plan

Although facility-run services are available for the SolarixXR FT-ICR mass spectrometer, collaborative research projects and researcher engagement are strongly encouraged. This necessitates the need for efficient dissemination of data to researchers across the University. A hierarchical structure centred on a project rather than a student/faculty member enables data from all types of analysis to be stored in one central location. This system reduces the redundancy of data sets stored across many PCs, identifies ownership, and provides access to all relevant persons. A read-only backup is accessible if the original data is ever needed. The workflow is presented in Figure 1.

Data analysis workflow

Once this data is managed and stored properly, the next major bottleneck is data analysis and access to appropriate software tools. The RVM is utilized to house all of the data analysis software packages in one location. With such a diverse user base in discipline and location, the RVM allows researchers to analyse, manipulate, and visualize their data from any computer on campus or externally through the VPN. This utility, combined with the data management system, creates an efficient workflow for mass spectral data analysis that reduces data reduction time, and increases the accessibility of the data (also refer to Figure 1).

Data visualization

In addition to increased accessibility and usability, the computing power of the RVM is vital for handling large data sets generated by this ultra-high resolution instrument. One example of such an experiment is matrix assisted laser desorption ionization-imaging mass spectrometry (MALDI-IMS). For this technique, a thin layer sample (e.g. sectioned human eye lens or mouse brain) is placed on a glass slide and coated with a matrix. A pulsed laser is used to generate ions, and a 2-D profile of the surface is collected by scanning the laser across the sample. One mass spectrum is collected at each point (10-200 um spacing), which results in 10,000–1,000,000 individual spectra. The individual mass spectra are then collated to produce a 2-D intensity profile (Figure 2). Typical data sets are 1–100 GB and the power of the RVM, combined with the centralized location of multiple software, allows users to manipulate and visualize their data efficiently.

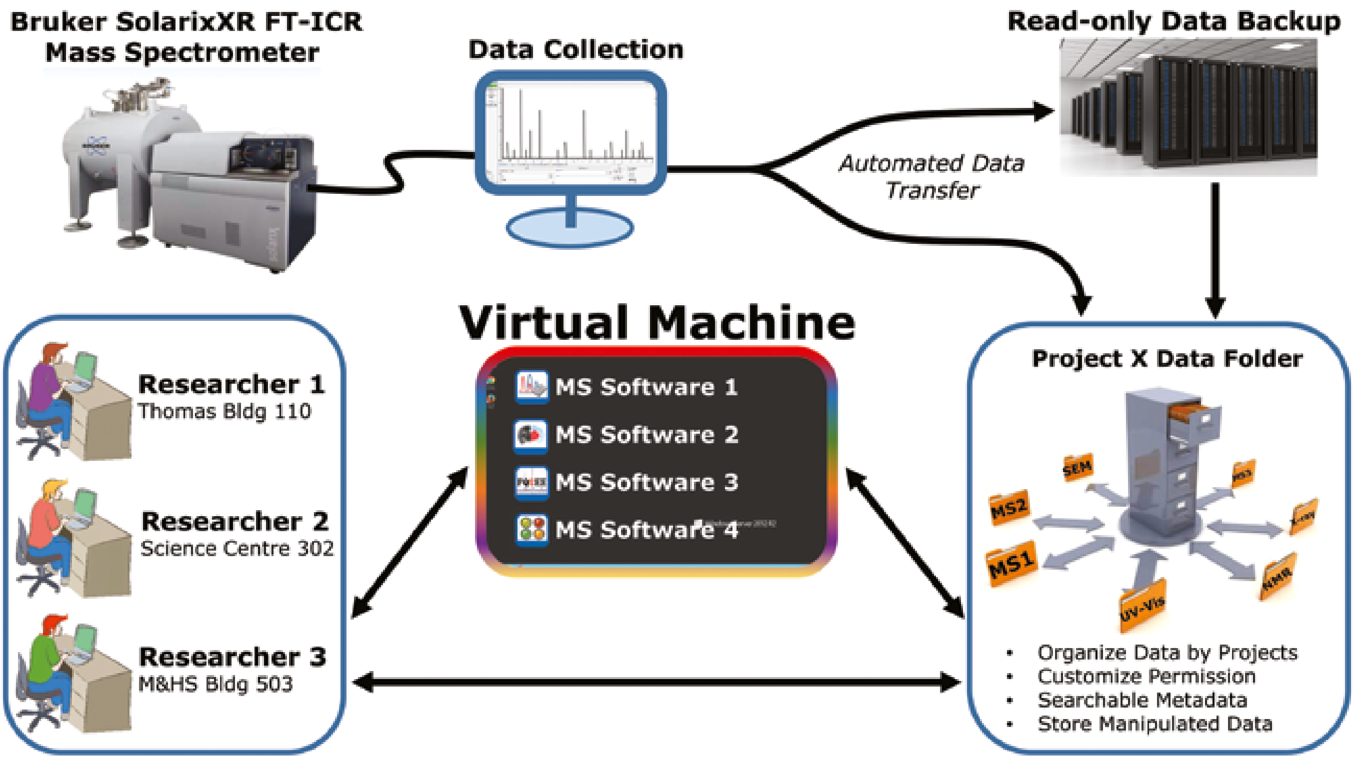


Figure 1. FT-ICR MS data management and analysis workflow. (From Top Left) Mass spectrometry experiments are carried out by researchers and data is collected on a local PC. This data is transferred to the relevant project folder and a backup is stored separately. Researchers with permission to access the project folder and/or mass spectrometry files can interact with the data directly or indirectly through the RVM. This can be done anywhere on campus or externally via access through the VPN. All relevant mass spectrometry software and visualization tools are available on the RVM.

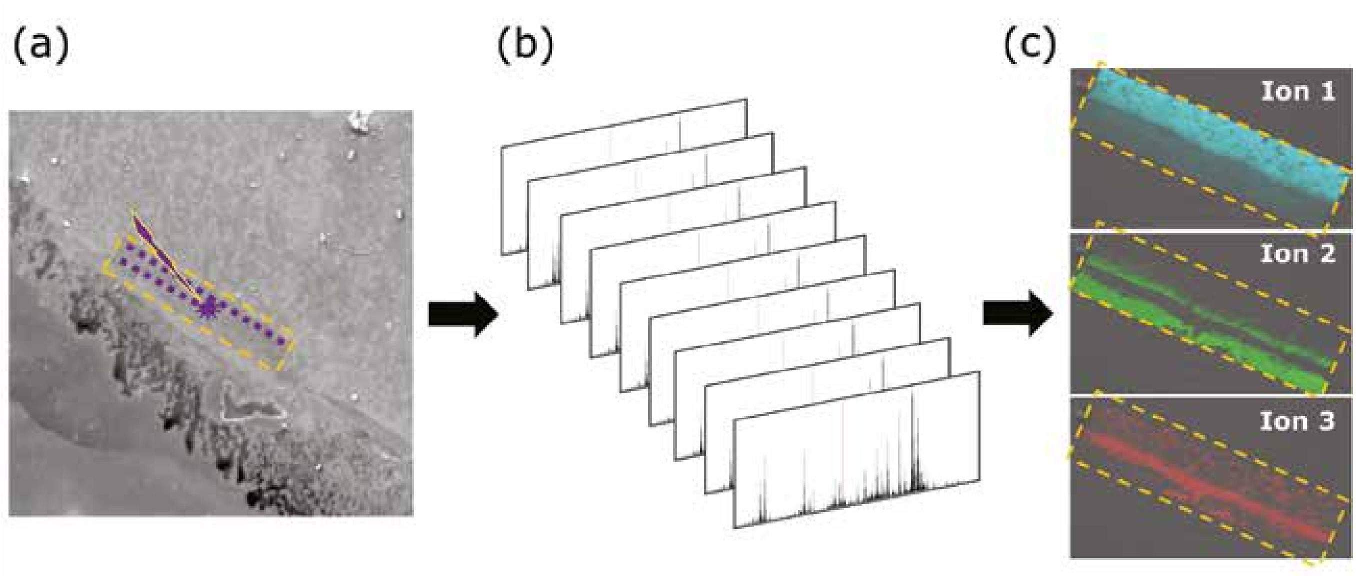


Figure 2. MALDI imaging mass spectrometry workflow. (a) A thin layer sample is placed on a glass slide and coated with a matrix. An area of interest is highlighted (dashed yellow), and scanned by a pulsed laser. (b) One mass spectrum is recorded for each spatial position. (c) The individual mass spectrum are collated to generate a 2-D image, and specific ions of interest are mapped onto the surface (brightness intensity). Three different special distributions are observed by plotting ions that correlate to three different molecular species.

Case study 10 | Research data management

MAKING STROKE RECOVERY PREDICTION TOOLS FREELY AVAILABLE

Associate Professor Cathy Stinear, Department of Medicine.

PRESTO platform and resource database

Stroke is a leading cause of adult disability worldwide. Most people who experience a stroke have weakness on one side of their body. The ability to live independently again after stroke depends largely on the recovery of strength and function on the affected side. Predicting how well someone will be able to use their hand and arm, and when they will be able to walk again, is an important but difficult task. Being able to predict recovery and outcomes allows better management of the patient's expectations, tailoring of rehabilitation goals, and more efficient use of time and resources. However, predictions based on clinical judgement and experience are inaccurate for many patients, especially those whose movement is more severely affected. Some of these patients make a good recovery and have good outcomes, while others don't improve at all, and it's impossible to tell them apart when relying on clinical impressions alone.

We've developed two clinical algorithms to address this challenge. The first is called PREP2, which predicts how well an individual patient will be able to use their hand and arm after stroke. The second is TWIST, which predicts when an individual will regain the ability to walk independently (Figure 1). We've combined algorithms in a project called Predict Stroke Outcomes (PRESTO).

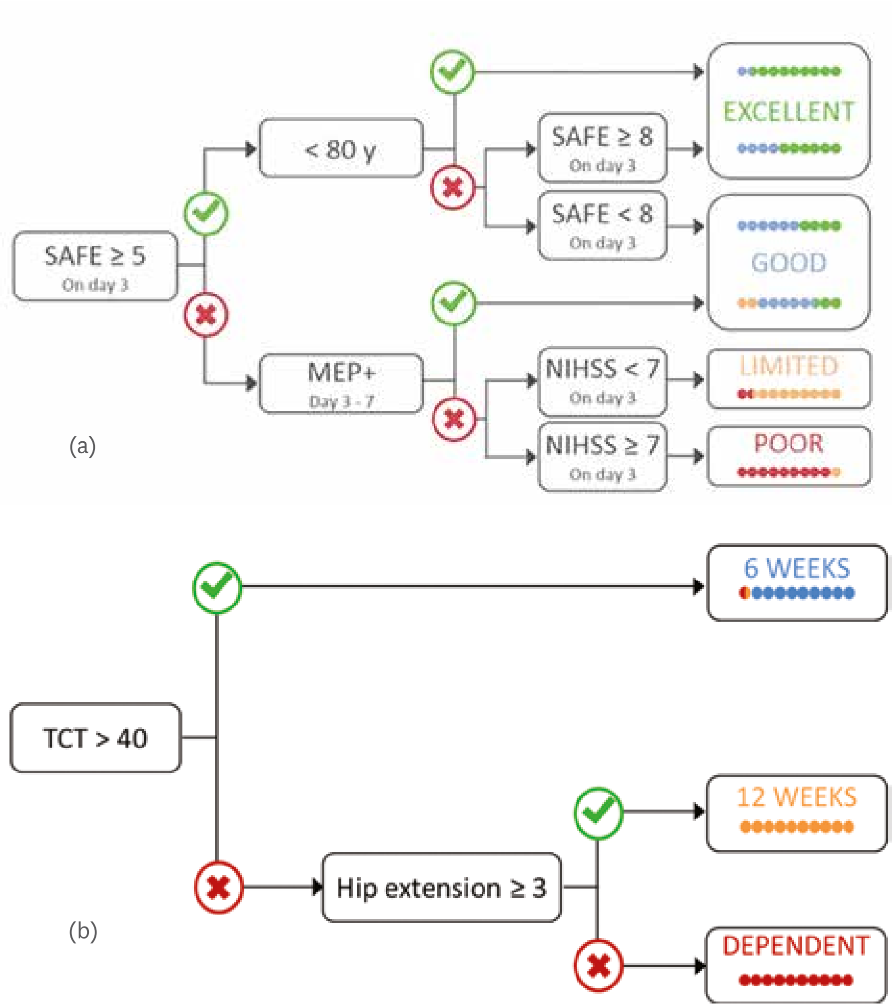


Figure 1. An overview of the two clinical algorithms
(a) PREP2 algorithm for predicting upper limb functional outcomes after stroke.
(b) TWIST algorithm for predicting the recovery of independent walking.

PRESTO aims to provide online information and resources to clinicians and researchers around the world to help with the implementation of both PREP2 and TWIST in a clinical or research context.

Using Figshare and Wikispace together for PRESTO

CeR has helped us to establish PRESTO with WordPress and the University of Auckland's Institutional Data Repository Figshare. WordPress provides us with a platform to host PRESTO and allows us to keep other groups up to date with our research, as well as explain how they can implement PREP and TWIST in a clinical or research environment (Figure 2).

Figshare provides a single place for us to upload all the resources related to PRESTO, and makes it easy for others to access and download these resources. Within Figshare we can easily track the impact of the resources in a real-time manner using altmetrics (alternative metrics) (Figure 3). Using WordPress and Figshare in combination makes it simple to go from finding and reading the information about PRESTO to downloading the resources.

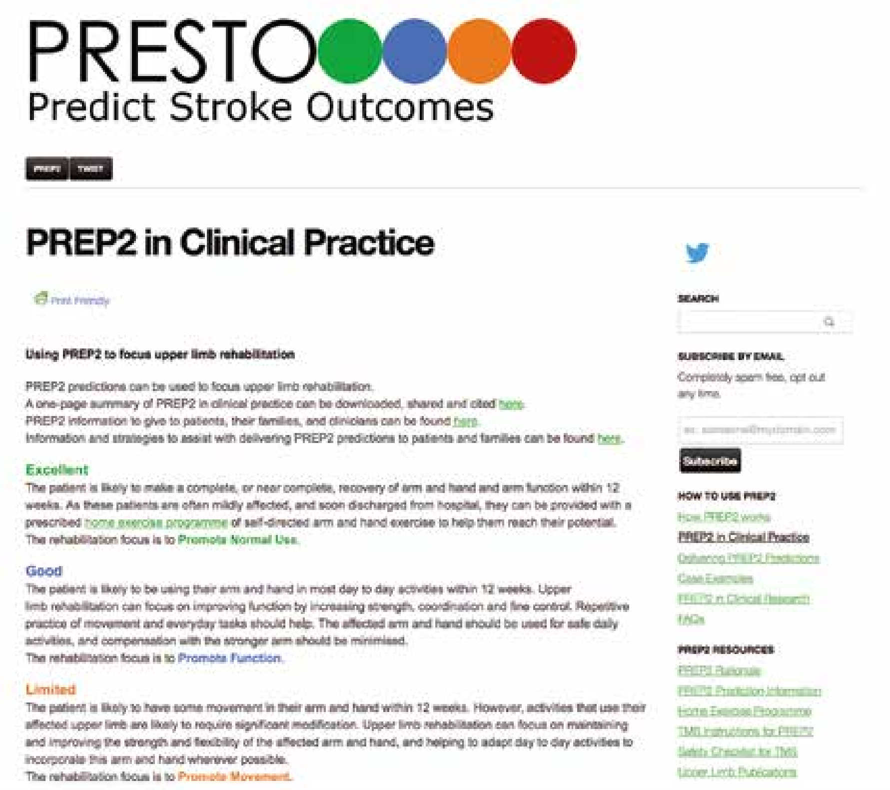


Figure 2. PRESTRO wordpress website. http://presto.auckland.ac.nz/

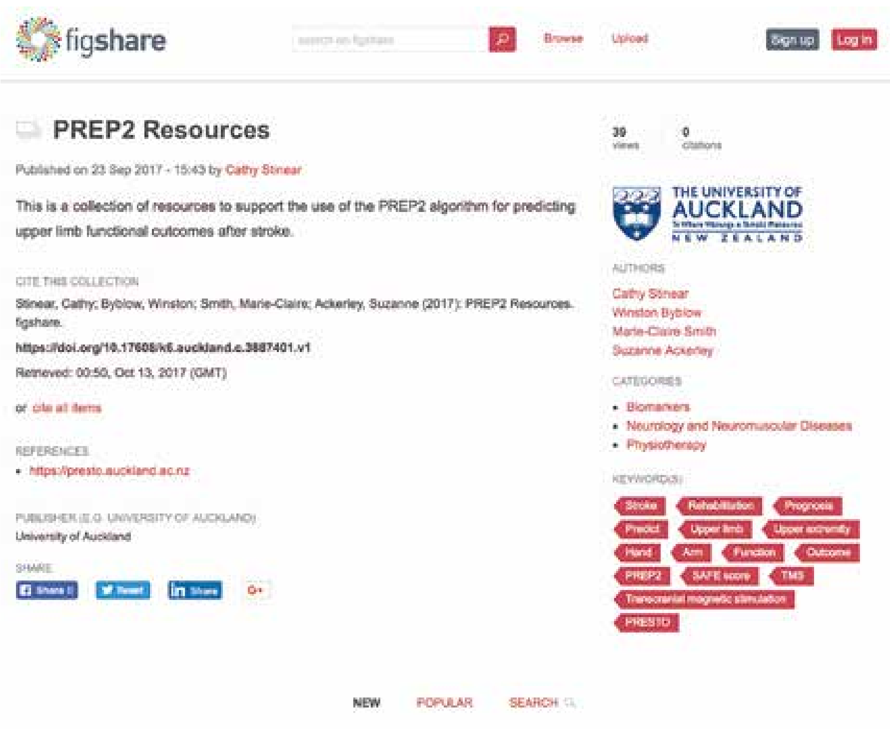


Figure 3. PREP 2 Resources in Figshare.
https://figshare.com/collections/PREP2_Resources/3887401

Case study 11 | Research data management

IMPROVING DIAGNOSIS FOR SCHISTOSOMIASIS BY USING THE 'METABOLIC FOOTPRINT' OF URINE SAMPLES FROM AN ANIMAL MODEL OF SCHISTOSOMA INFECTION TO IDENTIFY POSSIBLE BIOMARKERS

Rodrigo Loyo, Masters student, Brazil; Dr Augusto Barbosa, Supervisor, Biology Sciences; Dr Constança Simões Barbosa, Supervisor, Brazil; Dr Erica Zarate, Senior Technician GC-MS, Mass Spectrometry Centre.

Overview

Schistosomiasis is a neglected tropical disease caused by a trematode of the genus Schistosoma. The majority of human infections are attributable to the three species: S. haematobium, S. mansoni and S. japonicum (1). For S. mansoni, some eggs leave the body in the faeces and hatch in water to liberate miracidium larvae, which infect certain types of freshwater snails (2). Within the snail, the parasites multiply asexually to produce free-swimming cercariae larvae, and these then infect people by skin penetration (3). The adults do not multiply in the body but instead live there for several years, producing eggs (4).

Schistosomiasis infection constitutes a major public health problem, particularly in countries where the disease is endemic. Worldwide (Figure 1), it is estimated 779 million people at the risk of contracting schistosomiasis, while about 210 million are infected with the disease (5). The acute or short-term consequences of schistosomiasis infection include skin rashes, fever and fatigue, while chronic or long-term effects involve damage to internal organs such as the liver, spleen and gallbladder.

Traditionally, the schistosomiasis diagnosis has been performed by direct parasitological techniques, such as the Kato–Katz method (6). However, in cases of a low infections the Kato-Katz method is not efficient and may be detected by means of serological immunodiagnostic tests or molecular techniques (7–10). Alternative methods for the schistosomiasis diagnosis can potentially be identified through metabolomics studies, wherein the metabolites found associated with

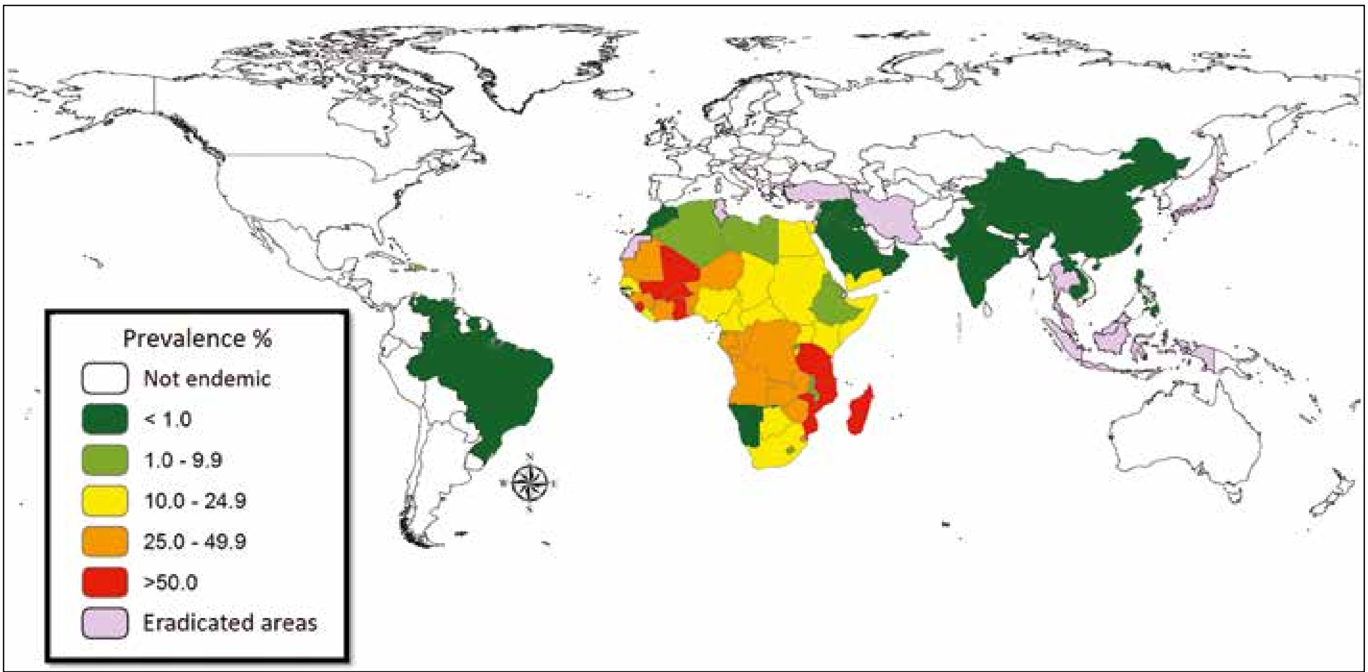


Figure 1. The world map showing the prevalence data of schistosomiasis in several countries.

Figure 2: Example of training session using the virtual machine in this Metabolomics analysis.

Schistosoma infections (e.g., in an animal's blood, stool or urine) are profiled to identify characteristic biomarkers.

Research collaboration

This study was a partnership between the Oswaldo Cruz Foundation (https://portal.fiocruz.br/pt-br) and the University of Auckland (UoA). The samples were obtained at the Aggeu Magalhães Institute (Brazil) and shipped to the UoA Mass Spectrometry Facility. The Mass Spectrometry facility provides the service to do metabolomics analysis for many researchers and students (Figure 2) around the University of Auckland. The GC-MS instruments and the data analysis computer are constantly in use. For this reason, this facility together with the Centre for eResearch created a virtual machine (VM) that enables a multi-user platform making the data processing easier and faster. This VM made possible a quickly way to work with all the required software and data generated by the equipment from everywhere and any computer to complete this metabolomics project.

Results

This study used the methyl chloroformate (MCF) methodology as described by Smart et al., 2010 (11) to do the metabolomics analysis. The samples were obtained from three groups of 5 mice (Figure 3). Two groups infected (low and high parasitological load) and one group not infected (used as a control).

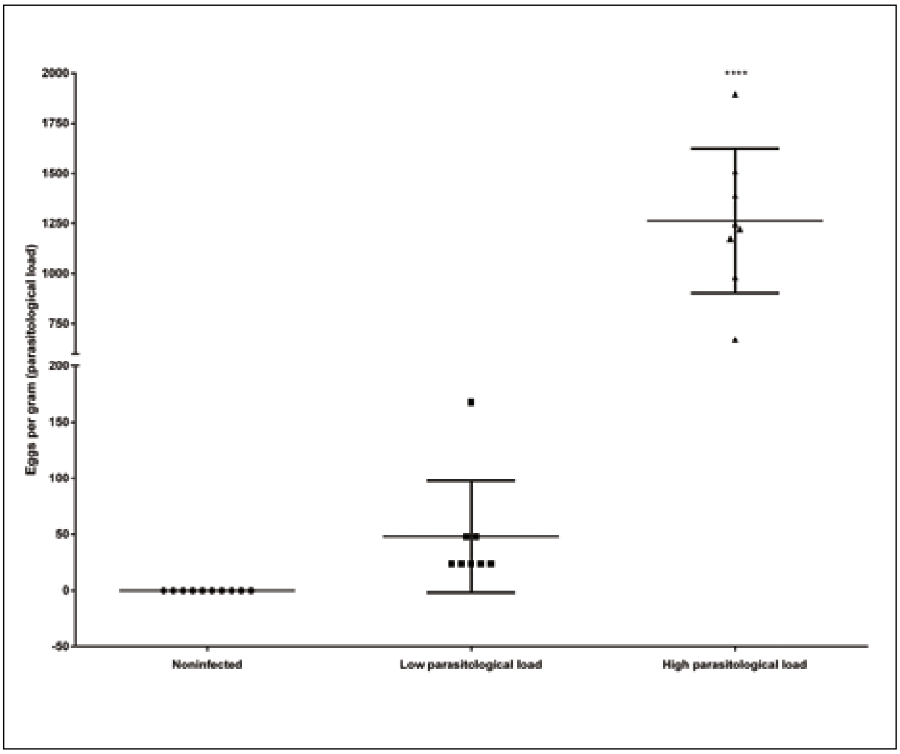


Figure 3. Result of a parasitological stool examination showing the difference between the groups.

The metabolic profile was different between groups with some metabolites contributing to these differences (Figure 4), like Hippuric acid and Alanine. The Hippuric acid was related to a low load infection as the Alanine was related to a high load infection (Figure 5).

Previous studies found similar results using different techniques (Wang et al., 2004, García-Pérez et al., 2008 and Li et al., 2011 (12–14)) showing the relevance of these metabolites to possible be used as biomarkers for schistosomiases diagnosis.

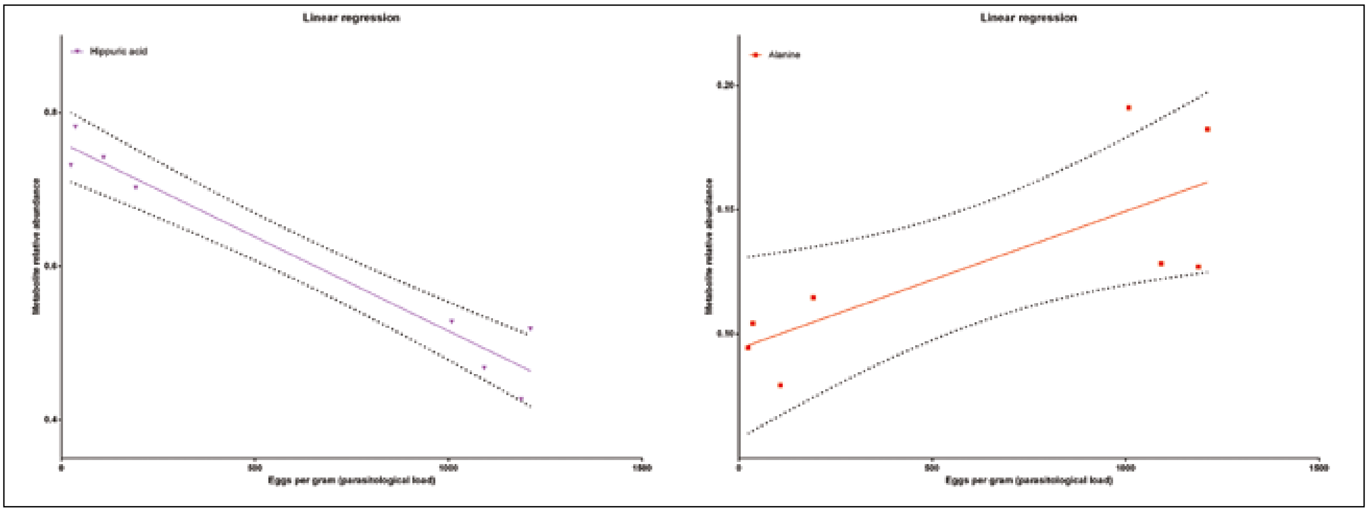


Figure 5. The linear regression showing a negative correlation between the Hippuric acid relative abundance and the parasitological load and a positive correlation between the Alanine relative abundance and the parasitological load.

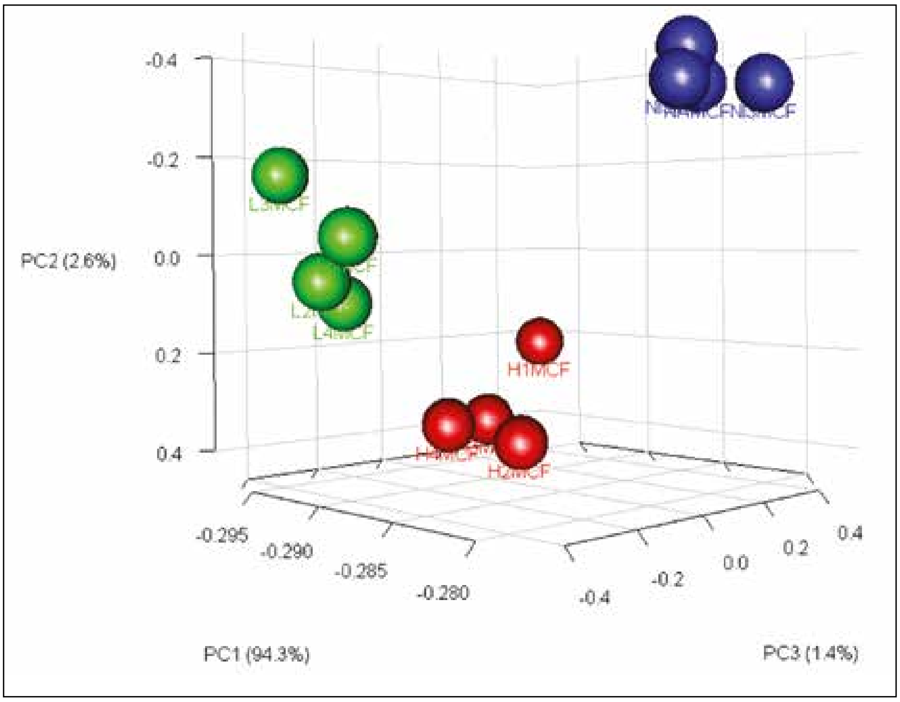


Figure 4. Principal Component Analysis (PCA) showing the separation and the groups formation.

References

1. *Rollinson D, Simpson AJ. The Biology of Schistosomes: From Genes to Latrines. Parasitology. London: Academic Press; 1987. 472 p.*
2. *Gryseels B, Polman K, Clerinx J, Kestens L. Human schistosomiasis. Lancet. 2006;368(9541):1106–18.*
3. *Wilson RA, Coulson PS. Schistosoma mansoni: dynamics of migration through the vascular system of the mouse. Parasitology [Internet]. 1986 Feb [cited 2016 Jan 14];92 (Pt 1):83–100.*
4. *Loverde PT, Chen L. Schistosome female reproductive development. Parasitol Today [Internet]. 1991;7(11):303–8.*
5. *Steinmann P, Keiser J, Bos R, Tanner M, Utzinger J. Schistosomiasis and water resources development: systematic review, meta-analysis, and estimates of people at risk. Lancet Infect Dis [Internet]. 2006;6(7):411–25.*
6. *Katz N, Chaves A, Pellegrino J. A Simple Device for Quantitative Stool Thick-Smear Technique in Schistosomiasis mansoni. Rev Inst Med trop. 1972;14(6):397–400.*
7. *Rabello A. Diagnosing Schistosomiasis. Mem Inst Oswaldo Cruz [Internet]. 1997 Sep [cited 2017 Mar 7];92(5):669–76.*
8. *Pontes LA, Oliveira MC, Katz N, Dias-Neto E, Rabello A. Comparison of a polymerase chain reaction and the Kato-Katz technique for diagnosing infection with Schistosoma mansoni. Am J Trop Med Hyg. 2003;68(6):652–6.*
9. *Cavalcanti MG, Silva LF, Peralta RHS, Barreto MGM, Peralta JM. Schistosomiasis in areas of low endemicity: A new era in diagnosis. Trends Parasitol. 2013;29(2):75–82.*
10. *Gray DJ, Ross AG, Li Y-S, McManus DP. Diagnosis and management of schistosomiasis. BMJ. 2011;342(May):d2651.*
11. *Smart KF, Aggio RBM, Van Houtte JR, Villas-Bôas SG. Analytical platform for metabolome analysis of microbial cells using methyl chloroformate derivatization followed by gas chromatography-mass spectrometry. Nat Protoc. 2010;5(10):1709–29.*
12. *Wang Y, Holmes E, Nicholson JK, Cloarec O, Chollet J, Tanner M, et al. Metabonomic investigations in mice infected with Schistosoma mansoni: an approach for biomarker identification. Proc Natl Acad Sci U S A [Internet]. 2004;101(34):12676–81.*
13. *García-Pérez I, Whitfield P, Bartlett A, Angulo S, Legido-Quigley C, Hanna-Brown M, et al. Metabolic fingerprinting of Schistosoma mansoni infection in mice urine with capillary electrophoresis. Electrophoresis. 2008;29(15):3201–6.*
14. *Li J V, Saric J, Wang Y, Keiser J, Utzinger J, Holmes E. Chemometric analysis of biofluids from mice experimentally infected with Schistosoma mansoni. Parasit Vectors [Internet]. 2011;4(1):179.*

Case study 12 | Research data management

GIVING PACIFIC RESEARCH GREATER REACH

Dr Gerry Cotterell, Research Operations Manager, New Zealand Institute for Pacific Research.

New Zealand Institute for Pacific Research

Launched in March 2016, the New Zealand Institute for Pacific Research (NZIPR) is a national institute aimed at promoting and supporting excellence in Pacific research, in order to deliver a world-class research programme focused on Pacific development, investment and foreign-policy issues. Key support partners are the Ministry of Foreign Affairs and Trade and the Pacific Cooperation Foundation. The NZIPR comprises a consortium made up of the University of Auckland, Otago University and the Auckland University of Technology. Together the university partners bring expansive and unique Pacific research capacity and relationships, as well as educating three quarters of New Zealand's Pacific students and employing ninety percent of New Zealand's academics of Pacific Island descent.

Pacific research repository

As part of its mandate, NZIPR was charged with establishing a repository to enable the collection, collation, and dissemination of research data and outputs. NZIPR worked with the Centre for eResearch to connect with the Analysis & Policy Observatory (APO) in order to host the Repository (Figure 1). APO is a not-for-profit, open access research collection and information service curating key resources to support evidence-informed policy and practice.

This collection contains a wide range of digitised documents, government records, research reports, postgraduate theses and more. As a repository for the New Zealand Institute for Pacific Research, the collection plays a key role in disseminating public policy information from the NZIPR partners, together with relevant content by researchers in the Pacific and beyond. Using the collection repository with the website portal ensures there are tools to publish, search, manage and track content, people and organizations. Access to the repository is seamlessly integrated from within the portal (http://www.nzipr.ac.nz) to the repository (http://apo.org.au/collections/pacific-research), making it simple to go from searching to discovery and utilization of the information. The repository will be expanded over time as more resources are located.

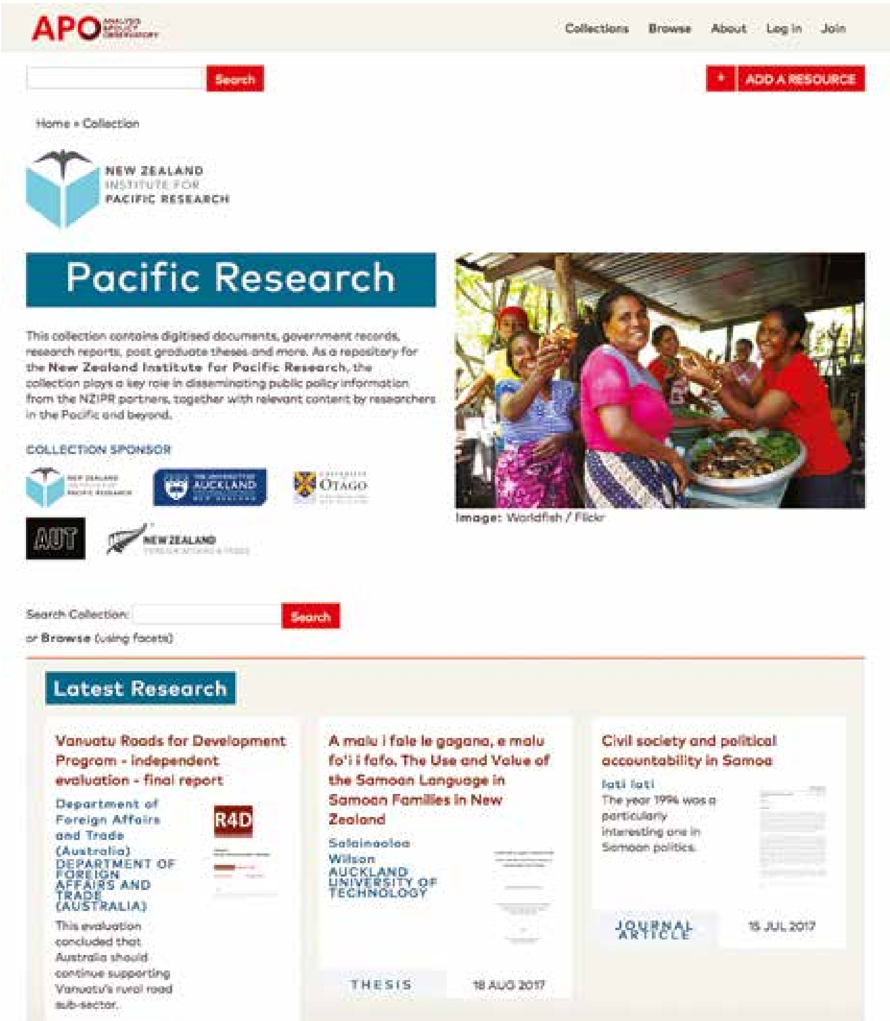


Figure 1. NZIPR Pacific Research Collection hosted at APO.

Case study 13 | Visualisation & analytics and collaboration with researchers

DISPOSITION OF MICROSOFT HOLOLENSES FOR A POP-UP REALITY SHOP TO DEMONSTRATE THE PROGRESS OF A RESEARCH PROJECT

Associate Professor Uwe Rieger, School of Architecture and Planning.

The research team is working closely with Datacom NZ on new hybrid retail concepts, combining physical and online shopping. The project is embedded in the research at the Lab for Digital Spatial Operations [arc/sec] at the University of Auckland. A team of four Architectural postgraduate students, Anita Chin, Linus Goh, Ricky Tung and Bevin Liang are working on a one-year project to investigate how we can merge physical properties with digital information to form a responsive architecture. The team looks to utilise AR headset tools to actively link physical touchable matter with digital materiality, to create a unique user experience in a haptic digital space.

The Pop-up Reality Shop is situated in Auckland's first and leading innovation hub for augmented reality and virtual reality technologies; the 'AR VR Garage'.

Application

These devices allow us to simultaneously reinvent and reimagine the future of interactive space, going beyond our traditional understanding of both architecture and technology, towards a collaborative and experimental mode of practice.

The research direction for Pop-Up Reality Shop is to test new tangible interactions and conditions that could occur within the context of a retail experience.

This is executed by the use of a physical structure that is designed to be flexible and becomes a framework that acts as an external input device to control and instruct our digital designs with the utilisation of the augmented reality headset HoloLens.

Microsoft HoloLens

The HoloLens and its initial development kit has been a pivotal technological tool for this investigation. The HoloLens will be used in conjunction with other existing tools such as Optitrack Motion Capture and Unity3D in order to create performative physical environments that embody programmable virtual characteristics. It is the flexible and highly manoeuvrable nature of the HoloLens which makes it a great asset for Pop-up Realities; being able to experience a space untethered from a computer makes the architectural experience much more seamless, immersive and naturalised. Looking ahead, the arc/sec Lab will explore the use of two HoloLenses for extended streaming purposes as well as improved user interaction between multiple users.

Thanks for the advice and technology demonstration by the specialist from the Centre for eResearch which proven to be highly valuable for the progress of the project. We expect to show first results at the ACADIA Conference at MIT in Boston in November 2017.

Case study 14 | Visualisation & analytics and collaboration with researchers

ALTER: BETWEEN HUMAN AND NONHUMAN – A VR ART EXHIBITION

Deborah Lawler-Dormer, Doctoral Candidate, Faculty of Creative Arts and Industries.

Overview

Alter: Between human and nonhuman is a curatorial project that has a component which is a Virtual Reality exhibition. The core curatorial project consisted of a group exhibition featuring selected works by international artists. The physical exhibition was staged at the Gus Fisher Gallery at the University of Auckland in May 2016 and reconfigured as a VR exhibition in collaboration with the Centre for eResearch (CeR) in July 2017.

This exhibition, ***Alter***, foregrounded the works of international artists working within a technoscientific space, teasing out the relations between corporeality and technologies that allow a rethinking of the posthuman. Each artwork was a product of the shifting collaborative process involving artists, engineers, computer scientists, neuroscientists and medical practitioners often incorporating the artist's corporeality as data or characteristics as part of the technoscientific art work.

Alter featured the following artists and their respective artworks: Jane Prophet (UK, USA): Neuro Memento Mori (2015); Nina Sellars (Aus): Scan (2012); Agatha Haines (UK): Drones with Desire (2015); Elena Knox (Aus): Canny (2013) and Comfortable and Alive (2014); Stelarc (AUS): Prosthetic Head (2003) and Deborah Lawler-Dormer and Mark Sagar (NZ): Leah (2016).

Donning a VR headset entails immersion. By choosing the HTC-Vive platform the participant has full body movement and can navigate the space using handheld remotes. The VR environment is both a performative and a computational environment. The participant is still engaging with the particularities of both content and interactivity of each selected art work.

Curating, and recreating, virtual spaces for contemporary artworks, that require some form of engagement other than standing and viewing, is a relatively new curatorial challenge and has not been well theorized. This practice-led research project engages with this transdisciplinary provocation.

Design requirements

The requirements for the VR system included:

1. users to experience each artwork interactively and explore a virtual engagement with its content and unique experience;
2. the art exhibition to be placed within a replicated version of the University of Auckland's Gus Fisher Gallery in the same configuration as the physical exhibition staged in 2016;
3. to be staged at CeR for examination by external examiners;

The Centre for eResearch visualisation suite

4. to be accessible to the artists and selected galleries for virtual experience and documentation of artworks and exhibition.

Engaging with CeR for solution

Over the course of several consultative meetings, the CeR team of Nick Young and Bianca Haux, worked alongside the curator. Following this, the curator brought in experts, supervisors and participants, who were drawn from a range of disciplines across the University, including from the Dance Studies Programme, Elam Fine Arts School, School of Psychology and the Centre for Art Studies. Throughout, the CeR staff were thorough and professional in their approach and worked with the curator to ensure an immersive and creative experience for all of the participants.

CeR adapted a prototype developed by Werner Ollewagen from the Laboratory for Animate Technologies at the Auckland Bioengineering Institute. CeR took the prototype, built in the gaming engine Unity, and designed additional functionality for each of the artworks in consultation with the curator. The curator provided a range of sources to ensure that the artworks were virtually close to the original artifacts and experience. This included using files supplied by artists from a range of sources and programmes. CeR researched specialized functionalities required for interactivity and contributed unique solutions. This enabled such features as:

1. adapting a cellphone interface and accessing an animated sequence of MRI images;
2. the ability to pick up objects and view from a variety of angles;
3. transferring rapid prototyping data files into Unity;
4. rebuilding older interfaces and accessing online databases for real-time conversations with virtual agents;
5. transferring interactive generative sound into a virtual triggered project gallery;
6. developing an avatar virtual body for user's virtual engagement in exhibition. This encouraged a range of user experiences from virtual hypnosis, keyboard and cellphone use, standing, lying, floating and talking with a conversational agent;
7. enabling listening to audio compositions while interacting with each artwork with limited sound leakage of other displayed works.

Additional features relating to the architectural surfacing of the Gus Fisher Gallery were added that closely represented the physical attributes of the existing campus building.

Future project aspirations

The project has been displayed for examination and is currently under assessment. It is anticipated that the project will be made available to the participating artists, Gus Fisher Gallery and other selected galleries by the curator in 2018. ***Alter*** activates a methodology for curatorial practices relating to transdisciplinary art-science-technology projects that move between physical and virtual experiential approaches.

Case study 15 | Visualisation & analytics and collaboration with researchers

WANDERING AROUND THE MOLECULAR LANDSCAPE: EMBRACING VIRTUAL REALITY AS A RESEARCH SHOWCASING OUTREACH AND TEACHING TOOL

Professor Juliet Gerrard, Department of Biology Sciences; Professor Mark Gahegan, Centre for eResearch; Dr Justin O'Sullivan, Senior Research Fellow, Liggins Institute; Matt Pendred, Motion Graphic Artist, Department of Media Production.

Virtual reality on campus

Virtual reality in science is an exciting, effective and rapid expanding field in today's research. Auckland molecular research at the School of Biological Sciences aims to build capability in the use of virtual reality technology for molecular visualisation, i.e. to create an experience which enables the user to actually climb inside a molecule and walk around it.

The required hardware was funded by the MacDiarmid Institute, and a workshop is run by PhD student Kyle Webster to build interest in using the technology to showcase research on campus.

The team has run a small pilot study using existing software to provide a proof of concept that is indeed possible to climb inside a virtual molecule. To keep up with the rapid growing technology, a dedicated software developer and a motion graphic designer are part of the Auckland team to connect local and international expertise so that the Auckland molecules are professionally rendered to maximise the impact as a research, teaching and outreach tool.

Application

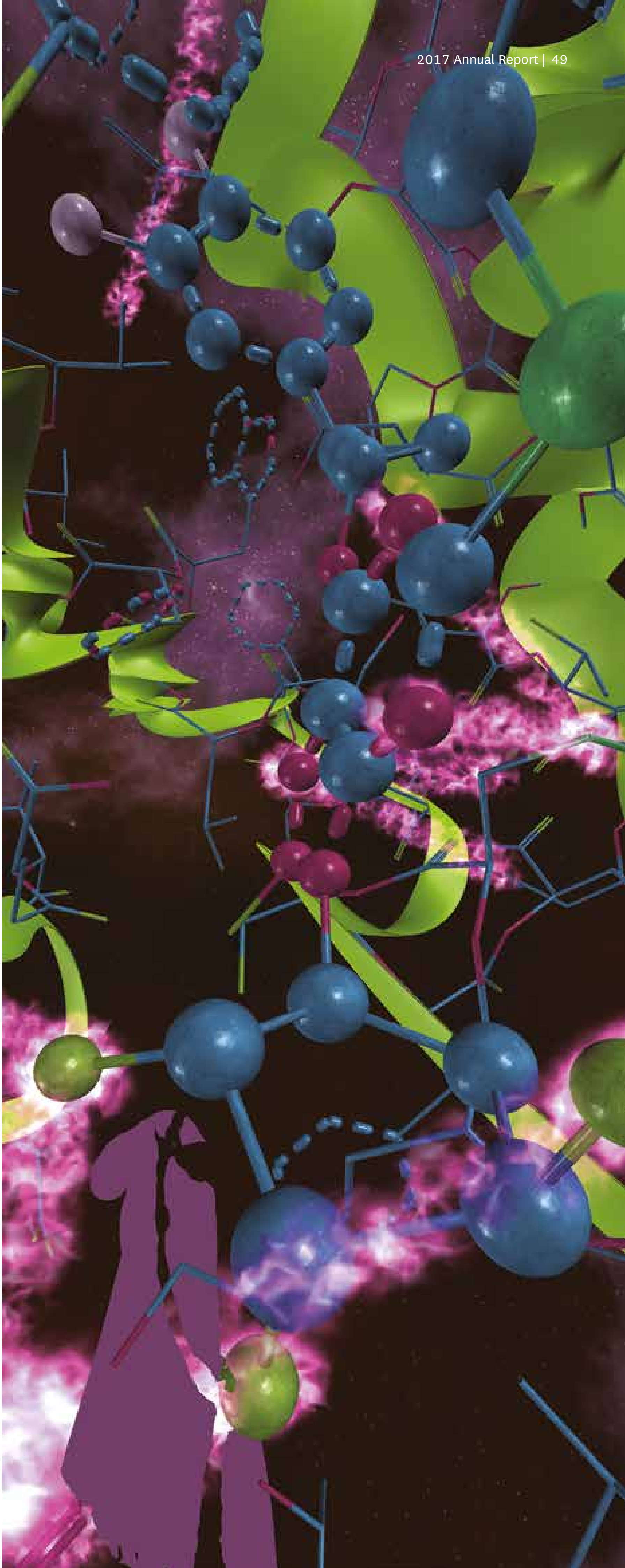
The team will create awareness on campus of the technology, so that researchers can turn their molecules into virtual reality experiences and stroll around them. In the first instance, the team will use proteins in the Lab from Professor Gerrard and her collaborators as case studies, moving onto the more complex case of Dr O'Sullivan's complex DNA structures in whole chromosomes. We envisage that many other projects will be nucleated by the programmer, researchers and PhD students who will be excited to turn their 3D data into a virtual reality experience.

Collaboration

Professor Gerrard was introduced to the Vive VR technology at the Googleplex – an international SciFoo Conference – and was able to keep updates of other institutions that are rapid adopters of this technology. For example, Professor Tom Davies, Director of the ARC Centre of Excellence in Convergent Bio-Nano Science and Technology (https://www.cbns.org.au/) and Professor Paul Bonnington from CAVE-2 Centre (http://www.monash.edu/mivp) are some of the more advanced players in the field who have agreed to share their expertise as part of this programmer.

More locally, the team have connected with the VR Garage and the Media Design School, where Steve Dorner, Associate Dean (Programme Leader of Bachelor of Art and Design) is interested in assisting with scaling the technology for large undergraduate audiences (perhaps through augmented reality).

This would provide an opportunity to showcase Auckland molecular research at the centre of the virtual reality experience, and tailor it to attract students and researchers to the University of Auckland.



The role of CeR

CeR has an excellent facilities and expertise in visulisation and analytic area. The Centre has been working with researchers across the University including Auckland Bioengineering Institute (ABI), Faculty of Medical and Health Sciences, Population Health, Mathematics, Marine Science, New Zealand Institute for Pacific Research names but a few, to build visual spatial maps, generate animated model, 3D modelling tool etc.

The benefits to the University

The project is linked to the University strategic priorities and provide the following benefits:

1. Attracting talented undergraduates and postgraduates throughout NZ and internationally.
2. Enhancing the University's ability to attract externally funded research and in particular, offshore funding.
3. Enhancing the student experience, including the experience of international students.
4. Enhancing learning and teaching.
5. Enhancing the University's international relationships and standing.

Case study 16 | Visualisation & analytics and collaboration with researchers

VISUALISING THE NEW ZEALAND INDEX OF MULTIPLE DEPRIVATION (IMD)

Dr Daniel John Exeter, Senior Lecturer, Epidemiology and Biostatistics; Visualisation and Analytics service are provided by Nick Young and Bianca Haux, Research IT Specialists, Centre for eResearch.

Background

The New Zealand Index of Multiple Deprivation (IMD) allows one to look at disadvantage profile in overall terms, as well as in terms of seven domains of deprivation: employment, income, crime, housing, health, education and access. The seven domains are weighted to reflect the relative importance of each domain in representing the key determinants of socio-economic deprivation, the adequacy of their indicators and the robustness of the data that they use. The IMD uses routinely collected data from government departments, census data and methods comparable to current international deprivation indices to measure different forms of disadvantage.

The IMD measures deprivation at the neighborhood level using custom data zones that were specifically developed for social and health research. The New Zealand land mass has 5,958 neighborhood-level data zones that have a mean population of 712 people. In urban settings, data zones can be just a few streets long and wide. Data zones are ranked from the least to most deprived areas (1 to 5958) and grouped into five quintiles for mapping purposes.

The Index of Multiple Deprivation was developed by the IMD team based at Epidemiology & Biostatistics at the School of Population Health: Dr. Daniel John Exeter, Dr. Jinfeng Zhao, Dr. Sue Crengle, Dr. Arier Chi Lun Lee and Michael Browne, with help and support from numerous individuals and organizations.

The vision of the IMD

The research is to look into the impact of deprivation and health inequalities, and improve understanding of service quality and the degree of disparities across population groups, so that it can be used to inform agencies and policy makers, and to prompt system change, a greater responsiveness and equity of health service provision. The aim for the IMD is to be widely used for community advocacy, research, policy and resource allocation. It provides a better measurement of area deprivation, equity of service provision, and a more consistent approach to reporting and monitoring the social climate of New Zealand.

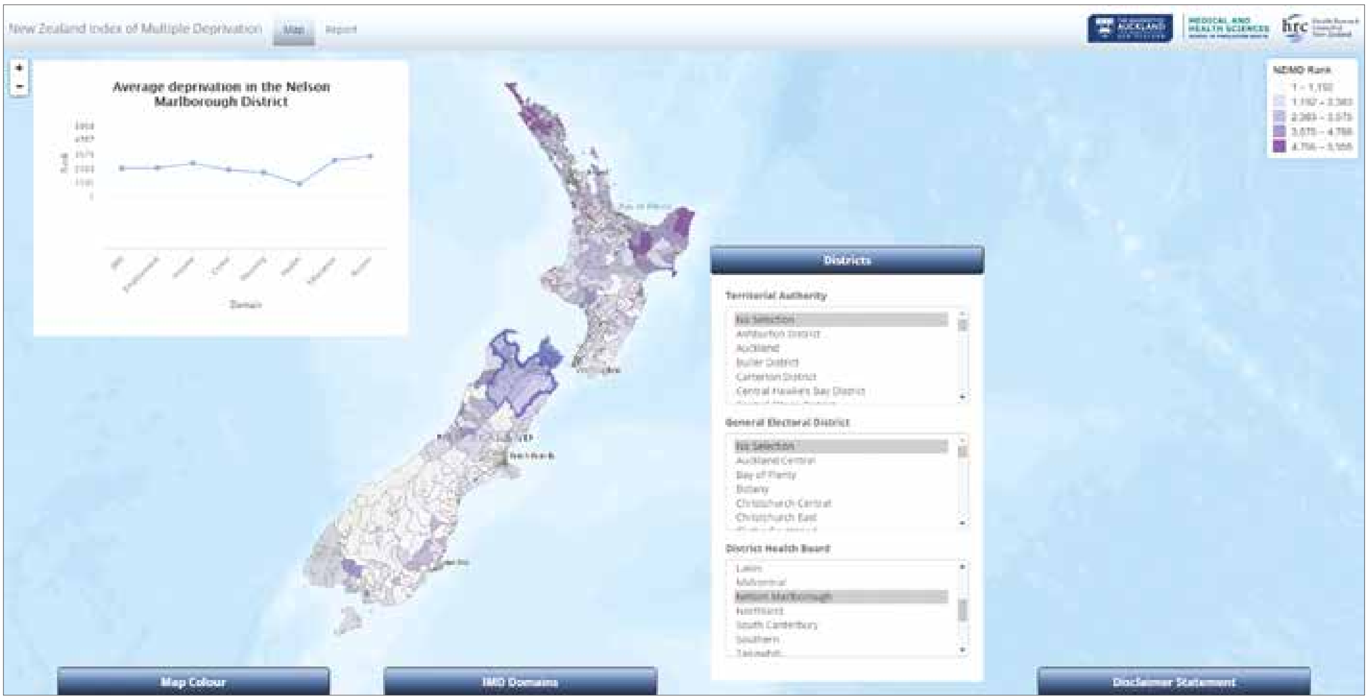
Creating a visual interface for dynamic research analysis

As part of the research project, the IMD team developed a series of online interactive atlases using some proprietary software. The existing atlases are very flexible and informative, but do not provide the ability to create a dynamic report for users to download. The principal investigator Dr. Daniel Exeter therefore sought advice from the visualisation experts at the Centre for eResearch to provide an open-source mapping solution that provided such functionality.

An interactive map interface which can be used to explore the geography of deprivation and its association with a given health or social outcome over the seven domains was developed. It illustrates the IMD by ranking the data zones from the least to most deprived areas (1 to 5958) and grouping it into five color coded quintiles. Quintile 1 (Q1 - light shading) represents the least deprived 20% of data zones in the whole of NZ; while Q5 (dark shading) represents the most deprived 20%. When a user zooms into a place of interest and selects some data zones of interest, the line graph updates to highlight the rank of overall deprivation (the IMD) and its 7 domains.

This is useful in highlighting that the drivers of deprivation can differ markedly between neighborhoods, even if they have similar IMD ranks overall.

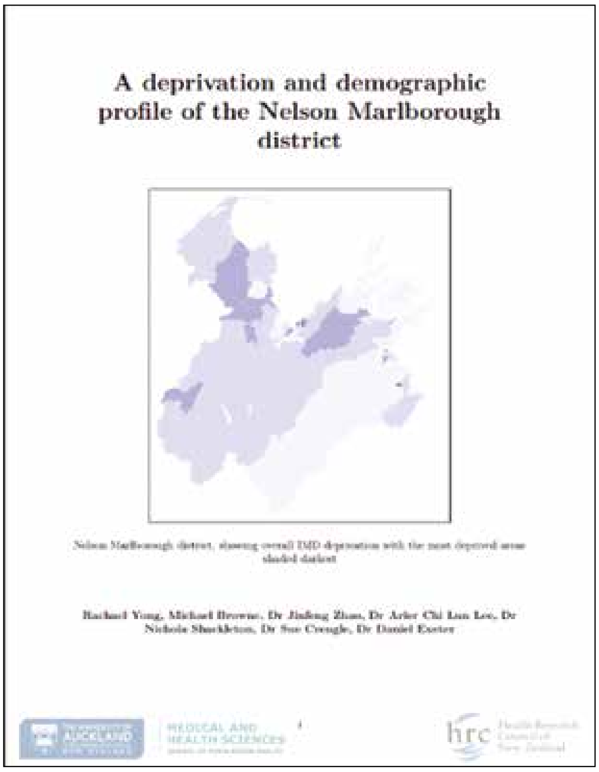
In addition to the interactive map, staff at the Centre for eResearch created an application that provides the means to download dynamically created reports for a particular administrative area of interest. At this stage, users can select from District Health Boards, Territorial Authorities, or General Electoral Districts to zoom into. Every report includes



several maps and summary statistics calculated on the fly depending on the user's choice of district. This combination allows for a great overview while also providing the means to gather selective and detailed information depending on the viewer's area of interest.

Who is the intended user community?

This project was prompted in part by a request from a District Health Board (DHB) Funding and Planning agency to have reports available within their DHB rather than for the entire population. This atlas will be added to the existing suite of visualisation tools that the IMD team has developed, and to provide outreach for the general public, as well as researchers, policy analysts and organisations who are interested in better understanding the socio-economic circumstances of the communities they serve.



Case study 17 | Visualisation and NeSI High Performance Computing platform

MODERNISING MODELS TO HELP DIAGNOSE OR TREAT DISEASE AND INJURY

David Ladd, Doctoral Candidate, Dr John Rugis, Dr Chris Bradley, Professor Peter Hunter, Auckland Bioengineering Institute.

Computer modelling techniques and visualisation software packages have been important resources for biomedical engineering research for decades. In New Zealand, a group of University of Auckland researchers have been at the heart of a movement to re-envision some of these tools with open-source approaches and standards-compliant practices. The results will not only enable new combinations and better linking of computational models, they will also improve researchers' understanding of physiological functions and abilities to diagnose or treat injury or disease.

Continuum Mechanics, Imaging, Signal Processing and System Identification (CMISS)

Since 2005, Chris Bradley, a Senior Research Fellow at the Auckland Bioengineering Institute (ABI), has been leading an open-source reworking of the CMISS programme. CMISS is a set of libraries and applications widely used in bioengineering to support computational modelling and visualisation. It was created in 1980 by Peter Hunter, the current Director of the ABI and a Professor of Engineering Science at the University. Professor Poul Nielsen and Professor Andrew McCulloch also made major early contributions to the original CMISS packages.

Integration through OpenCMISS

Bradley's updated version of the programme, called OpenCMISS, aims to integrate modern programming languages, data structures, and high performance hardware.

"This significant re-engineering effort represents a complete upgrade in functionality and modelling capability, particularly in terms of increased ability to optimise simulation performance on high performance, and in particular distributed, architectures," Bradley says.

Today, the OpenCMISS project is a collaborative effort between groups based at the University of Auckland, the University of Stuttgart, and the University of Melbourne, and is funded by both European and New Zealand research funding agencies. Developers are not only modernising OpenCMISS's packages, they are making the tools more accessible and functional for wider use. For example, in response to demands for better versatility across systems, Bradley and other developers have been building OpenCMISS packages to run on a variety of platforms, from Mac and Windows machines to Linux systems. NeSI's Pan cluster was used to support some of this work, as Bradley worked with Chris Scott from NeSI's Solutions team to design the underlying framework to implement OpenCMISS in a High Performance Computing (HPC) environment.

Another key OpenCMISS developer at the ABI is David Ladd. Ladd has been refining OpenCMISS's fluid mechanics package, which aligned well with his doctoral thesis, exploring how computational fluid dynamics (CFD) models could be used for more detailed and clinically relevant analysis of fluid mechanical scenarios in cardiovascular systems. "As computing advances have made CFD models of the vasculature more tractable, their adoption into clinical application has not been equally forthcoming," Ladd says. "In my doctoral work, I approached vascular modelling from two sides as steps toward closing this gap."

His project was a perfect example of a key challenge OpenCMISS hopes to address: How can researchers build models hierarchically to support better communication between the models and glean better insights from the results.

Working with patient MRI velocimetry data and using open-source and Physiome standards compliant methods, Ladd created multiscale CFD models to improve clinicians' ability to study a vascular system. In addition, he also created an adaptable framework for vascular fluid mechanical modelling that could be used, shared, and extended. In one application, Ladd modelled the development of a stenosis in an iliac bifurcation, where the descending aorta splits into the two main blood vessels for the legs. "You can't usually look at a model of an artery on its own, you

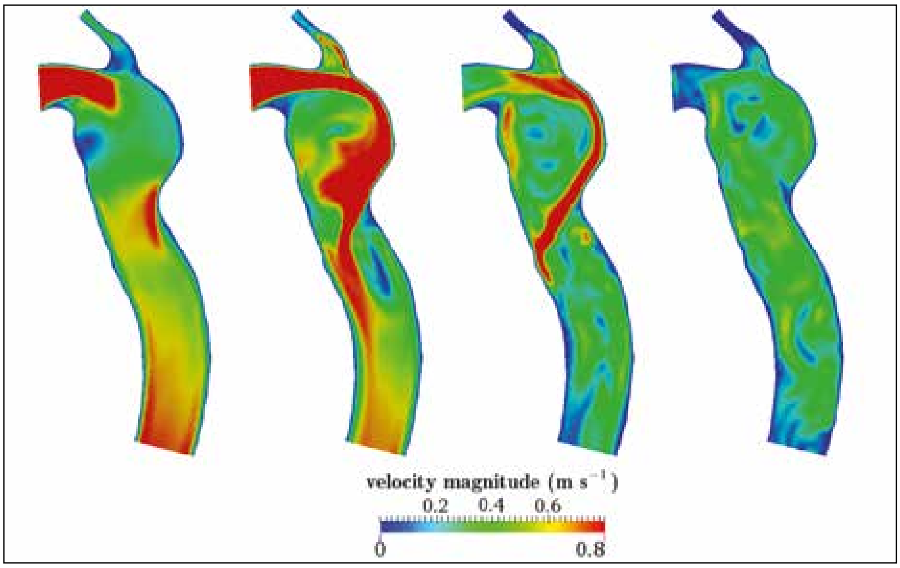


Figure 1. Illustrates a cross-section of development of flow in an aortic aneurysm, images 100ms apart.

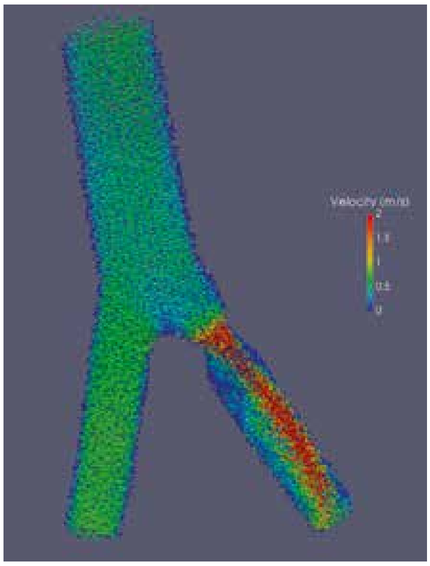


Figure 2. Iliac bifurcation visualization

also have to consider the upstream forcing mechanisms of the heart and the perfusion into downstream vascular beds," he says. "The novel aspect of my work was that we were doing it in an open source and standards-compliant way." Ladd says.

The use of NeSI High Performance Computing

To return results in days rather than weeks, Ladd harnessed the computational power of NeSI. "When you're running several big CFD problems and looking at problems on the order of hundreds of thousands or millions of degrees of freedom, you need HPC," he says. "NeSI's resources can give researchers access to thousands of cores to compute on, versus the few that you can get from a desktop or tens that you can get from a small cluster. It's really useful because you can run several big jobs at once."

By the end of his thesis work, Ladd had created multiple open-source methods and solvers for the OpenCMISS library.

"We're hoping it's a solid base upon which we can start constructing multi-physics, multi-scale models in a standards-compliant way," he says.

As OpenCMISS sets the stage for researchers to create more comprehensive models, other tools such as data visualisation software will help researchers analyse those models on a deeper level. John Rugis, Visualisation Specialist at the Centre for eResearch, knows this realm well. "The goal of any visualisation is to show 3D dynamic data in a way that is intuitively obvious," he says. "With visualisation, you know it works when someone looks at it and they don't have to think about what they're looking at, they're just seeing something happening."

Rugis recently created a dynamic 3D visualisation of Ladd's iliac stenosis simulation (See Figure 2). A normal iliac bifurcation is featured next to one with a stenosis, making the effects of a narrowed vessel easy to see. Blood flow velocity is colour coded with blue-green indicating low velocity and yellow-red indicating high velocity. This single heart-beat animation shows the high velocity blood flow stream that results from blood vessel constriction. "It's showing the actual dynamic effect of that condition," says Rugis. "It's real science data, not just a pretty picture."

Together, the teams and resources of the ABI, Centre for eResearch, and NeSI are helping researchers to gain broader insights and better understanding of physiological processes.

With programmes like OpenCMISS improving model creation and performance, HPC resources powering increasingly complex simulations, and data visualization techniques aiding with model analysis, researchers are brought steps closer to developing models with greater predictive capabilities and direct clinical applications.

Case study 18 | NeSI High Performance Computing platform

AERODYNAMICS MODELLING PAVES THE WAY FOR IMPROVED YACHT DESIGNS

Stefano Nava, Dr Stuart Norris and Dr John Cater, Faculty of Engineering.

Study of aerodynamics of high performance yachts

As any experienced seafarer knows, good aerodynamics is the key to smooth sailing. In order to improve the design of sails, hulls, and masts, we need to better understand the properties of sail aerodynamics in a range of conditions.

At the University of Auckland, Stefano Nava, Stuart Norris and John Cater in the Faculty of Engineering are using NeSI computing resources to study the aerodynamics of high performance yachts.

"The fluid dynamics that characterises sailing yachts is extremely complex, due to the fact that the yacht is partly immersed in water and partly in air, with the flow being three-dimensional and turbulent," the group notes in their latest paper, published in the International Journal of Heat and Fluid Flow. "Experimental and numerical studies have created a large body of knowledge of the physics of the problem, but at the same time have highlighted the limits to the predictive methodologies available for designers."

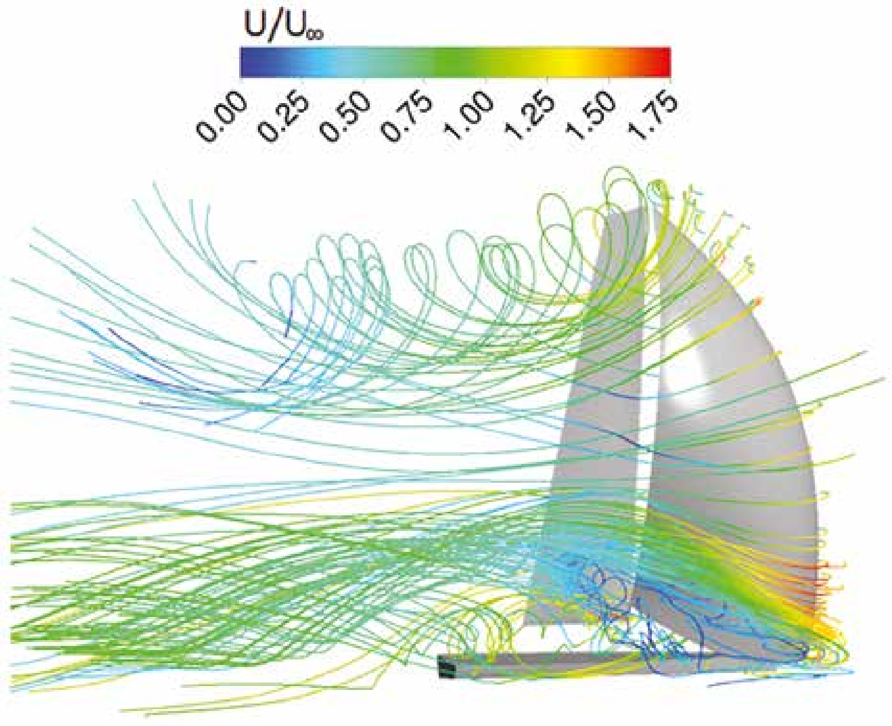


Figure 1. Velocity streamlines of the time averaged flow around the yacht model.

The University of Auckland wind tunnel set up for a downwind sailing experiment.

Modelling turbulent flow

Their paper compared two methods for modelling turbulent flow – Large Eddy Simulation (LES) and Reynolds Averaged Navier-Stokes (RANS) – to see if the extra computational resources required for the first method would enable a more accurate prediction of the aerodynamics of upwind sailing.

At first, the computational demands involved with this project posed a challenge. Complications with both the hardware and software components for their simulations led to performance issues and long wait times for results.

"Without computational resources such as the Pan cluster at NeSI, we would be unable to use accurate, yet computationally demanding, predictive methods such as Large Eddy Simulation."

Dedicated resources from NeSI's Pan cluster

Jordi Blasco, a member of NeSI's Solutions team, stepped in to help. First, he created a short-term reservation on Pan cluster to provide a dedicated set of resources. He then found ways to improve the performance of simulation software Nava was using and the resilience of his workflow.

As a result, Nava's computations were nearly five times more efficient and used less resources than expected, which freed up computational time on Pan for others.

"Given that Stefano had already spent weeks of computing hours on this project, the 500% efficiency improvement meant a lot, not only for him but for other users of the Pan cluster as well," said Blasco.

Thanks to these adjustments, Nava was able to compute and collect the results he needed. The simulations were compared with experimental data previously gathered in the wind tunnels of the University of Auckland's Aerodynamics Laboratory.

Ultimately, when they compared the LES and RANS results with their experimental data, they found the LES model's capabilities to be superior.

"In both cases the LES model has been shown to be able to predict the correct pressure distribution for this separated flow, while the RANS simulations shows less accurate results," their paper stated.

These results add to the growing body of work studying yacht aerodynamics. The researchers are continuing their study into the more demanding case of modelling downwind sails such as spinnakers, to be presented at the INNOV'SAIL conference this June in Lorient, France. Organised by the Cité de la Voile Eric Tabarly and the French Naval Academy Research Institute, INNOV'SAIL is an international forum for the presentation and discussion of the latest scientific and technologic research and its application in the field of high performance yachts and competitive sailing.

"Access to supercomputing resources is critical to advancing this field of study," says Norris. It is currently not possible to use the methods on a desktop computer, due to their large CPU and memory requirements. The use of the methods enables a better understanding of the flow around yacht sails, enabling the design of faster yacht sails."

References

S. Nava, J.E. Cater and S.E. Norris. Modelling leading edge separation on a flat plate and yacht sails using LES. International Journal of Heat and Fluid Flow,2016. doi:j. ijheatfluidflow.2016.12.004

S. Nava, J.E. Cater and S.E. Norris. A comparison of RANS and LES for upwind sailing aerodynamics. In Proceedings of the 22nd Chesapeake Sailing Yacht Symposium, Annapolis, Maryland, 2016.

S. Nava, J.E. Cater and S.E. Norris. Large eddy simulation of downwind sailing. Accepted for the International Conference On Innovation In High Performance Sailing Yachts, 4th Edition, Lorient, France, 2017.

Case study 19 | NeSI High Performance Computing platform

SHEDDING NEW LIGHT ON DARK MATTER

Professor Richard Easther, Department of Physics.

Professor Richard Easther, Head of the Physics Department at the University of Auckland, has been asking questions about the universe for more than 20 years. He's uncovered just as many new questions as answers in that time, but he and colleagues at Auckland and the University of Canterbury, along with a number of international collaborators, used NeSI resources to shed some potential new light on a dark part of the evolution of the universe.

"The universe is continually making space between the galaxies," says Easther. "The explosion that started our universe is still making space today."

What happened in the seconds following the Big Bang remains a source of many hypotheses and unanswered questions. In particular, researchers propose there was a period of accelerated expansion, called inflation, just after the Big Bang during which the universe expanded dramatically, before settling into a slower state of expansion.

To better understand what happened during that initial inflationary period, a research team including Easther, Grigor Aslanyan, and Layne C. Price from Auckland, and Jenni Adams from Canterbury, looked closely at the connections between the physics of inflation and the existence of ultracompact minihalos (UCMHs) – small dense clouds of dark matter that could form soon after the Big Bang.

Scientists believe that dark matter acts like supporting infrastructure or a skeleton for the universe, providing the extra gravitational force necessary to hold galaxies together and bind them to others. However, scientists have no idea what the dark matter is actually made of, and this poses a deep challenge for both particle physics and astronomy.

There are many competing ideas about what dark matter might be, and not all of them predict the existence of UCMHs. Consequently, finding UCMHs would shed light on the nature of dark matter while revealing information about the universe soon after the Big Bang, similar to the way animal and plant fossils can provide us with information about the earth's conditions in earlier geological times.

What is turning heads about this recent work is the connection it makes between

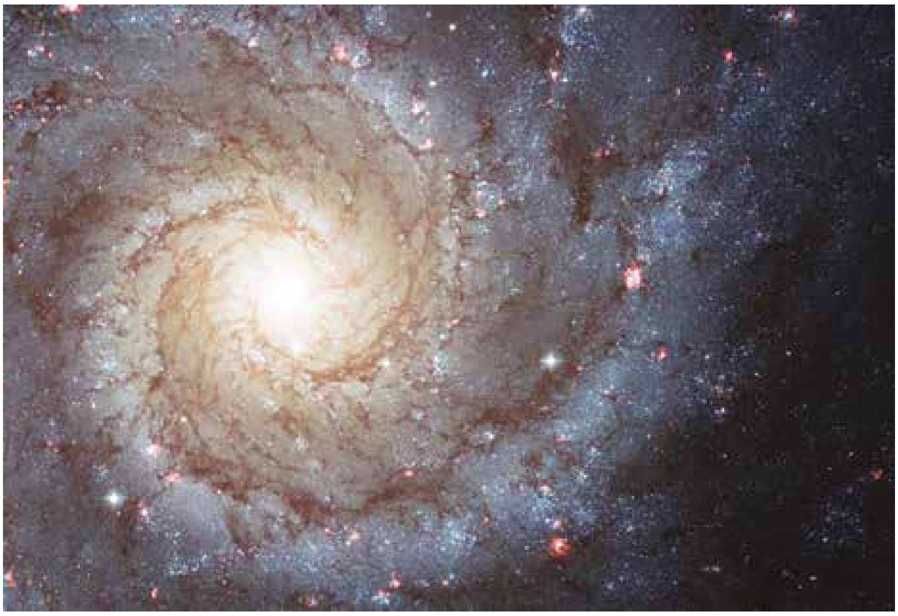


Image Credit: NASA, ESA, and the Hubble Heritage (STScI/AURA)-ESA/Hubble Collaboration

"This was a computationally challenging project and would have been impossible without the serious muscle NeSI provided. The NeSI platform tied together a collaboration between scientists based in New Zealand, Australia, Norway, and the United Kingdom."

the distribution of UCMHs and inflation, providing information to help develop, or maybe completely rule out, some existing inflationary models while refining our understanding of how and when we would expect minihalos to form.

Space.com published an article on this work, and it was also highlighted as an Editor's suggestion in the Physical Review Letters (PRL) journal.

"By simulating the evolution of the perturbations during inflation, and then predicting the mass and abundances of UCMHs, they find that their non-detection is still a useful constraint on the details of inflation," says David Parkinson in PRL's summary of the findings. "This approach significantly widens the available observational pathways to understand the early Universe, and so may provide a future key piece of information as to exactly how inflation took place."

The simulations that led to these key insights were run on NeSI's High Performance Computing systems.

Next, Easther's group is looking at ways in which newly released astrophysical data can further constrain models of inflation and NeSI's platform and services will play an important role in that project as well.

"NeSI's resources makes it possible for New Zealand based scientists to do work at the forefront of modern cosmology," he says.

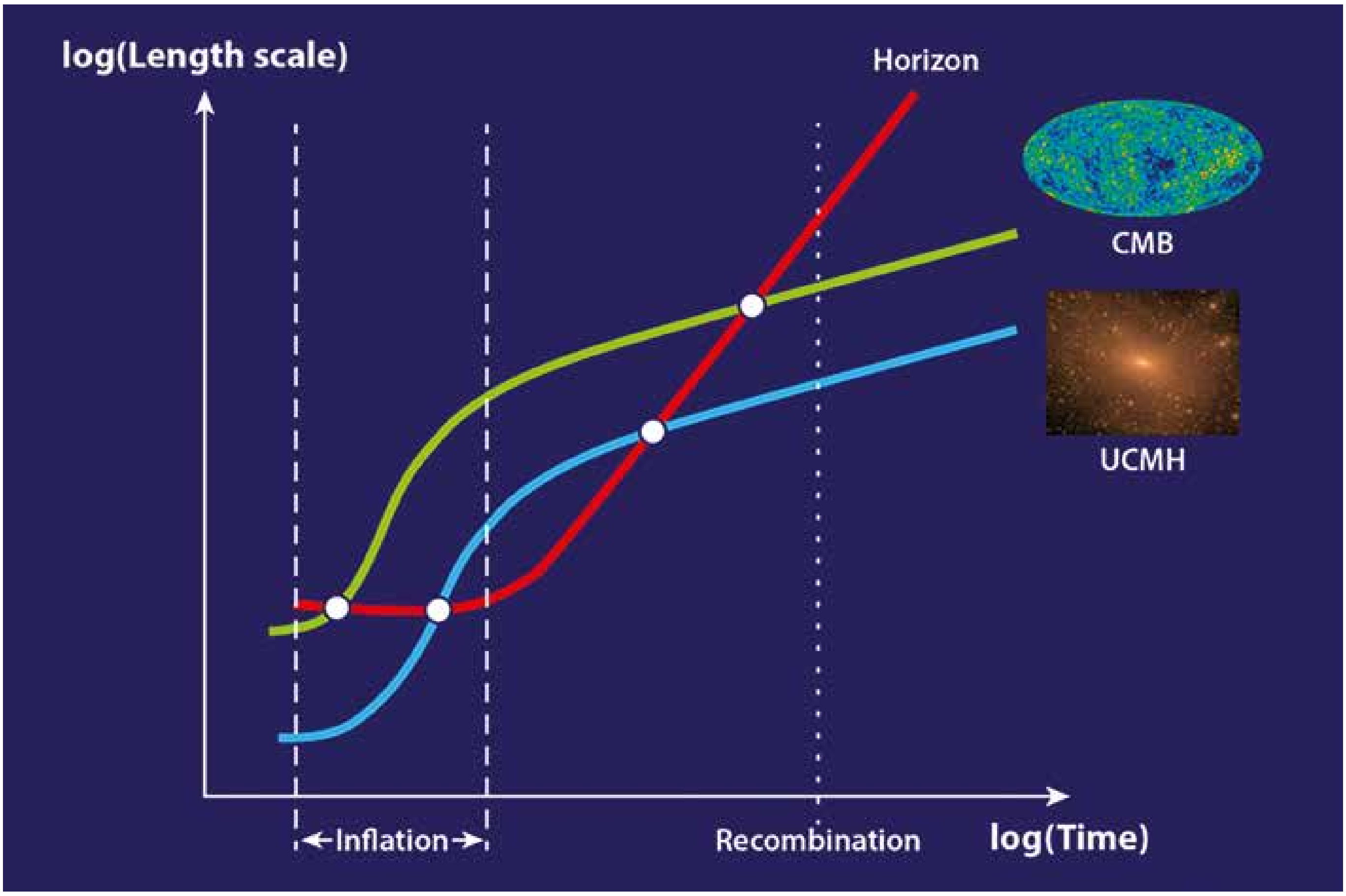


Figure 1. A schematic showing the length scale of two different perturbations: the green line corresponds to the cosmic microwave background (CMB), while the blue line represents ultracompact minihalos (UCMHs). During inflation, these scales expand exponentially, and at some point they surpass the causal horizon (shown in red), which means the perturbation's density becomes temporarily frozen in value. The CMB scales cross the horizon at an early period of inflation, whereas UCMH scales access a much later period of inflation.

Case study 20 | NeSI High Performance Computing platform

USING GPUS TO EXPAND OUR UNDERSTANDING OF THE SOLAR SYSTEM

Dr Philip Sharp, Department of Mathematics.

Using High Performance Computing (HPC) to study the early evolution of the Solar System

Dr Sharp is collaborating with William Newman, a Professor in Physics & Astronomy at UCLA, to investigate whether the Nice model, the leading model for the Solar System's evolution, accurately predicts the present day orbits of the giant planets, Jupiter, Saturn, Uranus and Neptune. In addition to the giant planets, the Nice model also includes the Sun and N-5 small bodies called planetesimals (see Figure 1).

"Put simply, the work would not be possible without HPC," Sharp says. "We chose NeSI because its HPC resources are the best in the country."

As part of his investigation, Sharp uses mathematical methods called N-body simulations. His latest simulation program, running on a single Graphics Processing Unit (GPU), has performed simulations of 20 to 100 million years with N=1024 astronomical bodies. Although this value of N has been used by other research groups, Sharp and Newman are the first to perform accurate simulations with this value. Results of that work were published in the Journal of Computational Science.

"Collisionless N-body simulations over tens of millions of years are an important tool in understanding the early evolution of planetary systems," says Sharp. "Our method is significantly more accurate than symplectic methods and sufficiently fast."

So far, results from these simulations show the Nice model is unstable for the number of bodies researchers use with the model. This instability therefore limits its predictive power. To strengthen this argument, Sharp and his colleagues wanted to perform accurate simulations with larger values of N, however they were limited by their existing simulation program because the elapsed time would be too large.

The help of NeSI' s Scientific Programmer

This is where Chris Scott from NeSI's Computational Science Team stepped in to help. Although Sharp's simulation program was already well optimised, further gains could be made by modifying the program to run on multiple GPUs at once. With Scott's help, the researchers were able to convert their simulation program to make use of all available GPUs within a node from a single CPU core. They applied this to a range of

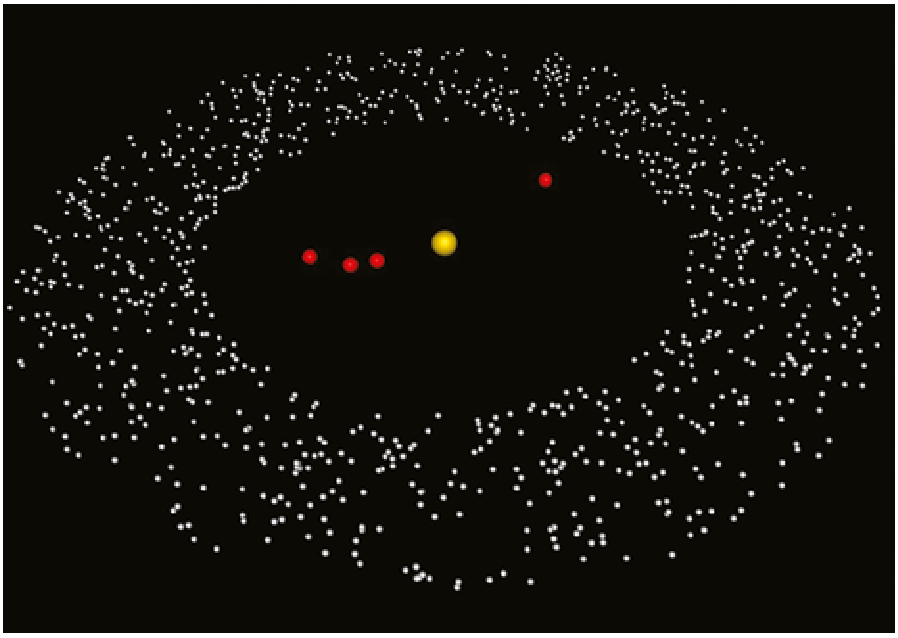


Figure 1. Initial configuration of a simulation (not to scale), showing the Sun (gold), large planets (red) and planetesimals (silver).

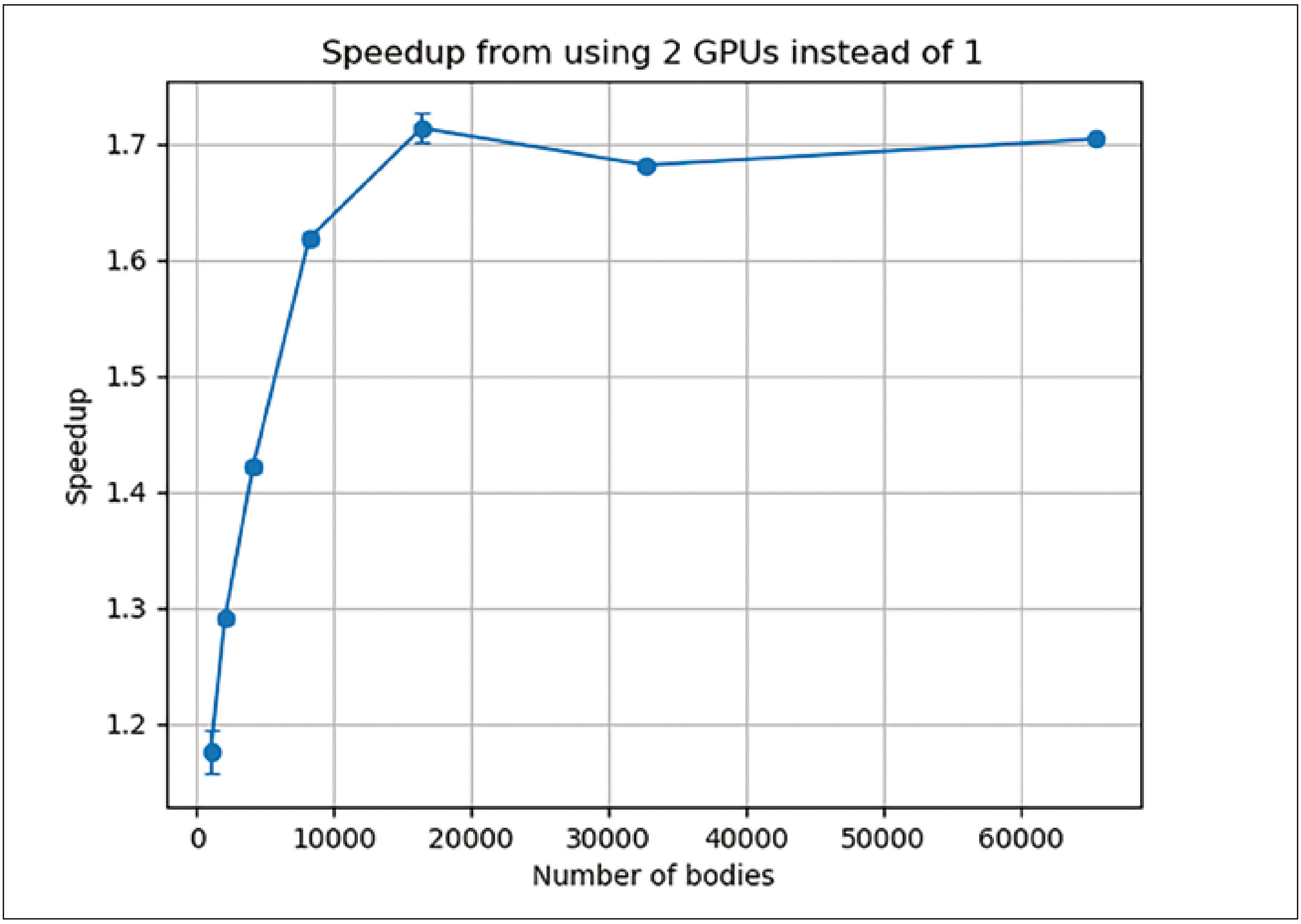


Figure 2. The speedups obtained by running on two GPUs compared to one.
The initial increase is due to the GPUs not being fully utilised at lower numbers of bodies.

numbers of bodies, from 1,024 to 65,532, and immediately observed a modest improvement in performance when working with the lower numbers of bodies. As they switched to larger numbers of bodies, which are of more interest to Sharp, the gains were even greater, with a 1.7x speedup observed for simulations with 16,384 or more bodies. Scott then experimented with introducing OpenMP, an Application Programming Interface (API) commonly used for parallel programming, and the combined speedups from using two GPUs and OpenMP were 1.9x for 4,096 bodies, 2.0x for 16,384 bodies, and 1.8x for 32,768 bodies. As a result of these changes, the simulation program is now running up to two times faster than previously," says Scott. "Also, these gains will allow researchers to run simulations of larger systems in a more reasonable time and verify the results obtained from their calculations."

"Tackling these types of research questions relies heavily on access to High Performance Computing," says Sharp.

As they continue their work in this area, the insights uncovered by Sharp and his colleagues will further strengthen the Nice model as a tool for explaining how the Solar System is evolving.

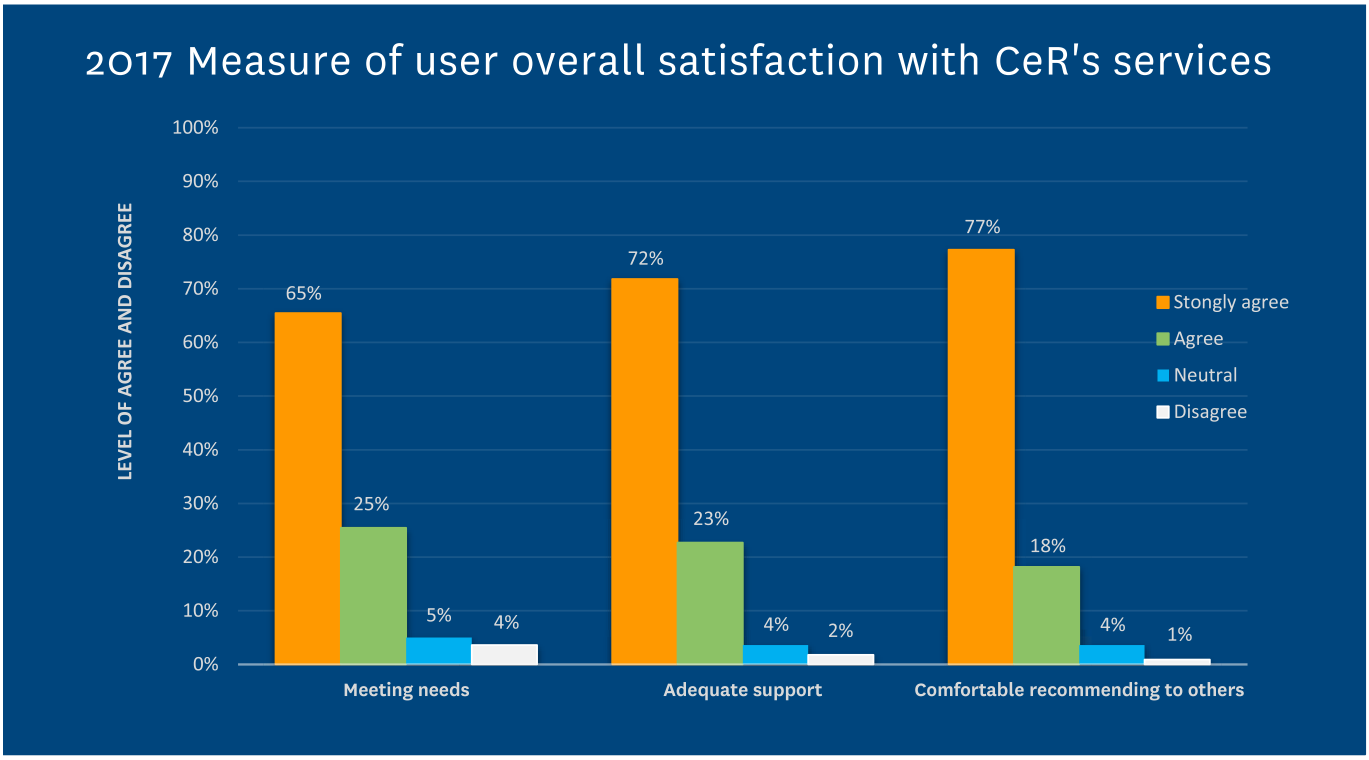
"Our results clearly signpost what to do next - perform simulations with a lot more bodies," Sharp says. "We expect, as happens with galactic simulations, that increasing the number of bodies will make the Nice model more stable and hence improve its predictive power."

RESEARCHER SURVEY

We have conducted a survey in November 2017 for more than 160 researchers who have used research computing (including Nectar but excluding HPC), data management and visualisation & analytics services that the Centre provides. Of which, there were 69% response rate. The following graph demonstrates the measure of user satisfaction of using these services and some selected feedback and constructive suggestions from researchers about these services.

Overall, 94% (rounding) average satisfaction in the following areas: 1) if these services meeting their needs, 2) provided adequate support, and 3) if they comfortable recommending to others.

Measure of satisfaction



Feedback - research virtual machines / research data management

The VM for this project was set up to allow us to have a central storage and processing repository for the MRI data collected as part of the larger DPRC project. This allows multiple researchers to access the data from different labs, and in the future be able to process the data in a standardised way in a Linux environment, which is the standard OS for imaging software. It has also allowed external collaborators at Otago University to access the data easily. We are also collaborating with a researcher in the UK and the VM was an ideal way to view the data collected in Auckland. It also saved considerable set-up time by providing a Linux environment with all the software I needed installed to process the data. I would like to add that the support in terms of setting up the VM and installing the software we would like has been fantastic. *– Dr Catherine Morgan, Research Fellow, Optometry & Vision Science.*

After I initiated this project with CeR in 2015, they managed to get ChemStation running in a Win 10 environment, which was not an easy task. By 2016, the VM had really sped up our ability to process omics data. I don't spend all my time installing software and troubleshooting, because we have stable versions of all of our software on the VM. It allows my collaborators at AUT and North Shore Hospital to log in and process their own data, which wasn't possible before, and has saved me having to book a computer and sit with them. It's also great when

I have short-term students who don't have their own computer desk allocated. I honestly don't know what I would do without it now. *– Dr Elizabeth McKenzie, Research Fellow, Liggins Institute.*

Currently, anyone within the UoA can access MPAI (http://mpai-speak.uoa.auckland.ac.nz/). However, we'd really like to make it accessible to people outside UoA, so we can get a greater number of users, and therefore feedback. So, I would like to work with the Centre of eResearch and see what I need to do security wise so that we can have access outside the University. – A/Prof Catherine Watson, Electrical and Computer Engineering.

The VM has made a massive contribution to my work, and to what is possible both for myself and for my research postgraduate students. This service is invaluable and will become increasingly useful over time - particularly in terms of verifiable research outputs. *– Dr Kane Meissel, Lecturer, Education Psychology.*

This VM machine has given me the flexibility that I did not have before. Previously, I could not do the same work on the desktop machine, and had to depend on NeSI all the time. It was not efficient at all. I'm grateful for the support and attention that CeR provides. *– Tharanga Jayathungage Don, Doctoral Candidate, Engineering Science.*

Being able to put future software on this powerful computing platform would be very helpful. *– Margaret Coe, Senior Research Technician, Liggins Institute.*

I prefer the flexibility of the more recent Nectar infrastructure. *– Dr John Walter Rugis, HPC Scientific Programmer, Mathematics and Auckland Bioengineering Institute.*

The availability of VM capacity has been very helpful, as it has provided a flexible and highly portable work environment that can be used, even while travelling internationally, to provide additional processing power. It has been useful to share 'live' work environments with students, and to work continuously on a project from multiple machines. *– Dr Jon Francis Tunnicliffe, lecturer, Environment.*

The research VM was very easy to use, had everything we needed, and was obviously most vital to our project! I'm actually not sure how we could've done the project without it, since usually requesting a VM directly from the University takes way too long in relation to the project's time frame. So yeah, I'd recommend it to other people working on similar projects. *– Prof Robert Amor, HoD, Computer Science.*

I really appreciate the service provided to others and me by the Centre for eResearch. It really makes a difference to be able to use this software, and without CeR we would not be able to run and manage it. I appreciate that everyone I have talked to has been patient enough to explain things to me as well. To further improve the service, it would be great to see the status somehow, including the storage space as well as an error. I am not sure how we are tracking with storage space but we will need continued support in this area. We also discussed automated transfer of data onto the virtual machine and it would be great to set this up. *– Erica Burns, Doctoral Candidate, Biological Sciences.*

We are planning on expanding our collaborative work with other universities offering varying levels of security to the research datasets (depending on the sensitivity of the content) and are hoping to be able to manage this through the VM systems you provide. *– Peter Tricker, Data and Systems Manager, Growing up in New Zealand, UniServices.*

The research VM has allowed collaborators in the UK to access our study data, which our research agreements and ethical approvals dictated could not leave the UoA. Formal outputs are currently being developed, but after receiving the results of the UK group's work a couple of days ago, we are confident they will have a major impact on our field of medicine as well as in the wider methodological field behind this work. *– Dr Katrina Poppe, Senior Research Fellow, Medicine.*

The support levels are excellent, and have enabled us to resolve problems in a timely manner. It also has provided system admin so that new software can be properly installed for all users. We are very reliant on the eResearch facilities to progress our bioinformatics/genomics research, and are grateful for the facilities and level of support they have provided. In terms of future needs, double the RAM would be useful, and we expect to have significantly more storage requirements in the coming 12-24 months. *– Dr Austen Ganley, Senior Lecturer, Biological Sciences.*

This project received considerable media attention. Please let me know if you'd like a list of media coverage, as it is directly linked to the existence of this web tool. *– Dr Kiti Suomalainen, Research Fellow, Energy Centre.*

You solved the big need — a server. The second-biggest need is competence for installing and configuring the software (because while I can learn, there doesn't appear to be a good learning pathway and I don't have the continual engagement to maintain fluency), and I still haven't raised funds to do the customisation that I need. *– Dr Alistair Kwan, Lecturer, Centre for Learning and Research in Higher Education.*

I have found CeR to be extremely helpful in all interactions I have had with the team. I cannot think of much else they can do to help us our research. Thank you! *– Meg Spriggs, Doctoral Candidate, Psychology.*

Support from CeR and NeSI staff has been fantastic, un-bureaucratic and lightening-fast. This is a clear point of difference (to other platforms) and very important for our research. Many thanks for all the great assistance. *– A/Prof Klaus Lehnert, Biological Sciences.*

The Centre for eResearch offers a fantastic service. I have recommended it to several other researchers from different faculties. *– A/Prof Beryl Plimmer, Computer Science.*

Very useful service. I find copying large files starts off fast, but then becomes very slow; is this the caching layered file storage. *– A/Prof Andrew Mason, Engineering Science.*

You are always very helpful and extremely responsive to queries. We're very happy with the support we get from you. *– Dr Anna Santure, Senior Lecturer, Biological Sciences.*

I would like to thank Sean Matheny who provided me with an exceptional support during my research project. *– Farzan Kolini, Doctoral Candidate, Information Systems and Operations Management.*

Thanks for providing a quality service! So easy to work with and reasonably quick to establish. You understood our needs and worked hard to ensure things worked. *– Dr Daniel Exeter, Senior Lecturer, Epidemiology and Biostatistics.*

Thank you for providing such a powerful platform to me. Although there's no research output so far via your services, I have got some fantastic result and will write a manuscript soon. Next year you will probably see some journal articles and conference reports coming out and I will surely acknowledge your contribution. *– Shuqi Wang, Doctoral Candidate, Biological Sciences.*

Everything is perfect. *– Dr Johannes Dimyadi, Research Fellow, Computer Science.*

Feedback - visualisation and analytics

This is an awesome product and the additional functions - the geocoding engine - was an awesome addition. The team understood my needs and were able to develop an on-the-fly app that could generate a deprivation-related report for different geographical scales. This was immensely practical and time-saving from having done the DHB-level reports manually for each of the 20 DHBs! I'd like to do some more modifications in terms of the geocoding engine etc. By having an open-source option to do this will be a real cost-saver! Thanks to Nick and Bianca and others for their patience and excellent work output! *– Dr Daniel Exeter, Senior Lecturer, Population Health.*

It is great to go to a place in the University where people really want to help, have great suggestions, are not condescending, know what they are doing, then do it. This particular project has benefited immensely from the involvement of the Centre for eResearch. The ability to show people the animation of the whales' movements over the course of their 8 1/2 months of tag transmission time, including the seaice extent moving as well, was an enormous 'aha' moment for anyone viewing the animation. I have used it in science talks, public presentations and it has been transmitted via social media (Twitter and Facebook) by me and NGOs working in this area. Nick Young has been helpful all the way through this work. *– A/Prof Rochelle Constantine, Biological Sciences.*

I use the visualisation tools to illustrate my work to people in the field and to the public. *– Dr Gilles Bellon, Senior Lecturer, Physics.*

The visualizations have been very helpful in demonstrating the potential of VR technology in its application to structural and molecular biology. They have inspired a number of biologists to get involved in the VR project and to write code to develop applications in this area. Additionally, this project provided a valuable proof of principle for further VR-Molecular biology projects currently underway. It would be awesome to try out some Augmented Reality tech adapted for use in the lab. e.g. Google glass with video recording and review of experiments, built in safety glasses, heads-up display for protocols, calculations etc. *– Kyle Webster, Doctoral Candidate, Biological Sciences.*

They were imaginative and helped show the data pattern in space and time well. At present, a graduate student is doing detailed statistical analysis of the data and when we are ready to write a paper, we will twin it with the visualisations somehow. *– A/Prof Mark John Costello, Institute of Marine Science.*

The visualisation services were used to create an animation of the 3D scans of the reef island simulated in the physical modelling experiments. *– Megan Tuck, Doctoral Candidate, Environment.*

The key focus of my project was to test user experience of seeing captions for recorded lectures in augmented reality (AR) environments. Without the assistance of Prof Mark Gahegan and Nick Young the study was not possible. Their advice and actual implementation of the critical application as well as use of the lab for user testing were invaluable. Furthermore, eResearch was able to introduce the resident UX expert Dr. Doris Jung to assist us with questionnaire and UX testing design. The study is still ongoing at the time of writing, and we aim to publish the results in a number of outlets during the course of 2018. The study allowed the PI to make oral presentations nationally and internationally, facilitating to promote and seek collaboration for the new lines of inquiry in Translation Studies research where AR is little explored. *– A/Prof Minako O'Hagan, European Languages and Literature.*

Used as part of an analysis on Transport issues and how they relate to the broader University Student Experience. We appreciated the collaborative approach and willingness to undertake this task at short notice. *– Tessa Jones, Campus Life.*

Suggestions - research virtual machines / research data management

We recently had issues raised about compliance with transferring data from DHB's, which has to meet some government specifications in relation to encryption etc. It needs to ensure systems continue to meet changing and increasing security requirements of other government partners. *– A/Prof John Thompson, Paediatrics, Child and Youth Health.*

Consecutive runs of my simulation on the VM can lead to run times that vary a lot (no other programs running, CPU and memory limits not reached). Usually by doubling, but this doesn't happen on my personal computer. It would be better if the run times were consistent, especially for reporting these times in the results of computational experiments. *– Samuel Ridler, Doctoral Candidate, Engineering Science.*

The data being analysed can sometimes be 10, 50, 100 GB, which is a bit slow to upload and CNA stall the programs. Any advancements to speed-up download times would be great. The lag in programs is likely a computing power issue that could be resolved. If it becomes frequent, I will contact CeR. *– Dr Nicholas Demarais, Research Fellow, Biological Sciences.*

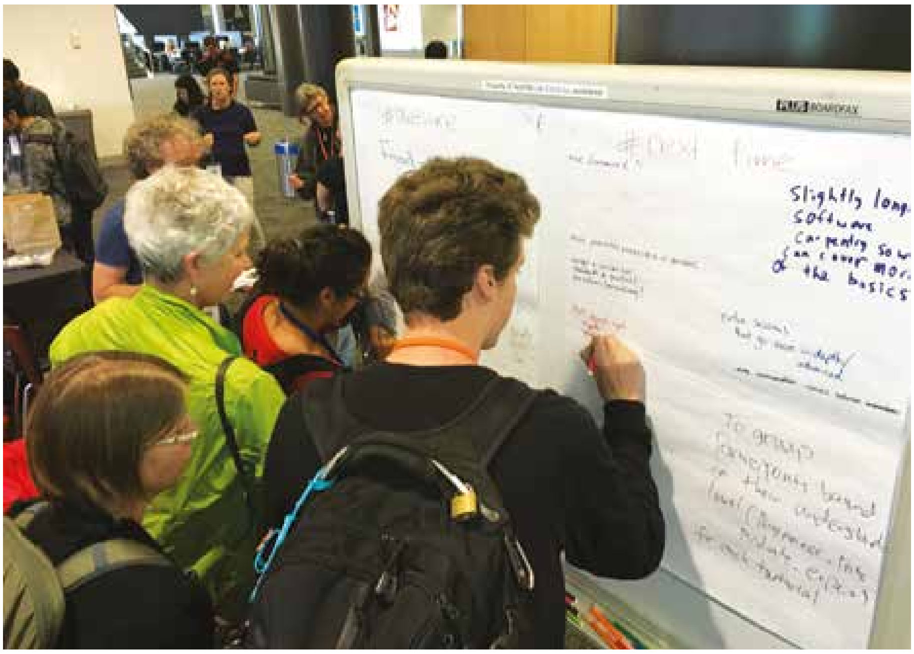
The only thing I feel inconvenient is data transfer. I wish I could use Globus to transfer my data between my own MacBook and your Virtual Machine but your platform doesn't allow such connection. Maybe I have an alternative way to transfer data with a large size. Please let me know if so. *– Shuqi Wang, Doctoral Candidate, Biological Sciences,*

Transferring data into and from the VM is still a challenging for me because of the amount of time it consumes. Sharing my data with researchers overseas has been also an additional limitation. *– Ivan Braga Campos, Doctoral Candidate, Biological Sciences.*

The memory is a problem with the VM and I experience considerable lag on many occasions, which could be improved. *– Teena Gamage, Doctoral Candidate, Medicine.*

Not to have parallel machines clogged with serial jobs. *– Dr Stuart Norris, Senior Lecturer, Mechanical Engineering.*

An increase in disk quota would be very helpful for the ongoing and expanding research in this project. *– A/Prof Klaus Lehnert, Biological Science.*

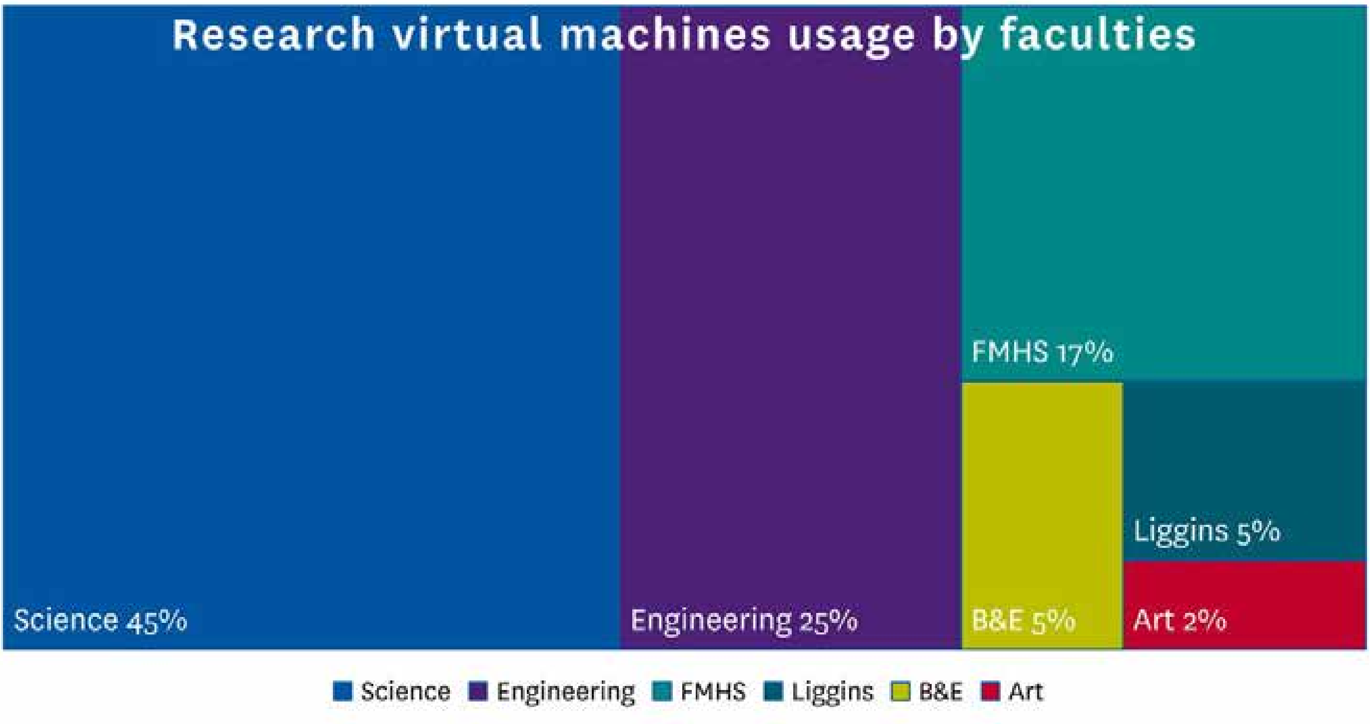


Summary

The following are the key reasons that the RVM service is helpful to researchers. In addition, the researchers consider the usefulness in terms of computational aspects: CPU 42%, Memory/RAM 40% and Disk Space 18%.

Main reasons that the research virtual machines are useful	% of total
It offered a more powerful computational environment for analyses/simulation than office/lab	32%
It enabled me to share data and work collaboratively with my students and colleagues	18%
I didn't need to buy a physical machine	16%
It freed up my desktop/laptop by offloading computations to the VM	14%
It provided a more reliable work environment than I have in my lab	11%
Other	9%
Total	100%

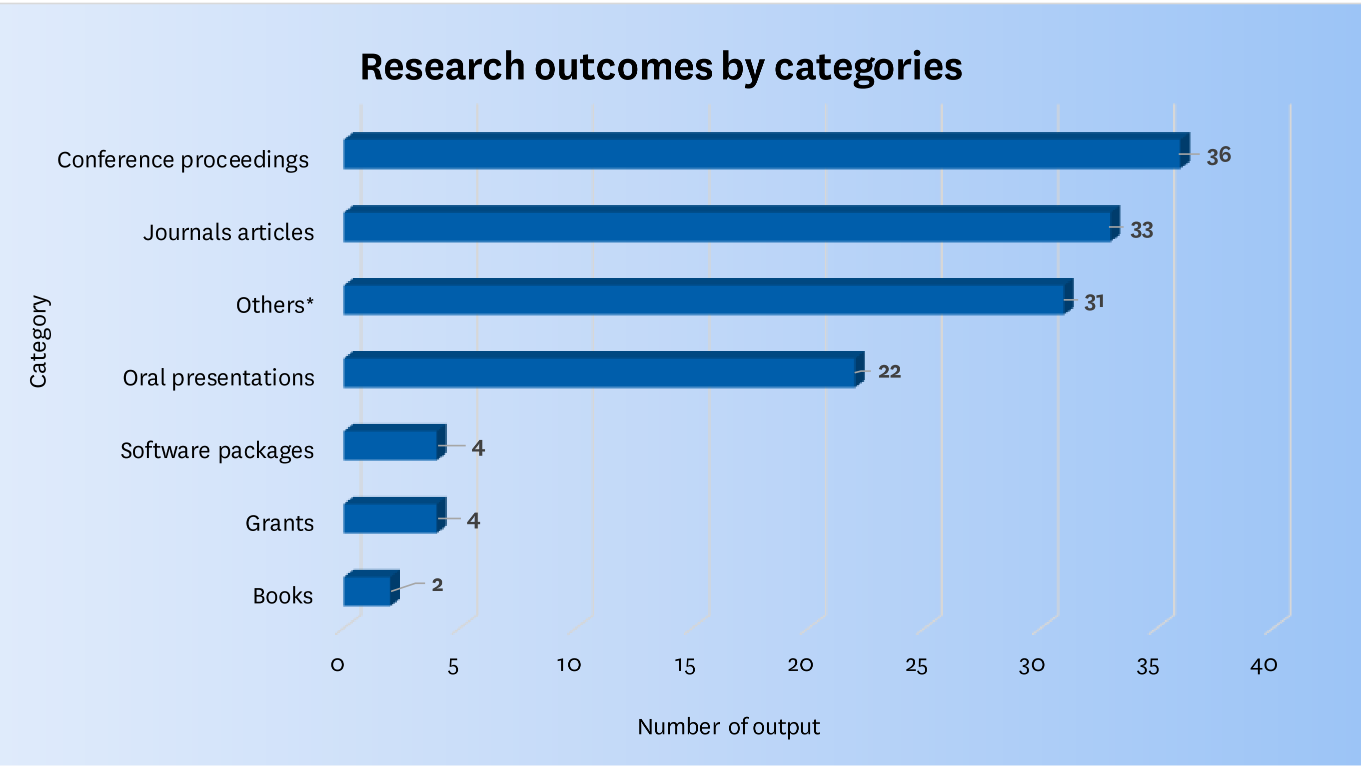
The graph below shows a breakdown by faculties of over 320 researchers to-date that are supported by the Centre for eResearch with the RVM service, of which 59% occurred in 2017.



Our RDM service has focused on securing storage space, assessing data for migration to a properly managed state and hosting it in an appropriate platform this year. In addition, the Figshare publishing service that collaborates with the Library has been in production since August 2017 and has reached over 76,000 views and 13,000 downloads. Altogether, there were 29 workshops conducted with over 350 researchers this year to begin an intensive process of education and engagement around research data and the importance of its planning. The following are some brief findings from researchers' perspective of how they prioritise the needs for "research data management".

Main reasons that research data management are helpful	% of total
It provided me with a safe place to store data	35%
It enabled me to collect and share with internal/external collaborators	35%
It helped me to better organise my data	15%
It made me think about copyright and intellectual property (IP)	5%
It enabled me to store instrument data	5%
Others	5%
Total	100%

The graph below shows a breakdown of a total over 130 research outcomes by categories that are supported by CeR's services in 2017. Please refer to the 'RESEARCH OUTCOME' section from page 66 for detailed lists.



** Others include abstracts, theses, posters, papers in preparation and technical reports etc.*

RESEARCH OUTCOMES

The following shows a summary of research outcomes (brief format) collected by using research computing platforms and/or supported by visualisation & analytics services provided by CeR. These outputs are based on the Centre's November survey, and only represent those who responded to the survey, so it is likely that some research outcomes are not listed.

Grants

1. **Dr Jessie Jacobsen**, "Genetic diagnosis of autism spectrum disorder," in the Minds for Min ds and Growing up in New Zealand cohorts, CureKids. 10 projects were funded across the four specific areas through National Science Challenge.
2. **Ivan Braga Campos**, "Listening to nature – A passive acoustic approach for protected areas monitoring and biodiversity assessment," in Birds NZ Research Fund (BNZRF).
3. **Dr Christopher Walker**, "Naturally biased? Exploring neuropeptide signal pathway bias in pain." in Health Research Council of New Zealand, Sir Charles Hercus Health research Fellowship.
4. **Christopher Walker**, "Exploring Neuropeptide Signal Pathway Bias in Pain," in Royal Society of New Zealand, Marsden Fast-Start Grant.

Journal articles

1. **Fliegauf, M., Bryant, V. L., Frede, N., Slade, C., Woon, S. T., Lehnert, K., Grimbacher, B.**, "Haploinsufficiency of The NF-KB1 Subunit P50 in Common Variable Immunodeficiency," in American Journal of Human Genetics, 2017 97(3), 389-403.
2. **Holdaway, S. J., Phillipps, R., Emmitt, J., Wendrich, W.**, "The fayum revisited: reconsidering the role of the Neolithic Package, Fayum North Shore, Egyp," in Quaternary International, 410A, 173-180.
3. **Chua, E. W., Cree, S. L., Ton, K. N. T., Lehnert, K., Shepherd, P., Helsby, N., Kennedy, M. A.**, "Cross-comparison of exome analysis, next-generation sequencing of amplicons, and the iPLEX(®) ADME PGx panel for pharmacogenomic profiling," in Frontiers In Pharmacology, 7, 1.
4. **McKay, N. S., Iwabuchi, S. J., Häberling, I. S., Corballis, M. C., Kirk, I. J.**, "Atypical white matter microstructure in left-handed individuals," in Laterality, 22(3), 257-267.
5. **Phillipps, R., Holdaway, S., Ramsay, R., Emmitt, J. J., Wendrich, W., Linseele, V.**, "Lake level changes, lake edge basins and the Paleoenvironment of the Fayum North Shore, Egypt, during the Early to Mid-Holocene," in Open Quaternary, 2, 12 pages.
6. **Phillipps, R., Holdaway, S., Emmitt, J., Wendrich, W.**, "Variability In the neolithic settlement patterns of the egyptian nile valley," in African Archaeological Review, 33(3), 277-295.
7. **Thompson, J. M., Stone, P. R., Sanders, M., van der Zee, H., Borman, B., Fowler, P. V.**, "The incidence of orofacial cleft in live births in new zealand," in New Zealand Medical Journal, 129(1440), 64-71.
8. **Geraerts, F. C. A., Snell, R. G., Faull, R. L. M., Williams, L., Jacobsen, J. C., Reid, S. J.**, "Comparison of huntington's disease cag repeat length stability in human motor cortex and cingulate gyrus," in Journal of Huntington's Disease, 5(3), 297-301.
9. **Zhang, Y., Munir, M. T., Yu, W., Young, B. R.**, "Modelling batch bioreactions with continuous process simulators," Korean Journal of Chemical Engineering, 33(12), 3343-3349.
10. **Dimyadi, J. A., Pauwels, P., Amor, R.**, "Modelling and accessing regulatory knowledge for computer-assisted compliance audit," in Journal of Information Technology in Construction, 21, 317-336.
11. **Nava, S., Cater, J. E., Norris, S. E.**, "Modelling leading edge separation on a flat plate and yacht sails using LES," International Journal of Heat and Fluid Flow, 65, 299-308.
12. **Emmitt, J. J., Sefton, B., Phillipps, R., Wendrich, W., Holdaway, S.**, "Reimag(in)ing the past: adding the third dimension to archaeological section drawings," in Advances in Archaeological Practice, 5(1), 44-53.
13. **Wendrich, W., Holdaway, S. J.**, "Basket use, raw materials and arguments on Early and Middle Holocene mobility in the Fayum Egypt," in Quaternary International, 1-10.

14. **Whitford, W., Hawkins, I., Glamuzina, E., Wilson, F., Marshall, A., Ashton, F. Jacobsen, J. C.,** "Compound heterozygous SLC19A3 mutations further refine the critical promoter region for biotin-thiamine-responsive basal ganglia disease," in Cold Spring Harbor molecular case studies.

15. **Bertin, S., Groom, J., Friedrich, H.,** "Isolating roughness scales of gravel-bed patches," in Water Resources Research, 53(8), 6841-6856. 2017.

16. **Wu, C. C., Hamm, J. P., Lim, V. K., Kirk, I. J.,** "Musical training increases functional connectivity, but does not enhance mu suppression," in Neuropsychologia, 104, 223-233.

17. **Grey, A. C., Demarais, N. J., West, B. J., Donaldson, P. J.,** "A quantitative map of glutathione in the aging human lens," in International Journal of Mass Spectrometry.

18. **Costello, M. J., Tsai, P., Wong, P. S., Cheung, A. K. L., Basher, Z., Chaudhary, C.,** "Marine biogeographic realms and species endemicity," in Nature communications, 8(1), 1057.

19. **Kwan, A. M.,** "Re-Defining 'evidence': appraising for historical value as historians turn to media and materiality," in Archifacts, (April-October), 86-94.

20. **F. Hafiz, A. Swain, N. Patel, and C. Naik,** "A two-dimensional 2-dlearning framework for particle swarm based feature selection," in Pattern Recognition, Elsevier Vol 76, Apr 2018, 416-433.

21. **F. Hafiz, A. Swain, N. Patel, and C. Naik,** "Efficient feature selection of power quality events using two dimensional (2D) particleswarms," in IET Generation, Transmission & Distribution" (Under Review).

22. **Fowler PV, Azza AH, Thompson JMD.,** "Comparison of reliability of categorical and continuous scales for radiographic assessments of bone infill following secondary alveolar bone grafting," in The Cleft Palate-Craniofacial Journal (in press).

23. **Riekkola, L. et al.,** "Large-scale movements and life history characteristics of a highly mobile long-distance migrant,' Manuscript in review.

24. **Ameratunga, R., Lehnert, K., Woon, S. T., Gillis, D., Bryant, V. L., Slade, C. A., Steele, R.,** "Review: Diagnosing common variable immunodeficiency disorder in the era of genome sequencing," in Clinical Reviews in Allergy and Immunology, 1-8.

25. **Zhang, Y., Saleman, T. L. H., Li, G. K., Xiao, G., Young, B. R., May, E. F.,** "Non-isothermal numerical simulations of dual reflux pressure swing adsorption cycles for separating N2 + CH4," in Chemical Engineering Journal, 292, 366-381.

26. **May, E. F., Zhang, Y., Saleman, T. L. H., Xiao, G., Li, G. K., Young, B. R.,** "Demonstration and optimisation of the four dual-reflux pressure swing adsorption configurations," in Separation and Purification Technology, 177, 161-175.

27. **Davies, M., Kiemer, K., Meissel, K.,** Quality talk and dialogic teaching- an examination of a professional development programme on secondary teachers' Facilitation Of Student Talk. In British Educational Research Journal, 43(5), 968-987.

28. **Phillipps, R., Holdaway, S. J., Ramsay, R., Wendrich, W., & Emmitt, J. J.,** "Approaches to paleoenvironment and landscape use. The desert fayum reinvestigated: the early to mid-holocene landscape archaeology of the Fayum North Shore, Egypt," in Los Angeles: Cotsen Institute of Archaeology Press.(pp. 17-50). 2017.

29. **Bellon, G., Bony, S.,** "Tropical and subtropical cloud systems," in Clouds and Climate. Cambridge University Press. 2017.

30. **Brewer, B.,** "Computing entropies with nested sampling," in Entropy, 19(8), 422.

31. **De Seymour, J., McKenzie, E. J., Baker, P.,** "Using the Food metabolome to understand the relationship between maternal diet and gestational diabetes.," R. Rajendram, V. R. Preedy, V. B. Patel (Eds.), in Nutrition and Diet in Maternal Diabetes An Evidence-Based Approach (pp. 263-274). Cham: Humana Press.

32. **Holdaway, S., Phillipps, R. S., Koopman, V., Linseele, V., & Wendrich, W.,** "The L Basin archaeological record. The desert fayum revisited: the eartly to mid-holocene landscape archaeology of the Fayum North Shore, Egypt. in Los Angeles: Cotsen Institute of Archaeology Press. (pending verification)

33. **Wendrich, W., Holdaway, S., Phillipps, R. S., & Emmitt, J.,** "The k basin archaeological record. The desert fayum revisited: the eartly to mid-holocene landscape archaeology of the Fayum North Shore, Egypt," inLos Angeles: Cotsen Institute of Archaeology Press. 2017 (pending verification)

Conference proceedings

1. **Lu, H., Geertsema, P.,** Is Real Earnings Management Real?. In 2017 New Zealand Finance Colloquium. Auckland, New Zealand.

2. **Arriola Ortiz, J., Tripathi, A. K.,** "Audience acquisition in online resource mobilization: quantifying the relationship with influential actors," In PACIS 2017 Proceedings. Langkawi, Malaysia.

3. **Bretscher, E., Norris, S. E., Mason, A. J., MacFarlane, G. J., Denier, J. P.,** "Parametric search and optimisation of fast displacement hull forms using RANS simulations of full scale flow," In M. Visonneau, P. Queutey. D. Le Touzé (Eds.), MARINE 2017 Computational Methods In Marine Engineering VII (pp. 515-528), Nantes, France: International Center for Numerical Methods in Engineering (CIMNE).

4. **Nava, S., Cater, J. E., Norris, S. E.,** "Large eddy simulation of downwind sailing," In INNOV'SAIL 2017 International Conference On Innovation In High Performance Sailing Yachts: 4th EDITION. Lorient, France.

5. **McKenzie, E. J., McKenzie, N. A. L., Pook, C. J.,** "Long-term storage of human skin surface samples for metabolomics," in The 26th Biannual Conference of the Australia and New Zealand Society for Mass Spectrometry (pp. In Press). Adelaide, Australia.

6. **Watson, C. I., Keegan, P., Maclagan, M. A., Harlow, R., King, J.,** "The motivation and development of MPAi, a Māori pronunciation aid," Stockholm, Sweden: International Speech Communication Association. (pp. 2063-2067).

7. **Dimyadi, J. A., Governatori, G., Amor, R.,** "Evaluating legaldocml and legalruleml as a standard for sharing normative information In The AEC/FM Domain," in Lean Computing in Construction Congress (LC3 2017) Vol. 1 (pp. 639-646). Heraklion, Crete, Greece.

8. **Dimyadi, J. A., Amor, R.,** "Automating conventional Compliance audit processes," The 14th International Conference on Product Lifecycle Management. 2017 Seville, Spain.

9. **Watson, P., Rubie-Davies, C. M., Meissel, K., Flint, A., Garrett, L., McDonald, L.,** "Mathematics self-concept in new zealand primary school students: intervening to lower a barrier?" 17th European Association for Research on Learning and Instruction (EARLI 2017). Tampere, Finland.

10. **Dimyadi, J. A., Amor, R., Thomas, G.,** "Enabling automated compliance audit of architectural designs," in M. Aurel (Ed.), The 51th International Conference of the Architectural Science Association 2017. Wellington, New Zealand.

11. **Blumenstein, M., McDonald, J., Gunn, C., Leichtweis, S.,** "Let's talk student engagement: connecting teachers and students using learning analytics," 17th Biennial EARLI Conference. Tampere, Finland.

12. **Tunnicliffe, J. F., Leenman, A., Reeve, M.,** "The impacts of landsliding on a gravel bed river network: the remarkable case of new zealand's east cape," in Gravel Bed Rivers 8: Gravel Bed Rivers And Disasters. Available online Jan 2017.

13. **Tunnicliffe, J. F.,** "A sediment budget framework to support river management in Tairāwhiti," in Tairāwhiti Freshwater Conference 2017. Hauiti Marae, Tolaga Bay.

14. **Tunnicliffe, J. F.,** "Mechanisms of channel and floodplain evolution," in The Tapuaeroa River, East Cape, NZ. The 17th Australian and New Zealand Geomorphology Group Conference. Greytown, New Zealand.

15. **Bayati, S., Tripathi, A. K.,** "Analysis of open source software projects popularity among developers.," in R. Santanam, V. Yoon (Eds.), Proceedings of Twenty Seventh Annual Workshop on Information Technology and Systems (WITS). Seoul.

16. **Hu, T., Tripathi, A. K.,** "Is there a free lunch? Examining the value of free content on equity review platforms," M. Shaw, B. Yoo (Eds.), in Proceedings of Sixteenth Workshop on E-Business (Web). Seoul.

17. **F. Hafiz, S. Abecrombie, A. Eaton, C. Naik and A. Swain,** "Power quality event identification using wavelet packet transform: a comprehensive investigation," TENCON 2017, IEEE Region 10 Conference.

18. **F. Hafiz, A. Swain, N. Patel, and C. Naik,** "Feature selection for power quality event identification," TENCON 2017, IEEE Region 10 Conference.

19. **O'Sullivan JM, Jayasinghe T, Chiavaroli V, Jayan S, Ekblad C, Derraik J, Hofman P, McKenzie E.,** "The gut microbiome and insulin resistance in children born preterm," in Cutfield WS Developmental Origins of Health and Disease Research, Tokyo, 2016. Public record Jul 2017.

20. **Campos, I.B., Landers, T.J., Lee, K., Lee, W.G., Friesen, M.R., Gaskett, A.C., Ranjard, L,** "preliminary results of an automatic identification of seabird calls from multispecies field recordings," in New Zealand Bird Conference 2017, Te Anau, New Zealand.

21. **Landers, T.J., Brendon, D.J., Campos, I.B., Matt J. Rayner, M.J., Ranjard, L.,** "The challenges of seabird monitoring by local government and what auckland council is doing about it," in New Zealand Bird Conference 2017, Te Anau, New Zealand.

22. **Campos, I.B., Ranjard, L., Lee, W.G., Gaskett, A.C.,** "Soundscape and acoustic analyses for biodiversity monitoring in protected areas," in International Congress for Conservation Biology - ICCB 2017, Cartagena, Colombia.

23. **Kolini, Farzan and Janczewski, Lech,** "clustering and topic modelling: a new approach for analysis of national cyber security strategies," in PACIS 2017 Proceedings. 126.

24. **Steven Dakin; Samuel Schwarzkopf; Geraint Rees; Catherine Morgan; Elaine Anderson,** "Cortical magnification factor of human V2 predicts individual susceptibility to letter-crowding," in Journal of Vision Aug 2017, Vol.17, 396. doi:10.1167/17.10.396

25. **Gamage T, Schierding W, Tsai P, Ludgate J, Macaulay E, Weeks R, Chamley L, and James J.,** "Global methylation profiling reveals intragenic methylation as a novel regulator of human trophoblast differentiation," in Placenta 2017: 57; 261.

26. **Gamage T, Schierding W, Hurley D, Tsai P, Ludgate J, Macaulay E, Weeks R, Chamely L, and James J.,** "Gene methylation regulates the acquisition of an invasive phenotype during extravillous trophoblast differentiation," in Placenta: 57; 304.

27. **Bayati, S., & Tripathi, A. K.,** "Analysis of open source software projects popularity among developers," R. Santanam, V. Yoon (Eds.) in Proceedings of Twenty Seventh Annual Workshop on Information Technology and Systems (WITS) Seoul. Dec. 2017.

28. **Hu, T., & Tripathi, A. K.** Is There A Free Lunch? Examining The Value Of Free Content On Equity Review Platforms. M. Shaw, B. Yoo (Eds.) Proceedings of Sixteenth Workshop on E-Business. Web 2017 Seoul.

29. **Shahbaznezhad, H., Dolan, R., & Tripathi, A. K.,** The Power Of Facebook And Instagram Fans: An Exploration Of Fan Comments And Their Effect On Social Media Content Strategy. M. Shaw, B. Yoo (Eds.) Proceedings of Sixteenth Workshop on E-Business.Web 2017 Seoul.

30. **Shahbaznezhad, H., & Tripathi, A. K.,** "Studying the effective factors of user engagement on social media fan pages," M. Shaw, B. Yoo (Eds.) in Proceedings of Sixteenth Workshop on E-Business. Web 2017 Seoul.

31. **Virues-Ortega, J., Lehnert, K., Swan, B., Taylor, M. W., Southee, A., Dougan, D. Jacobsen, J. C.,** "Research in Autism Spectrum Disorders," in The New Zealand Minds For Minds Autism Spectrum Disorder Self-Reported Cohort 36, 1-7.

32. Zhou, L., Swain, A., Naik, P. K., Hafiz, F., Nair, N. K. C., "Optimum design of series compensation to improve power system transient stability using DFQL and PSO Based controllers," in IEEE International Conference on Control and Automation, ICCA 2017 (pp. 1125-1130).

33. **Ganley, A. R. D., Chudakova, D., Fitch, J., Hermann-Ledenmat, S., Sharma, D., Diesch, J.,Hannan, R.** "The ribosomal rna gene repeats: a new paradigm for cancer research and therapy," in the 4th International Conference on Advanced Molecular Bioscience and Biomedical Engineering (ICAMBBE) 2017. Indonesia.

34. **Plimmer, B., Brown, B., Diprose, J., Du Preez, S., Luxton-Reilly, A.** "Tabletop 3D object manipulation with touch and tangibles," in Collaboration Meets Interactive Spaces (pp. 11-32).

35. **Willmott, G., Hammerschmidt, L., Grimson, M., Riddings, K., Hendy, S., Wills, P., Demarais, N.,** "Physics at the Interface," in AAPPS Bulletin (Vol. 27, Iss. No. 5, pp. 8-14). Association of Asia Pacific Physical Societies.

36. **Putland, R. L., Ranjard, L., Constantine, R., Radford, C. A.,** "A hidden Markov model approach to indicate Bryde's whale acoustics," in Ecological Indicators, 84, 479-487.

Oral presentations

1. **K. Suomalainen, V. Wang (presenter), B. Sharp,** "Estimation of solar potential in urban areas using lidar and GIS," Aug 2017. in ESRI User Conference, Auckland, New Zealand

2. **Hanh Nguyen, Margot Bador, Christophe Menkes, Alexandre Peltier, Gilles Bellon,** "Climate of the tropical Pacific Islands," in The Australian Meteorological and Oceanographical Society and The Meteorological Society of New Zealand. (AMOS-ICSHMO) Feb. 2018. Sydney Australia.

3. **Andrew Mason,** ORSNZ Annual Conference (with the NZSA / IASC-ARS conference) Dec 2017.

4. **Kate Lee, Patricia Brekke, John Ewen, Craig Miller and Anna Santure.** "Identifying snps with low coverage whole genome sequencing:

trials with Hihi genomes," in New Zealand Molecular Ecology Conference. Dec 2016.

5. **Kate Lee.** Finding Variation In Hihi Dna - The Next Step In Population Management. Miranda, 2017.

6. **Kate Lee, Patricia Brekke, John Ewen, Craig Miller and Anna Santure.** Conservation Management In The Genomics Era. A case study with Notiomystiscincta(Hihi), Hihi Recovery Group meeting talk 2017.

7. **Kate Lee.** Using Hihi Genomics In Reintroduction Biology, June 2016, Department seminar.

8. **Kate Lee, Patricia Brekke, John Ewen, Craig Miller and Anna Santure.** "Genomic architecture of hihi (stichbird) morphological traits," July 2017 Department seminar.

9. **Kate Lee.** "Hihi on Tiritiri Matangi", Public workshop for the Supporters of Tiritiri Matangi, Sept, Oct 2017.

10. **Kate Lee, Patricia Brekke, John Ewen, Craig Miller and Anna Santure.** "Looking at genomics of a wild population of endangered birds, a nz case study with Hihi (Stichbird)," in New Zealand Molecular Ecology Conference. Dec 2017.

11. **Debbie Hay,** "Towards novel amylin agonists for diabetes and obesity," in Asia Pacific Diabetes and Obesity symposium, Hong Kong, China October 2017.

12. **Thompson JMD, Ayrey L, Stone PR, Fowler P.,** "New Zealand National Cleft outcomes study: Overview and progress," in 13th International Congress of the of Cleft Lip and Palate and Related Craniofacial Anomalies. Feb 2017, Chennai, India.

13. **Thompson JMD, Ayrey L, Stone PR, Fowler P.,** "Reliability of the asher-mcdade and a visual analogue scale in assessing nasolabial outcomes," in 13th International Congress of the of Cleft Lip and Palate and Related Craniofacial Anomalies. Feb 2017, Chennai, India.

14. **Thompson JMD, Ayrey L, Stone PR, Fowler P.,** "Comparison of nasolabial outcomes In New Zealand," in a 10 Year Cleft Cohort. 13th International Congress of the of Cleft Lip and Palate and Related Craniofacial Anomalies. Feb 2017, Chennai, India.

15. **Thompson JMD, Ayrey L, Fowler PV.,** "Perception of surgical repair compared to quality of life," Australasian Cleft Lip and Palate Association annual meeting, May 2017, Brisbane, Australia.

16. **Merry AF.,** "Days alive and out of hospital," August, 2017. Perfusion Downunder – The Winter Meeting. Barossa Valley, South Australia.

17. **Megan Tuck,** "Wave Flume Island," PhD seminar series that included the animations created by the visulatisation services.

18. **Constantine, R.,** Environmental Defense Society Conference. August 2017, Auckland, New Zealand

19. Riekkola, L. et al. presentation at the New Zealand Marine Sciences Society Conference. July 2017, Christchurch, New Zealand

20. **O'Hagan, M.,** "Poetic entanglement with technology making university accessible and sustainable," in International Cultural Sustainability Conference. Aug 2017 Victoria University of Wellington, New Zealand.

21. **O'Hagan, M.,** Presentation in The first Cultural Sustainability Symposium in New Zealand, featuring special international guests. Aug 2017.

22. **O'Hagan, M.** Captions on Hololens. An invited speaker for a research seminar at the University of Roehampton, 11-DEC 2017.

Books

1. **Holdaway, S. J., Wendrich, W. (Eds.)** "The Desert Fayum reinvestigated: the early to mid-holocene landscape archaeology of the Fayum North Shore, Egypt," (Monumenta Archaeologica 39 ed.). Los Angeles: Cotsen Institute of Archaeology UCLA Social Sciences Division. 2017.

2. **Blumenstein, M., Liu, D., Richards, D., Leichtweis, S., & Stephens, J.,** "Data-informed nudges for student engagement and success," J. Lodge, J. Horvath, & L. Corrin (Eds.), Learning analytics in the classroom: Translating learning analytics research for teachers. 2018 Abingdon, UK: Routledge.

Abstracts

1. **Chudakova, D., Hannan, R. D., Hannan, K., Ganley, A. R. D.,** "Development of a ChIP-qPCR method to measure mitotic recombination in the ribosomal dna genes," in New Zealand Society for Oncology Conference 2017. Auckland, New Zealand.

2. **Chanprasert, W., Norris, S. E.,** "Turbulent natural convection in an air-filled square cavity," in 10th Australasian Natural Convection Workshop. Auckland, New Zealand.

3. **van Rijnsoever, M., McKenzie, E. J., Nadakkavukaran, I., Olynyk, J.,** "Fecal volatile organic compounds (VOCs) as a novel marker for

the detection of colorectal cancer," in New Zealand Society of Gastroenterology and NZNO Gastroenterology Nurses' College Annual Scientific Meeting 2017. Sky City Convention Centre, Auckland.

4. **Armfield, S. W., Kirkpatrick, M. P., Williamson, N., Lin, W., Norris, S. E.,** "The differentially heated cavity," in the 10th Australasian Natural Convection Workshop. Auckland, New Zealand.

Software packages

1. **(Reported by Wannes van der Mark)** Moss submission gateway, https://mossweb.cs.auckland.ac.nz/

2. **(Reported by Wannes van der Mark)** Moodle / Coderunner 2: https://coderunner2.auckland.ac.nz/moodle/

3. **(Reported by Samuel Ridler)** AmbulanceSim.jl [Ambulance simulation package for the Julia programming language]. (2017). Retrieved from https://github.com/samridler/AmbulanceSim.jl

4. **Mason, A. J.,** Optimisation Software For Real-Time Scheduling Of Field Officers For The 2018 Census [Computer Software].

Theses

1. **(Reported by Kathleen Campbell)** BSc Honours dissertation of Jessica Pelser, currently being examined.

2. **(Reported by Elizabeth McKenzie)** "Analysis of the metabolic profiles from faeces and plasma of children born very preterm" (Masters Thesis by Sachin Jayan, THESIS P17-046).

3. **(Reported by Klaus Lehnert)** The MSc thesis of Lydia Velzian is directly supported by this project.

4. **(Reported by Klaus Lehnert)** The MSc thesis of Chris Samson is directly supported by this project.

5. **(Reported by Margaret Coe)** PhD thesis submitted by Melinda Thomas.

6. **(Reported by Margaret Coe)** PhD thesis submitted by Jishan with David Greenwood.

7. **(Reported by Blake Matthew Seers)** PhD Thesis: "Investigating the climatic and oceanographic drivers of spatial and temporal variation in coastal turbidity and sedimentation".

8. **(Reported by Kane Lincoln Daniel Meissel)** "Advanced quantitative methodologies in education," Four PhDs are completing their analys using the Virtual Machines.

Posters

1. **Worrallo K, Rost D, Handley KM.,** "Impact of estuarine macro-algae on microbial communities and their biogeochemical cycles," in Annual Meeting of the New Zealand Microbiological Society.

2. **Boey JS, Couturier A, Mortimer R, Worrallo K, Handley KM.,** "The impact of terrestrial mud inputs on nitrogen cycling microbial communities in the intertidal zone of waiwera estuary," in Annual Meeting of the New Zealand Microbiological Society.

3. **M. L. Garelja, C.A. Walker, H.A. Watkins, D.L. Hay.,** "Probing the functional importance of the linker region of the class B GPCR, CLR, In CGRP and adrenomedullin receptors," in ASCEPT Conference, Melbourne, Australia.

4. **Riekkola, L. et al.,** "The great humpback whale trail: oceania to antarctica," The 6th International Biologging Conference, Sept 2017, Konstanz, Germany.

5. **Blumenstein, M., Leichtweis, S., & McDonald, J.,** "Learning analytics in practice: nudging students towards success," at the meeting of The Higher Education Technology (THETA) Auckland, New Zealand.

Papers in preparation

1. **Hu, T., Srinivasan, A., Tripathi, A. K.,** "Unstructured and short text classification."

2. **Bayati, S., Tripathi, A. K.,** "Recommending open source software projects to developers based on socio-technical activities"."

3. **Samin Aref.** "Balance and frustration in signed networks under different contexts," Paper in preparation for 2018.

4. **Michael L. Garelja, Christina A. Walker, Andrew Siow, Sung H. Yang, Paul W.R. Harris, Margaret A.** Journal Article in preparation.

5. **Brimble, Harriet A. Watkins, Joseph J. Gingell, Debbie L. Hay.** Journal article in late-stage preparation. "Receptor activity-modifying proteins have limited effects on the class B G protein-coupled receptor calcitonin receptor-like receptor stalk".

Others

1. **Reported by Robert Amor.** Two dissertations from the SOFTENG Part IV students submitted for their capstone project.

2. **Reported by Andrew Mason.** Finalist in NZSEA Spatial Awards;see http://www.nzspatialawards.org.nz/

3. **Reported by Ivan Braga Campos.** News report published in 12th of December 2016 at the Brazilian newspaper "O Globo". Link: https://oglobo.globo.com/sociedade/ciencia/meio-ambiente/pesquisa-usa-som-para-monitorar-cerrado-que-tem-alta-de-estudos-20628860.

4. **Debbie Hay** and co-applicants, 2017 HRC Programme Grant submission," Targeting peptide receptors to tackle migraine and pain"

5. **David Tomzik.** "Retro-fitting for data acquisition and analytics," a real-time 3D visualisation of a robot in the lab. see http://en-rvmf00132.uoa.auckland.ac.nz/

6. **Gunn, C., McDonald, J., Donald, L. C., Milne, J., & Blumenstein, M.,** "Building an evidence base for teaching and learning design using learning analytics," A Project report. Wellington: Ako Aotearoa

7. **Gunn, C., McDonald, J., Donald, L. C., Nichols, M., Milne, J., & Blumenstein, M.,** "Building an evidence base for teaching and learning design using learning analytics: Case studies," Wellington: Ako Aotearoa

Technical reports

1. **Dickson, M. E., Tunnicliffe, J.,** "Fan delta development on the western coast of the Coromandel Peninsula."

2. **Hafiz, F., Swain, A. K.,** "Selection of base wavelet and classifier for accurate identification of power quality events (694)."

Our thanks to all the authors from the University of Auckland who contributed the case studies. We'd also like to acknolwedge the New Zealand eScience Infrastructure (www.nesi.org.nz) for providing High Performance Computing platform.

OUR PEOPLE

Research IT Specialists - Centre for eResearch

Dr Cameron McLean, Team Lead
Research & Education

Sina Masoud-Ansari
Deep Learning

Dr Jamie Diprose
Software Development

Dr Doris Jung
User Experience

Martin Feller, Team Lead
Research Computing

Manish Kukreja
Research Computing

Dr Jason He
Research Computing

Sean Matheny
Research Computing

Yvette Wharton, Team Lead
Research Data

Robert Burrowes
Research Data Systems

Dr Prashant Gupta
Research Data Systems

Dr Dharani Sontam
Research Data

Nick Young
Visualisation and Analytics

Bianca Haux
Visualisation and Analytics

Dr John Rugis
Visualisation and Analytics

Leadership - Centre for eResearch

Professor Mark Gahegan
Director

Marcus Gustafsson
Manager - Operations

Jenny Lee Roper
Manager - Administration

eResearch Support - Centre for eResearch

Arron McLaughlin
eResearch Support

Warrick Corfe-Tan
eResearch Support

Doctoral Candidates - Centre for eResearch

Euan Forsyth
Doctoral Candidate

Adam Arvay
Doctoral Candidate

New Zealand eScience Infrastructure (NeSI) Auckland

Dr Brian Corrie
Solutions Manager

Gene Soudlenkov
Systems and Application Support

Dr Chris Scott
Scientific Programming

Jordi Blasco
HPC Architecture

Peter Maxwell
Application Support

Kumaresh Rajalingam
Analyst Programming

Shen Wang
Analyst Programming

Yuriy Halytskyy
Systems Engineering

Greg Hall
Systems Engineering

Anthony Shaw
Application Support

Anita Kean
Application Support

THE UNIVERSITY OF AUCKLAND
Te Whare Wānanga o Tāmaki Makaurau
NEW ZEALAND

Centre of eResearch
The University of Auckland
Science Building 302,
Room 590, 585, 580, 520
23 Symonds Street
Auckland 1010
New Zealand

Director
Mark Gahegan
Phone: +64 9 373 7599 ext 81891
DDI: +64 9 9231891
Email: mark.gahegan@auckland.ac.nz

eresearch.auckland.ac.nz