

INTERSECT

---

**Calendar 2014**  
Annual Report  
*Research . Technology . Impact*

# Welcome

## About Intersect

---

Intersect increases researcher impact through innovative products, services, technologies and expert advice. We work closely with our members and the wider research community to:

- Increase research productivity by decreasing cycle time from hypothesis to tested results;
- Expand research diversity by enabling collaborators to share data and experience across disciplines and across organisations; and
- Increase research longevity by storing and sharing the long tail of data beyond the research project lifecycle.



Productivity



Diversity



Longevity

## Mission

Intersect is Australia's premiere full-service eResearch support agency. We work closely with the research sector, including universities and public and private sector agencies.

Our portfolio of activities is designed to achieve our mission:

### Research Technology Impact

Intersect provides robust, innovative and collaborative technology to support the world-class research at our member institutions. Intersect delivers storage and analysis platforms, custom engineering, expert consultation and training programs to thousands of researchers every year. Members benefit from collaborative grant funding and economies of scale to deliver common infrastructure.

## NSW Chief Scientist's Message

"Looking back, Intersect Australia really was born into a perfect storm.

There were the great policy initiatives of open data right around the world, which were picked up very early here in New South Wales.

The Federal Government's policy emphasis was on collaboration between our research institutions and colleagues overseas.

And there were also scientific developments in big data, data analytics, and computational power to handle big data.

Intersect hasn't just worked along with those developments, it's ridden the wave of those developments – so that NSW now has a real presence in eResearch. That is so important for us because we are a long way from so many of the big centres of research.

Now, through effective eResearch, those centres can collaborate with us whatever time of the day, whatever time of the year. It's something that's been great for Intersect's members, the very visionary universities that founded it, supported of course by the State Government."

**Professor Mary O'Kane, NSW Chief Scientist and Engineer**



### Funded by



Trade &  
Investment



Australian Government  
Department of Education

### Member of



### Working with



# Foreword

by Dr Ian Gibson

---

"I am extremely proud of what we have achieved at Intersect. Starting in 2008 and growing to be the largest eResearch services organisation in Australia.

Intersect has delivered high quality services and a strong return-on-investment consistently for all members since day one. More importantly, we have enabled many thousands of research outcomes across many disciplines. Many, many of our undertakings have directly or indirectly impacted on the research outcomes of our members. I am deeply proud of our role in making the world a better place through those activities.

Intersect has made a difference. I also believe that Intersect is and will be ideally placed to have a growing role in enabling research in NSW and Australia.

Data analytics initiatives are what will drive research and innovation in the brave new big-data world. The future of world-leading research - across almost any discipline - is increasingly dependent on having access to the best data analytics and informatics skills and resources. Every organisation with global research ambitions needs to work out how they will achieve this. It is my belief that Intersect Australia is firmly positioned as the best possible partner for the journey".



A handwritten signature in white ink that reads "Ian Gibson".

CEO, 2008 - April 2015

## Intersect Members



# Table of Contents

---

|   |           |
|---|-----------|
| <b>Welcome</b>                                      | <b>2</b>  |
| <b>Table of Contents</b>                            | <b>4</b>  |
| <b>Changes in the Environment</b>                   | <b>7</b>  |
| <b>Chair's and CEO's Reports</b>                    | <b>8</b>  |
| <b>Governance</b>                                   | <b>10</b> |
| <b>Infrastructure: Storage and Compute</b>          | <b>12</b> |
| <b>Member Services</b>                              | <b>14</b> |
| <b>Exemplars</b>                                    | <b>16</b> |
| <b>Research Engineering, Projects and Platforms</b> | <b>20</b> |
| <b>Intersect Staff</b>                              | <b>26</b> |
| <b>Financial</b>                                    | <b>27</b> |



## Affiliate Members



Top: Intersect's successful showcase was hosted by the NSW Chief Scientist & Engineer, Professor Mary O'Kane, and revealed to those not so familiar with the company what Intersect offers. The showcase drew a crowd of over 140 people.



Middle left: Moves towards opening NSW data continued through 2014 with the NSW Open Data Summit in October, where speakers included then Intersect CEO Dr Ian Gibson.

Middle right: The showcase launched Intersect's large scale, high performance digital storage system for researchers, Space.intersect.org.au.

Below: Guest speaker at Intersect's Showcase was the Hon Dominic Perrottet MP, NSW Minister for Finance and Services who spoke on "how collaboration is changing technology within government, our relationship with industry and the economy itself. I don't have to tell anyone at Intersect the importance of collaboration. It's very much in the DNA of your organisation."

# Changes in the Environment

---

Since 2009, the Australian Government has invested approximately \$500m in eResearch through the National Collaborative Research Infrastructure Strategy, SuperScience and CRIS programs. These programs have delivered substantial developments in research infrastructure. Through university and direct channels, Intersect has received approximately \$20m of that investment, amounting to approximately 45 per cent of Intersect revenues (2008 - 2014). This includes infrastructure and software projects funded by the Research Data Storage Infrastructure (RDSI), the Australian National Data Services (ANDS) and the National eResearch Collaboration Tools and Resources (NeCTAR).

These funding programs have yielded fixed term projects. Long term, predictable funding remains unresolved. It is clear that eResearch infrastructure investments are unlikely to be funded in the same manner as they have been in the past. This will have substantial impacts on the opportunities Intersect leverages and the value proposition for Intersect members under the traditional subscription model.

Relative to five years ago, eResearch now has a much higher level of maturity across member universities both in appreciation of research enabled through eResearch methods and in the development of capabilities within universities. Perhaps more importantly, eResearch is becoming "mainstreamed" across a broad cross-section of research communities.

Most of Intersect's members now have internal organisational structures to deal with eResearch service provision. Their maturity in delivery eResearch services internally grows.

IT trends, the underlying drivers of eResearch methods, have also changed markedly. High-end computation is now achieved through much higher levels of parallelism which in turn has starkly exposed the "non-scalable" nature of a wide range of computational methods. Collaboration environments, five years ago a major focus for publicly funded research infrastructure, are beginning to be supplanted by market place services. The same can be said for mobile and shared data environments.

The following technology trends are apparent:

- delivery of IT services via cloud technologies is exploding;
- research using data analytics is increasingly practical, including machine assisted analysis of large datasets;
- information security is of growing importance, with failures in information security having greater impact;
- research solutions increasingly deploy technologies from an ever increasing number of vendors; and
- greater demand for managed data systems and eResearch as a service.

Changes in sector-wide funding opportunities, in eResearch maturity across Intersect stakeholders, in the breadth of disciplinary uptake of eResearch methods and in the need for continuously building capability due to shifts in IT trends, all conflate to challenge the status quo and business environment for Intersect Australia.

As Intersect diversifies its business to meet NCRIS obligations for infrastructure services, the need for a wider range of capabilities and more pluralistic value propositions for members has followed. This in turn raises the challenge of maintaining viability from a revenue base traditionally constrained to meeting NSW universities' needs.

# Chair's Report

Emeritus Professor Mark Wainwright AM FTSE

---

"Great technology platforms, new research infrastructure and a change of the guard were the themes which marked another big year for Intersect.

2014 saw Intersect's research data storage become fully operational, Intersect's NeCTAR Research Cloud built and a suite of virtual laboratories launched.

2014 also saw changes in the leadership team. I want to especially acknowledge the dynamic leadership of Dr Ian Gibson as CEO of Intersect. I wish him every success in his new endeavours.

Marc Bailey joined Intersect as Chief Operating Officer last year and in May 2015 was appointed by the Board to the position of CEO. Marc's appointment positions Intersect strongly for its next phase of its development: providing operational excellence and researcher focus across new product and service offerings.

A highlight of 2014 was Intersect's outstandingly successful showcase event hosted by the NSW Chief Scientist & Engineer, Professor Mary O'Kane. The showcase revealed to a number of people not so familiar with the Intersect enterprise the great success that it is due to the outstanding staff of Intersect. We appreciate the Chief Scientist's continued support.

We also welcomed the Australian Catholic University who became our twelfth university member. Intersect now represents twelve universities, with sixteen members in total.

The Board greatly appreciates the commitment and hard work of the Intersect team during 2014. My sincere thanks also go to the many member and independent representatives of the Intersect Board and Committees.

Intersect belongs to its Members. In turbulent times, commitment from the Members is even more crucial than in the founding years. The Board and I are committed to the long-term sustainability of Intersect, and we seek the same resolve from our Members. We look forward to the opportunities and challenges of the coming year".



# CEO's Report

## Mr Marc Bailey

---



"The one constant in research technology is change and 2014 was a watershed year for Intersect Australia. Members and customers increasingly voted with their feet to graduate from bespoke services to repeatable, reliable solutions and products that yield confident and consistent outcomes.

In calendar 2014, Intersect experienced "peak capital" as it converted \$5.2M of one-time grant funding into eResearch infrastructure. As National Collaborative Research Infrastructure Strategy, RDSI and NeCTAR project seed funding concluded, we created Space and Time - big data and big compute product lines - in anticipation of their strategic ongoing role in Australian research.

The NSW node of the National eResearch Collaboration Tools and Resources Research Cloud entered service in December 2014. Five national Virtual Laboratories powered by Intersect engineering and funded by NeCTAR were successfully launched in 2014. September saw us safely 'launch' your first one petabyte of research data into Space. Just as in a real space launch this culminated years of blood, sweat and tears of every staffer and engaged hundreds of Member data custodians. The research community now trusts [Space.intersect.org.au](http://Space.intersect.org.au) to store nationally significant collections in fields as diverse as medicine, particle physics and endangered languages.

2015 inevitably presents a new round of challenges. We are experiencing seismic changes in our business environment: reductions in research and higher education funding; a shifting level of maturity across member capabilities; disruptive developments in IT delivery; and changes in customer priorities all affect the value proposition Intersect creates.

My first strategic compact with Members is to make Intersect's activities more transparent and accountable by restructuring it into four business lines beginning with the 2015/16 financial year:

Space - big storage

Energy - big service

Time - big compute

Data - big analytics

Intersect faces difficult times and decisions in coming years in the wake of peak capital, but I remain optimistic about our future. My very personal and grateful thanks go to all staffers, Members, customers, office holders and friends who have contributed and supported us during my debut with Intersect.

My pledge is that together we will transform Intersect Australia into a more researcher relevant, financially resilient, and market viable organisation in the years ahead by leveraging the trust and respect upon which it is founded".

# Governance

---

Intersect's 2014 **Board of Directors** were:

- Emeritus Professor Mark Wainwright AM, Chair
- Dr Marilyn Sleigh, Independent
- Mr John Masters, Company Director, Independent
- Ms Anne Bell, University Librarian, University of Sydney
- Emeritus Professor Andrew Cheetham, formerly Deputy Vice-Chancellor of Research, University of Western Sydney
- Dr Ian Gibson, CEO, Intersect
- Professor Kevin Hall, Deputy Vice-Chancellor Research & Innovation, University of Newcastle
- Mr Michael Kirby-Lewis, Chief Information Officer, University of New South Wales
- Ms JoAnne Sparks, University Librarian, Macquarie University

Profiles of the Board can be viewed here: [www.intersect.org.au/board-of-directors](http://www.intersect.org.au/board-of-directors).

The role of the Committees is to provide advice and guidance to the Board

## **eServices and Strategy Committee:**

- Professor Andrew Cheetham, Chair and Board Member
- Dr Andrew Black, Director - Research Development and Collaboration, University of Sydney
- Mr Peter Gale, Information Technology Division, University of Technology, Sydney
- Mr Stephen Williams, Business Manager, Southern Cross University
- Dr Peter Sefton, Manager eResearch, University of Western Sydney
- Mr Ossie Richards, Manager Research Computing Services, IT, University of Newcastle

- Mr Grant Sayer, Development Manager, Informatics Development Group, Macquarie University
- Associate Professor David Miron, Research Director, Strategic Initiatives - PVCRC Office, University of New England
- Professor Judy Raper, Deputy Vice-Chancellor (Research), University of Wollongong
- Professor Mark Hoffman, Pro Vice-Chancellor (Research), University of New South Wales
- Professor Andrea Bishop, Director - Research, Charles Sturt University
- Dr Ksenia Sawczak, Director, Research Services Office, University of Canberra
- Mr Rick Van Haefen, Director, Information Technology, Australian Catholic University.

## **Infrastructure Committee:**

- Emeritus Professor Mark Wainwright AM, Chair and Board Member
- Professor Lindsay Botten, Director, National Computational Infrastructure, Independent technical expert
- Mr John Nicholls, Infrastructure Development Manager, AARNet, Independent technical expert
- Mr Steven Kuk, Acting Senior Manager ICT Infrastructure and Networks, University of Sydney
- Mr Jim Leeper, Acting Delivery Services Manager, University of New South Wales
- Ms Sarajane Hansen, CIO Cancer Institute, NSW State Government representative.

## **Resource Allocation Sub-Committee:**

- Professor Marc Wilkins, Ramaciotti Centre for Gene Analysis, UNSW
- Dr Ben Thornber, Mechanical Engineering, University of Sydney

## Intersect's committees provide advice and guidance to the Intersect Board

- Associate Professor Tim Langtry, Mathematical Sciences, UTS
- Professor Timothy Marchant, Dean of Research and Professor of Applied Mathematics, University of Wollongong
- Dr Craig O'Neill, Earth & Planetary Sciences Department, Macquarie University
- Professor Pablo Moscato, School of Electrical Engineering & Computer Science, University of Newcastle
- Dr Gregory Falzon, Computational Science, University of New England
- Professor Graham King, Professor of Plant Genomics and Epigenetics, Southern Cross University
- Professor Yang Xiang, Dean's Unit, School of Computing, Engineering & Math, University of Western Sydney
- Professor Terry Bossomaier, Director, Research Group for Complex Systems, Charles Sturt University
- Dr Joachim Mai, HPC Support Specialist, Intersect.

### **Storage Allocation Sub-Committee:**

- Dr Marilyn Sleight, Chair
- Associate Professor Grainne Moran, Director, UNSW Mark Wainwright Analytical Centre, UNSW
- Mr Ross Coleman, Director, Collections and eScholarship, University of Sydney
- Mr Peter Gale, Information Technology Services, University of Technology, Sydney
- Professor Nicki Packer, Chemistry & Biomolecular Sciences, Macquarie University
- Professor Timothy Marchant, Dean of Research, University of Wollongong
- Ms Lyn McBriarty, Director, Research Services, University of Newcastle
- Professor Heiko Daniel, Director, Research Services,

University of New England

- Mr Stephen Williams, Business Manager, Southern Cross University
- Mr Wayne Doubleday, Manager CSU Archives, Charles Sturt University
- Professor Deborah Sweeney, Associate PVC Research, Health & Science, University of Western Sydney
- Dr Ksenia Sawczak, Director Research Services Office, University of Canberra
- Ms Stefania Riccardi, Research Data Officer/ Project Manager, Australian Catholic University.

### **Board Audit and Risk Management Committee:**

- Mr John Masters, Chair and Board Member
- Mr Garry McLennan, Chief Financial Officer, FlexiGroup Limited, Independent
- Mr Michael Kirby-Lewis, Chief Information Officer, University of New South Wales
- Satish Nair, CFO, Sirca & Intersect
- Dr Ian Gibson, CEO, Intersect.

At year's end 2014 Intersect's Executive was as follows:

- Dr Ian Gibson, CEO
- Mr Marc Bailey, Chief Operating Officer
- Mr David Toll, National Member Services Manager
- Mr Johan Boshoff, NSW Member Services Manager
- Mr Rodney Harrison, Engineering and Professional Services Manager
- Mr Shane Youl, Operations and Support Manager
- Dr Derek Van Dyk, Director Strategic Development.

The qualifications, skills, experience and responsibilities of Intersect's management current team can be viewed here: [www.intersect.org.au/content/intersect-team](http://www.intersect.org.au/content/intersect-team).

# Infrastructure

---

2014 was a transitional year for Intersect: expanding its research offerings in big data, big compute and big service. Important research infrastructure came online, notably the virtual laboratories, Research Cloud and the Research Data Storage Infrastructure nodes (funded by NeCTAR, RDSI and NCRIS).

## Space.Intersect.org.au

The Intersect RDSI node entered mainstream production throughout 2014. At year's beginning there were 50 approved collections and 205 TB ingested. By December 2014 there were 79 approved collections and 1250 TB ingested.

In August, Intersect RDSI technology seeding was formally launched as 'Space', our big storage solution with a range of data storage options as product choices.



Space is Intersect's large scale, high performance, collaborative, and cost effective digital storage system. Space offers continuously growing capacity of up to 50 petabytes of fast, reliable and safe active and archive data retention.

In September we hit an historic Space milestone - Intersect safely launched one petabyte of research data into Space. The Intersect Team worked closely with dozens of data collection owners to expedite this

movement of 1000TB of research data, the equivalent of 4-door filing cabinets full of text.

In August 2014 the Research Data Storage Infrastructure project committed a further \$9.4M to support large research data collections at the Nodes. Intersect is involved in three of the five successful projects announced:

## Australian National Medical Research Data Storage Facility, now named Med.data.edu.au

This Intersect led initiative will, for the first time, enable the aggregation of health and medical data sets generated, curated and managed by many of the dozens of universities, medical research institutes and other government funded organisations who contribute to health and medical research in Australia. Data will be primarily located at Intersect, VicNode, QCIF and eRSA.

## National Genomics Data Storage Facility

VicNode, QCIF and Intersect RDSI nodes will establish a facility for storing and sharing data collections in genomics, proteomics and metabolomics. Each node will store data close to the instruments and compute power.

## Australian Coordinated Characterisation Data Space

VicNode, Intersect and QCIF will serve the Australian characterisation research community with instances of the Characterisation Virtual Laboratory at each node, supporting collaboration and streamlined data access and processing.

## NeCTAR projects and the Research Cloud

The NeCTAR Research Cloud expanded in 2013 with more research cloud nodes joining the federated effort and existing cloud nodes expanding their offerings.

## Research storage and compute over the Internet in the cloud

Intersect's submission for NeCTAR Research Cloud architecture was approved in 2013Q4, procurement began in 2014 and the Intersect Research Cloud node ([rc.NeCTAR.org.au](http://rc.NeCTAR.org.au)) entered service in December 2014.

Many of the virtual laboratories funded by NeCTAR were launched in 2014. Intersect was instrumental in:

- the All-Sky Virtual Observatory led by Astronomy Australia Ltd;
- Federated Archaeological information Management System led by UNSW (and now Macquarie University);
- Humanities Networked Infrastructure (HuNI) Virtual Laboratory, led by Deakin University;
- Alveo, the Human Communication Science Virtual Laboratory; and the
- Biodiversity Climate Change Adaptation Virtual Laboratory, led by Griffith University.

### Successful ARC LIEF grant

An Intersect led consortium of 11 universities was also successful with an ARC LIEF grant for funding commencing in 2014 entitled "Renewing Intersect's share of the National Computational Infrastructure's peak facility". \$1.025 was awarded by the ARC to fund the Intersect consortium merit share of the NCI facility for the next three years. The lead CI for this proposal was Associate Professor Evatt Hawkes, with UNSW the administering organisation.

This is the third successful ARC LIEF grant in sequence to fund High Performance Computing facilities for NSW researchers. This represents a leverage factor of 1.9 for a co-investment of \$540,000 by members.

### Help.intersect.org.au

Intersect also launched [Help.intersect.org.au](http://Help.intersect.org.au) so our

researchers can explore new services and seek out new answers and new products. Powered by ServiceNow.com, [Help.intersect.org.au](http://Help.intersect.org.au) is a service desk, deployed to support infrastructure users with HPC and the NeCTAR Research Cloud node (Time), and Research Data Storage (Space). It provides a self serve experience for requesting, monitoring and ordering Intersect services and products.



[Help.intersect.org.au](http://Help.intersect.org.au) is home to our catalogue of products and services going forward, following with resource and storage allocation. 'Help' takes Intersect service truly national by allowing any of the 1.2 million Australian Access Federation accounts to connect instantly.

### Acknowledgements

NeCTAR is an Australian Government project conducted as part of the Super Science initiative and financed by the Education Investment Fund.

The Research Data Storage Infrastructure (RDSI) Project, an initiative of the National Collaborative Research Infrastructure Strategy, is funded from the Education Investment Fund under the Super Science (Future Industries) initiative.

The Research Data Services (RDS) Project is an initiative of the National Collaborative Research Infrastructure Strategy (NCRIS).

Intersect share of NCI's Rajjin is funded through the Australian Research Council's Linkage Infrastructure, Equipment and Facilities (LIEF) scheme.

# Member Services

In reporting on Intersect's services during 2014 we present the aggregated values of services such as HPC, data storage, training and eResearch Analyst services, as well as highlights from the other services.

Member value reports for the calendar year 2014 were distributed to each eStrategy and Services representative in early 2015.

## Services delivered in 2014



## Services

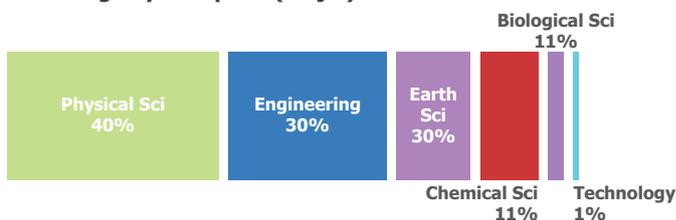
Services are conservatively valued on what it would cost a member for the same service at market rates. An aggregated summary of services delivered during the 2014 calendar year is below:

| Service Description                           | \$               |
|---|------------------|
| High Performance Computing (HPC)              | 1,305,407        |
| Data Storage                                  | 461,504          |
| Training                                      | 290,550          |
| eResearch Analyst Services                    | 899,080          |
| Hosted Services                               | 196,500          |
| Software Engineering                          | 2,106,588        |
| Grants enabled                                | 35,000           |
| <b>Total value of services delivered 2014</b> | <b>5,294,629</b> |
| Less: Payment received from members           | 3,737,019        |
| <b>Net costed value delivered in 2014</b>     | <b>1,557,610</b> |

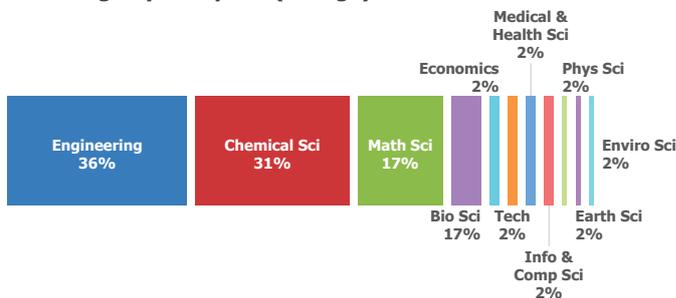
## High Performance Computing (HPC)

High Performance Computing time is valued at \$0.04 per core hour. More than 31,079,000 core hours were consumed in 2014. This is the combined usage of the 'Orange' as well as 'Raijin' supercomputers. High Performance Computing time is allocated on a merit basis and delivered to members at no additional charge.

### HPC Usage by Discipline (Raijin)



### HPC Usage by Discipline (Orange)



### HPC Usage by Member



## Research Data Storage

Data storage for the 2014 year was valued at \$45 per terabyte (TB) per month. At the end of 2014, Intersect hosted more than 1.7 petabytes of its members' valuable research data collections. Members paid a contribution fee to help fund the operational cost of the infrastructure.

# Annual subscription by all Members generates concrete researcher benefits



## Training

A total of 91 courses were delivered at no charge to researchers at Intersect member campuses, with 1033 individuals trained.

| Description   | Individuals trained |
|---|---------------------|
| Advanced Excel for Researchers (Excel Fu)                   | 575                 |
| Beginner to Intermediate HPC - Introduction to Unix for HPC | 107                 |
| Intermediate HPC - from Unix to HPC                         | 0                   |
| Advanced HPC - Parallel Programming                         | 24                  |
| Cleaning and exploring data with OpenRefine                 | 152                 |
| Data Visualisation with Google Fusion Tables                | 175                 |
| Introduction to Machine Learning                            | 0                   |
| <b>Totals</b>   | <b>1033</b>         |

## eResearch Analyst Services

- Intersect eResearch Analysts assist researchers to optimise their research outcomes by:
- Informing researchers on Intersect products and services to complement local university services.
- Analysing researcher needs with a view to developing new tools/software where existing solutions are inadequate.
- Providing advice on the access and best use of high

performance computing (HPC).

- Developing research data management plans to accommodate data storage and collaboration objectives.
- Recommending existing/new tools to improve workflow and research productivity.
- Sharing best practice information from other universities, disciplines and research groups.



Photo: Jackson Mann, UTS Library

More than 400 high level engagements occurred in 2014 related to significant research and projects. A sample of these engagements can be seen on pages 18-19.

## Hosted Services

Intersect hosts a significant number of software solutions on behalf of its membership. Solutions range from software projects developed by Intersect to third party solutions that require hosting. Intersect hosted 28 environments for its members in 2014.

## State and Federal funding opportunities

An important benefit of Intersect's membership is the capacity to leverage government eResearch funding. In 2014, Intersect received \$5.264m from NCRIS capabilities.

# Space Exemplars

Many research workflows require high-speed transfer of data, sharing and collaborating with others, managing access to data, and connecting data to high performance and cloud computing platforms. Intersect is answering those needs by developing a range of services to support sharing and collaboration around research data.

Announced in August 2014, [Space.intersect.org.au](http://Space.intersect.org.au) is a platform developed by Intersect to provide a large-scale, high performance collaborative and cost-effective digital storage system with the following benefits:

- it is networked to the Australian education/research sector at high speed
- it is easily and securely accessed with existing researcher IDs
- data can be read, written, copied or moved over the internet at high speed using SpaceShuttle.

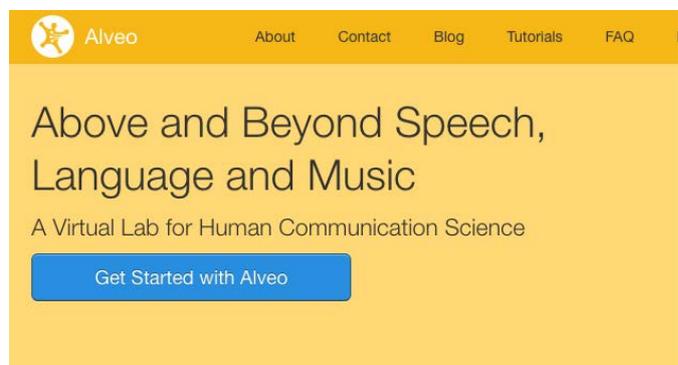
Here are a few of the 79 research collections using [Space.intersect.org.au](http://Space.intersect.org.au) at the conclusion of 2014. These collections were not available for reuse or collaboration prior to this facility's existence. All of these collections will likely support new research.

## Alveo for linguistics and cognitive sciences

The Alveo collection currently has 70 terabytes of cross-disciplinary datasets housed on Intersect's storage facility. Integrated with analysis tools they form the NeCTAR virtual laboratory, known as Alveo. Alveo serves up human communications data, consisting of language and music data, in their three most common modes – audio, auditory-visual and text.

Alveo supports research in automatic speech recognition (taxi ordering, directory assistance, etc.), hearing aids and cochlear implants, interactive learning programs for

children with learning disabilities, forensic determination of origin and background of particular accents, psycholinguistic studies of second language learning and pedagogy, socio-linguistic studies, anthropological research, and more.



The datasets are the result of researchers at the 15 collaborating institutions winning competitive grants, mostly ARC Discovery and LIEF grants, over the past 5-15 years.

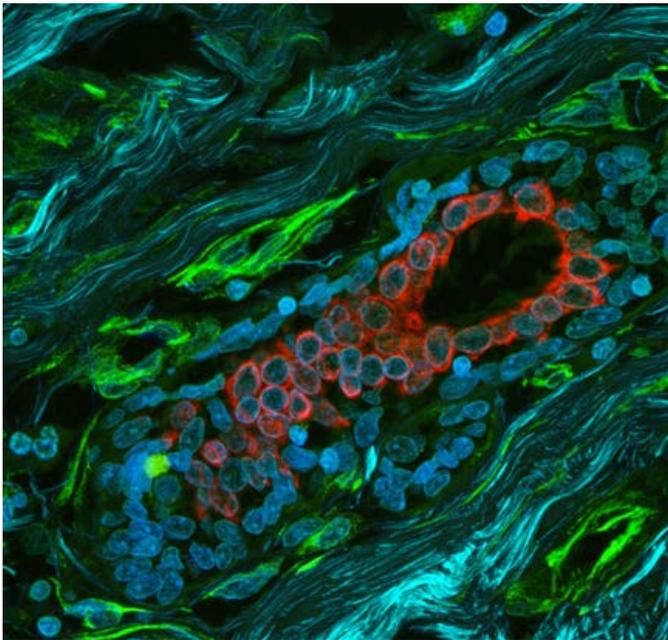
## The Australian Breast Cancer Tissue Bank

The Australian Breast Cancer Tissue Bank (ABCTB) exemplifies the innovation that Intersect's use of storage solutions can bring to the medical research community. The ABCTB is a breast cancer research resource that is supported by a collaborative network of cancer clinicians and researchers, with the aim to improve outcomes for breast cancer patients. It is Australia's only large scale collection of specimens and data dedicated to sporadic breast cancer.

These data collections comprise scans of tissue samples, associated clinical information and high-resolution images of tumour sections. The associated data can total up to 50 terabytes. As a result, the ABCTB is improving the availability of quality research for sharing and re-use

## Petabyte, scale digital storage makes the imaginable possible. These are just some of the 79 meritorious data collections now in [Space.intersect.org.au](http://Space.intersect.org.au)

among authorised researchers and clinicians, improved research efficiency and reduced institutional data storage costs.



A breast cancer biopsy stained and images using confocal microscopy by A Chien, E Kable and L Soon, University of Sydney.

### **PARADISEC: protecting endangered languages**

The Pacific and Regional Archive for Digital Sources in Endangered Cultures (PARADISEC) preserves recordings of hundreds of endangered languages from the Asia-Pacific region. PARADISEC is an exemplar for building secure infrastructure that enables citable, sharable and reusable data.

PARADISEC is an archive of over 80,000 items of data – including audio, video, image, and text – from around 850 endangered languages and cultures in the Asia-Pacific region, collected by researchers as part of their fieldwork over the past 60 years. Many of the recordings are unique, unrepeatable and represent the only known

documentation of particular communities from the Asia-Pacific region. In addition to being stored for posterity in digital form, the data is important for linguistic typology, comparative linguistics, anthropology, musicology and ethnomusicology.

Researchers use NABU, PARADISEC's online catalogue, to describe their research data, to cite their data in publications, and to facilitate access to the materials for other researchers, community members, or anyone who has an interest in endangered cultures.

Intersect SpaceLab provides secure infrastructure for the 20TB of data and a hosted environment for the NABU catalogue software to enable data discovery and encourage re-use from researchers around the world.

PARADISEC is funded by the University of Sydney, the University of Melbourne, the University of New England, ANU, the Australian Research Council and GrangeNet since its establishment in 2003.



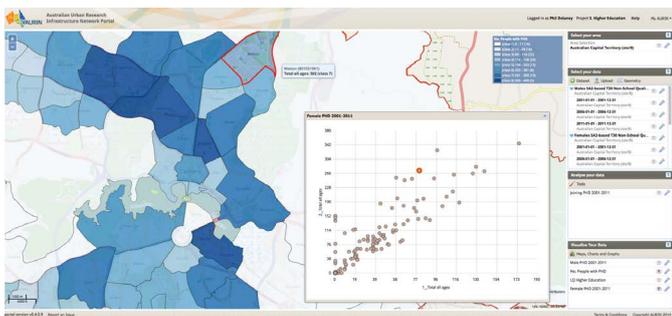
Photo: Jeremy Hammond

# Consulting Exemplars

Intersect has a team of experts which provides software engineering, advice and consultancy services across a broad range of areas. Intersect experts assist wherever there are opportunities to synergise. Here are a few recent examples.

## AURIN and Housemod

The National Centre for Social and Economic Modelling (NATSEM), based at the University of Canberra has



developed an online microsimulation model of housing in Australia as an eResearch tool. Named Housemod, the model combines Australian Bureau of Statistics (ABS) Census, survey and demographic projections to accurately replicate the current and future regional housing markets. Housemod includes very detailed data at a regional level on the demographic, social and economic circumstances of households.

Intersect worked with NATSEM to build an AAF-enabled web application that allows researchers to enter simulation parameters and pass them seamlessly to the underlying model. Once the simulation has finished, users receive an email notification that the results are ready to download and analyse.

Users of the model can forecast out to 2027 the effect of a change in GDP, mortgage interest rate, level of rent

assistance, etc. A simulation can generate results as fine grained as ABS Statistical Area 2 Level (populations of approx. 3,000 - 25,000 people). The simulation allows results to be compared against a “base world”, in order to determine the winners and losers of a change in policy or economic conditions.

Funding was awarded by AURIN, Australia’s Urban Research Infrastructure Network, to make the model available to researchers everywhere.

## Children’s Medical Research Institute

The Children’s Medical Research Institute (CMRI) is an independent research institute dedicated to discovering how to prevent or treat childhood diseases. Their medical research programs in cancer, embryology, cell signalling, and gene therapy use genomics, proteomics, imaging, and other technologies that generate extremely large data sets.

Head of Bioinformatics at the Children’s Medical Research Institute, Associate Professor Jonathan Arthur outlined the challenge, “the problem that we were trying to solve involved ...looking at the genetic makeup of a thousand tumour samples in order to try and identify potential therapeutic targets”.

This required using 560 CPU cores, 2 Terabytes of RAM and 20 Terabytes of disk storage. This was a very large request in terms of CPU and Storage, a large amount of compute over a short period, a tight project deadline, and a requirement to store the results to archival storage.

Intersect recommended Time product powered by rc.nectar. Intersect’s SpaceShuttle was used for the data transfer and storage, because it achieves speeds that are

## Intersect eResearch analysts, engineers, consultants and trainers augment local capabilities across disciplines and across organisations

hundreds of times faster than FTP and HTTP protocols. Intersect worked with the research team to revisit their requirements, a minor tweak in the workflow removed the requirement of 20 TB object storage. The analysis could be done with the 480 GB ephemeral disk that is allocated with each VM. The Space quota was increased and the allocation was provisioned.

CMRI was able to download the data from CGHub in the U.S.A. leveraging the AARNET 10Gbps link. The research team was set up the compute environment on Intersect Time (including software installation and testing), and was able to analyse the data in ~40 days and then stored the 8.5 TB of results in SpaceShuttle.

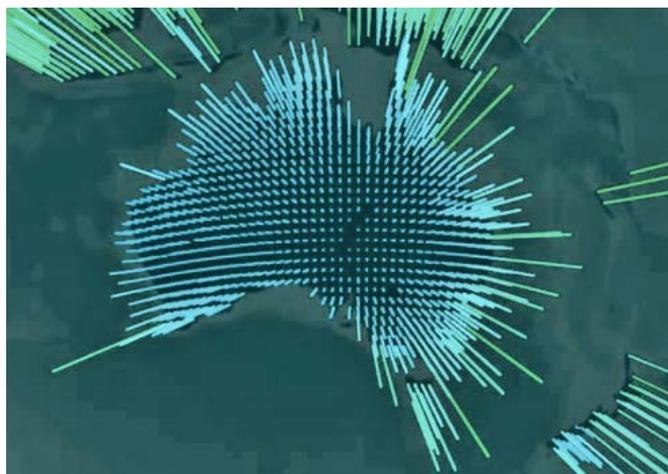
Arthur says, "The process has turned out to be very simple and very straightforward, contrary to my expectations. As a result I think we might make more use of the NeCTAR Research Cloud in the future".

### eMAST ecosystem modelling

The Ecosystem Modelling and Scaling Infrastructure Facility (eMAST) is a facility of the Terrestrial Ecosystem Research Network (TERN) that assembles data sets for integration into modelling applications. The foundation data are high-resolution estimates of the key climate variables: temperature, precipitation and vapour pressure. Users of the collection include ecosystem, climate and ecological scientists within Australia and internationally. Intersect members involved include those from UNSW, MQ, UTS, UWS, CSIRO, and the NSW Office of Environment and Heritage.

In September 2014, Intersect started working with eMAST on an ANDS Major Open Data Collection project. Intersect provided 145TB of online disk storage (Space

and 5 virtual machines (Time). On these virtual machines Intersect installed discipline specific open-source software, software developed by eMAST and software developed by Intersect. DIVER, a searchable file repository, was the primary software provided by Intersect. DIVER provides, among other features, AAF authentication, packaging of files into collections and production of RIF-CS Schema for harvesting via an Open Archives Initiative Protocol for Metadata Harvesting (OAI-PMH).



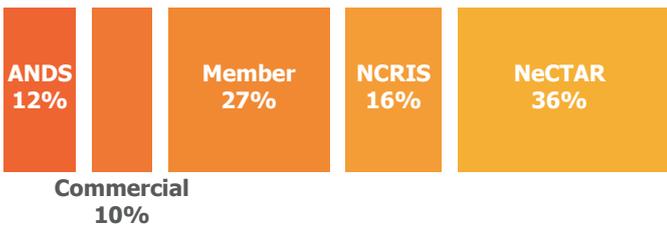
To support workflow, an iPython notebook was also developed that communicates with DIVER and provides a Vocabulary Service to aid discovery of eMAST data products. The eMAST-develop model can then run on Intersect Time and save its output directory to Space where it can be published and made available for other researchers to use.

Dr Brad Evans, Director of eMAST said, "This ANDS project has enabled eMAST and Intersect staff to collaborate on the challenges of connecting science with metadata and data services".

# Research Engineering

Organisationally embedded software engineering for research continues to challenge even the largest organisations. Intersect members benefited from numerous software engineering projects delivered in 2014. Our research engineering services are delivered by a team of professional business analysts, project managers, software engineers, user interface designers and testers.

## Funding Stats



The year was also marked by partnering with researchers on repeat business over an extended time, such as the AURIN projects and the FAIMS modules, to create sustainable research tools.

## Technologies Stats

Intersect's range of software capabilities continues to be broad, as the stats here indicate.



Ruby on Rails  
**44%**



Java  
**21%**



Python  
**15%**



Android  
**7%**



PHP  
**3%**



Other  
**10%**

## Alveo 2

Above & beyond speech, language and music Human Communication Science (HCS)



Professor O'Kane, NSW Chief Scientist and Engineer, launching Alveo at the University of Western Sydney.

Focusing on the manner in which humans communicate with each other, and with computers and machines via codified means, the Alveo virtual laboratory allows a diverse range of researchers to access corpora and analytical tools.

Intersect developed a range of tools enabling researchers to access human communication data sets (speech, texts, music, video, etc) and to search, analyse and annotate data previously spread across various sites with varying accessibility. The project is led by the University of Western Sydney.

## The All-Sky Virtual Observatory

Linking observational and theoretical capabilities Astronomy

The All-Sky Virtual Observatory (ASVO) provides a direct and vital link between the theoretical and observational aspects of data collection and analysis. Intersect worked with Swinburne University of Technology, the Australian

# Intersect partners, augments and accelerates engineering with best practice and best people for best outcomes.

National University, the National Computational Infrastructure, and Astronomy Australia Ltd.

The All-Sky Virtual Observatory provides access and add functionality to key astronomical resources of national significance, links observational data archives and theoretical infrastructure capabilities, and establishes a platform from which to exploit the growth in astronomical data in the coming decade.

## APProve-Lite

Mobile phone app developed for clinical trials  
Health & Medical Research

APProve: CAN Probiotics ImPROVE Breastfeeding Outcomes?

APProve is a randomised controlled trial that aims to assess the effectiveness of using probiotics to improve the health of breastfeeding women and their babies during the first two months after giving birth. Improving maternal health has the potential to increase the length of time women breastfeed their infants, resulting in long term benefits for you and your baby.

To support this trial, a smart-phone application is being trialed to collect information about you and your baby's health and well being. Along with daily and weekly text messages to link you to the APProve surveys, you will be able to access this webpage at any time. Feel free to scroll through the tabs to review important study information and reminders.

APProve-Lite is a novel mobile application to remind patients of their treatment needs, assess maternal breastfeeding outcomes, monitor maternal and infant infections and measure the acceptability of using the mobile app itself. APProve-Lite supports the APProve (CAN Probiotics ImPROVE Breastfeeding Outcomes?) Trial in order to collect daily information on treatment compliance and symptoms related to the health and well-being of trial participants.

Developed for the Kolling Institute, University of Sydney, APProve-Lite is a novel mobile application to remind patients of their treatment needs, assess maternal breastfeeding outcomes, monitor maternal and infant infections and measure the acceptability of using the mobile app itself. APProve-Lite supports the APProve (CAN Probiotics ImPROVE Breastfeeding Outcomes?) Trial in order to collect daily information on treatment compliance and symptoms related to the health and well-being of trial participants.

## AURIN Housing Lens and other projects

The new national urban intelligence initiative  
Built environment



AURIN, the Australia Urban Research Infrastructure Network, provides access to thousands of data sets—from Australian Property Monitors to the Australian Bureau of Statistics, from Geoscience Australia. It allows researchers, planners and policy-makers to answer many questions - what makes a neighbourhood walkable? Where are the hotspots for mortgage stress or affordable housing in Sydney?

AURIN is a national collaboration project delivering eResearch infrastructure to empower better decisions for Australia's urban settlements and their future development. The aim is to build a fully spatially co-referenced set of interlinked databases.

Intersect was involved as one of the technology providers in a range of AURIN projects: the Low Carbon Living business analysis report, Housemod (Urban Housing project) - an eResearch tool allowing online microsimulation modelling of housing in Australia, the ACT Data Hub for regional data, and the Urban Health Data project, developing infrastructure to provide medical research data and health data available from the National Health Services Directory into the AURIN online workbench.

# Software Exemplars

## Biobanking Stakeholders Network Pt2, Cancer Inst NSW

Harmonising image and IT platforms

Health and Medical Research



Photo: Dan Catchpoole

The Cancer Institute NSW's Biobanking Stakeholders Network (BSN) involves 23 biobanks. Intersect's work with the biobanks involved surveying the IT platforms/ data sharing and pathology image capture and analysis systems in use across the BSN, and ascertaining attitudes and barriers to the harmonisation of platforms. The resulting report suggested recommendations for implementing streamlined IT and imaging systems across the entire BSN.

## The Biodiversity & Climate Change Virtual Lab

Modelling biodiversity and climate change



The BCCVL allows researchers to project future outcomes using emission scenarios and climate models by combining experiment types, datasets and processing power, all within the cloud. It enables a phase shift in biodiversity modelling by providing an interface linking data and model algorithms with high performance computing infrastructure, and enabling novel analyses including potential climate change impacts on biodiversity. Headed up by Griffith University, Intersect was a development partner.

## Building Value

Realistic costs of building

Built Environment

Intersect continues work on this sophisticated IT system for the storage and management of benchmarking data based initially on the Building the Education Revolution (BER) program of school infrastructure projects. Building Value derives statistically valid comparative data on the cost, value and quality of those projects. Intersect designed and developed the database application to contain, manage and enhance the Building Value project data.

## CareTrack Kids

Health & Medical Research

CareTrack for Kids is a 2014-5 follow-on project from the Caretrack Australia system developed by Intersect in 2011 for use by the Australian Institute for Health Innovation at Macquarie University. The original was developed to aid the research team track medical records of participants with known conditions across multiple healthcare providers, analysing care and treatment. The same team has turned its attention to the field of paediatrics and

Each research project is unique. Intersect engineering crystallises research ideas and techniques into algorithms, databases and systems.



Newborn examination by Nevit Dilman, CC BY-SA 3.0

Intersect has extended the original Caretrack Australia application with a new module.

### CASDA Review

Architecture Review for CSIRO  
Astronomy

CSIRO engaged Intersect to perform an independent architecture review of the CSIRO Australian Square Kilometre Array Pathfinder Science Data Archive (CASDA). The CASDA system will be the principal data archive for the Australian Square Kilometre Array Pathfinder (ASKAP) radio telescope. The purpose of the review was to determine if the design for the CASDA system proposed by the CSIRO was capable of managing the substantial data volumes to be generated by the ASKAP telescope.



One of 36 12-metre dishes that make up the ASKAP telescope at the Murchison Radio-astronomy Observatory in WA. Image: CSIRO.

### CR8IT

Securing, archiving and citing long-tail research data  
Research Data Management

An OwnCloud plug-in, Cr8it is a University of Western Sydney initiative that provides the general community of researchers with a simple and consistent way to manage, archive and advertise their research datasets. It was developed in-house by UWS and was further developed by Intersect with funding from the University of Newcastle and UWS.

### FAIMS Phase2

Unearthing Android's potential for archaeology  
Archaeology



CSIRO researcher Alistair White using the FAIMS module for geochemical sampling in the Capricorn Distal Footprints Project.

The FAIMS application radically improves the workflow of archaeology. The Federated Archaeological Information Management System (FAIMS) includes an Android application and Ruby server built by Intersect. It is a comprehensive system, built for the digital collection and management of archaeological data, be it text, image, audio or video. The system allows field and laboratory data to be born digital using mobile devices, processed in local databases, extracted and exchanged online.

# Research Platforms

In late 2013 the FAIMS group was successful in attaining a LIEF grant from the Australian Research Council to deploy the tool Australasia wide, and returned for a second phase of development with Intersect. Take up in the field has Intersect creating modules for practising archaeologists. The android tool is also being used in other fields where off-line data collection is required, e.g by GNS Science, New Zealand, who are “keen to keep using it for subsequent sampling”.

## Humanities Networked Infrastructure (HuNI)

Australia’s cultural datasets online  
Creative Arts

Integrating 28 of Australia’s most important cultural datasets into a virtual laboratory, HuNI allows researchers from different disciplines to share findings and uncover new insights into Australia’s cultural landscape. HuNi was developed as a partnership between 13 institutions, led by Deakin University. Intersect provided user interface, semantic and business analysis capabilities.

## Launchpod

[intersect.org.au/content/launchpod](http://intersect.org.au/content/launchpod)

Build your software in the cloud  
NeCTAR porting tool

Launchpod is a tool developed by Intersect to allow anyone in an Australian research institution to quickly and easily deploy a virtual machine and run software for use in the cloud. Cloud computing means researchers don’t need high-powered computing to run complex analytical software.

Launchpod makes cloud computing even easier by taking care of the complex software management. It works like

a wizard: fill in some basic details, and it takes care of the technical details. You can build a machine to run R Studio or MATLAB® in the cloud without downloading CentOS packages or install Apache. Best of all, Launchpod uses the researcher’s personal NeCTAR Research Cloud allocation, meaning any Australian researcher can build a virtual machine quickly and easily.



Boeing X-37B inside payload fairing before launch, US Air Force.

## POSSUM

Android/web app for observing drug admin in hospitals  
Health & Medical Research

Do interruptions and multi-tasking impact on medication administration errors rates? Another great tablet based project, POSSUM is for Precise Observational System for the Safe Use of Medications. Fully mobile, the POSSUM tablet app keeps collecting data when offline, syncing with a central database when back in wifi range.

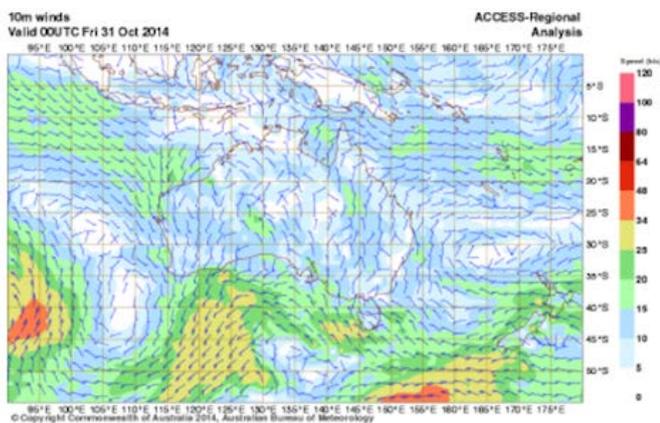
# We seek to identify common practice, experience and techniques to leverage quality software engineering into reusable, redeployable platforms

The POSSUM study tests an observational method for measuring medication administration errors suitable for use in Australian hospitals, and compares the type, rates and severity of medication administration error rates before and after various interventions. POSSUM was developed for the Centre for Health Systems and Safety Research (CHSSR), Australian Institute of Health Innovation - Faculty of Medicine and Health Sciences, Macquarie University.



## TAPPAS, Tool for Assessing Pest and Pathogen Aerial Spread

Plant Biosecurity  
Environmental Science



TAPPAS V1.0 has the basic functionality of dispersion modelling, ability to view previously run models and

visualisation of those modelled runs. TAPPAS allows end-users (State & Federal Governments and industry partners) to appreciate new biosecurity threats posed by wind dispersal to guide surveillance, preparedness, incursion response, and resource prioritisation. The prototype TAPPAS was developed by CSIRO IT services, and further developed by Intersect through a series of intensive sprints alternating with extensive user testing and consolidation work.

## UTS Metadata Stores Project (MS22) now 'Stash'

A university-wide metadata store  
Data Management

This university-wide metadata store is now integrated with existing enterprise systems to administer and plan data storage and other infrastructure. 'Stash' makes research data available for reuse within UTS and the wider community via data harvest and discovery services. Intersect performed the analysis, design, building and testing of the discrete units. The metadata store, or Research Data Catalogue supports registration of datasets, the maintenance of metadata, and search and discovery of data collections by UTS staff and via Research Data Australia.

## Acknowledgements

The Alveo, FAIMS, BCCVL and Launchpod projects acknowledge funding from the NeCTAR project. NeCTAR is an Australian Government project conducted as part of the Super Science initiative and financed by the Education Investment Fund.

The UTS Metadata Stores and Urban Health Data projects are supported by the Australian National Data Service (ANDS). ANDS is supported by the Australian Government through the National Collaborative Research Infrastructure Strategy Program and the Education Investment Fund (EIF) Super Science Initiative.

AURIN is funded through the National Collaborative Research Infrastructure Strategy (NCRIS).

# Intersect Staff

May our Force be with you

Behind all our technology stands an energised, enthusiastic staff that understands research, higher education and science.

We're here to serve and to make a difference.

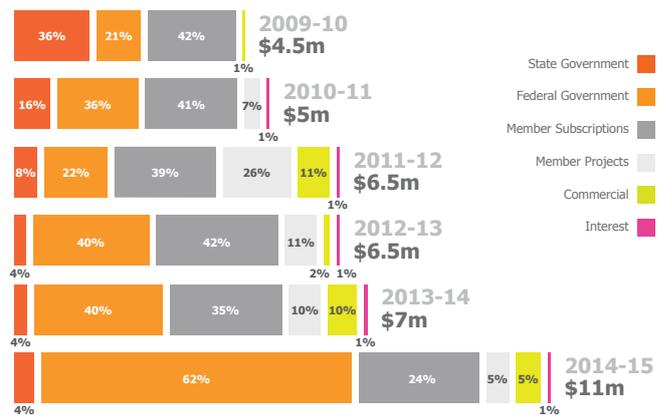


## Staff Numbers by financial year (2009-2014)



Source: Intersect Human Resources Department, based on actual headcount numbers.

## Revenue Sources by financial year (2009-2014)



Source: RSM Bird Cameron audited accounts

# Financial

## Independently audited financial statement Profit & Loss July 2013 - June 2014

|  | Jul 13 - Jun 14   | Budget            | Variance          |
|--|-------------------|-------------------|-------------------|
| <b>Ordinary Income/Expense</b>         |                   |                   |                   |
| Income                                 |                   |                   |                   |
| Membership Subscription Fees           | 2,495,000         | 2,495,000         | 0                 |
| Affiliate Membership Fees              | 40,000            | 110,000           | -70,000           |
| Professional Services – Other Research | 307,164           | 150,000           | +157,164          |
| Projects Income – Member               | 496,566           | 556,200           | -59,634           |
| Projects Income – Other Research       | 2,417,202         | 2,922,049         | -504,847          |
| Projects Income – Non-Member           | 308,071           | 640,943           | -332,872          |
| Hosting Services – Member              | 19,155            | 4,000             | +15,155           |
| Other Asset Reimbursement              | 39,964            |                   |                   |
| Other Income                           | 10,000            | 20,000            | -10,000           |
| NeCTAR RC                              | 280,172           | 440,521           | -160,349          |
| NeCTAR RC Assets Reimbursement         | 0                 | 1,026,595         | -1,026,595        |
| NECTAR Member Contribution             | 0                 | 166,000           | -166,000          |
| RDSI SLF Funding                       | 440,000           | 440,000           | 0                 |
| RDSI Member Contribution               | 0                 | 265,000           | -265,000          |
| RDSI ReDs                              | 1,158,500         | 900,000           | +258,500          |
| RDSI Asset Reimbursement               | 2,141,441         | 1,980,000         | +161,441          |
| Other RDSI Income                      | 752,344           | 280,000           | +472,344          |
| Recharge Income                        | 17,727            | 0                 | +17,727           |
| <b>Total Income</b>                    | <b>10,923,306</b> | <b>12,396,308</b> | <b>-1,473,003</b> |
| Gross Profit                           | 10,923,306        | 12,396,308        | -1,473,003        |
| Expense                                |                   |                   |                   |
| Resourcing Costs                       | 6,488,023         | 6,812,454         | -324,431          |
| Network & Telcom Costs                 | 762,437           | 662,475           | +99,962           |
| Membership & Licence Costs             | 21,168            | 30,300            | -9,132            |
| Occupancy Expenses                     | 368,468           | 368,320           | +148              |
| External Professional Services         | 234,545           | 145,000           | +89,545           |
| Marketing & Promotions                 | 44,049            | 50,500            | -6,451            |
| Travel & Entertainment Expenses        | 210,647           | 140,100           | +70,547           |
| Insurance                              | 47,460            | 51,373            | -3,913            |
| Office Expenses                        | 60,755            | 42,000            | +18,755           |
| Other Expenses                         | 26,348            | 6,000             | +20,348           |
| NeCTAR RC Expenses                     | 0                 | 193,150           | -193,150          |
| RDSI Expenses                          | 502,593           | 386,780           | +115,813          |
| <b>Total Expense</b>                   | <b>8,766,492</b>  | <b>8,888,452</b>  | <b>-121,960</b>   |
| Net Ordinary Income                    | 2,156,813         | 3,507,856         | -1,351,043        |
| Other Income/Expense                   |                   |                   |                   |
| Other Income                           |                   |                   |                   |
| Interest Income                        | 23,736            | 36,000            | -12,264           |
| RDSI Interest Income                   | 29,016            | 36,000            | -6,984            |
| Total Other Income                     | 52,752            | 72,000            | -19,248           |
| Other Expense                          |                   |                   |                   |
| Depreciation & Amortisation            | 95,500            | 158,203           | -62,703           |
| NeCTAR Deprecation                     | 0                 | 342,198           | -342,198          |
| RDSI Depreciation                      | 341,034           | 660,000           | -318,966          |
| <b>Total Other Expense</b>             | <b>436,534</b>    | <b>1,160,401</b>  | <b>-723,867</b>   |
| <b>Net Other Income</b>                | <b>-383,781</b>   | <b>-1,088,401</b> | <b>+704,620</b>   |
| <b>Net Income</b>                      | <b>1,773,032</b>  | <b>2,419,455</b>  | <b>-646,423</b>   |

Source: RSM Bird Cameron audited accounts, 2013-2014

