Opening new ways to treat malaria

Tim Thwaites, Science in Public

In 2012, according to World Health Organization estimates, about 207 million cases of malaria worldwide resulted in about 627,000 deaths, mainly of young children in Africa.

We are rapidly running out of options to treat the disease. In South-East Asia, some malaria parasites have already developed resistance to the present drug treatment of choice, artemisinin.

Open Source Malaria (OSM) is a data sharing project that lets scientists pool their knowledge, ideas and expertise to find viable new drug leads. And it’s already proving effective.

In its first couple of years, about 40 contributors, from undergraduates to senior pharmacologists, have investigated four promising series of compounds. Two were discarded because they were deemed too difficult to develop into useful drugs. A third is being pursued privately elsewhere.

The fourth series of compounds, the triazolopyrazines, look very promising indeed, says organic chemist Associate Professor Matthew Todd of the University of Sydney, creator and leader of the OSM project. Todd decided to pursue this fourth series of compounds using an open source approach.

“The idea of open source is all data, all ideas, all discussions are shared online. Anyone can participate at any level. And there are no patents.”

In practice that means anyone who wants to take part can log onto opensourcemalaria.org, read Wiki pages about the project and its progress, join the team, consult the ‘to do’ list, and begin submitting data or ideas in an electronic Lab Notebook, which is searchable through Google.

The default licence for everything in the OSM project is a Creative Commons CC BY licence, which means participants can use anything lodged on the site for any reason – even to make money – provided they cite OSM.

Somewhat to Todd’s surprise, this has not discouraged people in the pharmaceutical industry from contributing. In fact, private companies have been very forthcoming.

Continued on page 8.
Boosting your research profile through your research data

Karen Visser, ANDS

Research data are increasingly recognised by both institutions and researchers as valuable assets, well worth managing, citing and publishing.

As research data become more and more easily discoverable, it has become clear there is considerable value in adding data to the impact of other research outputs to enhance the research profile of data producers.

Understanding how research data can contribute to a researcher’s profile is of interest not just to higher research degree students and early-mid career researchers, but also to those who enable data discoverability (e.g. data managers and librarians) and those interested in research impact.

It is not only the sciences that are now recognising data can be a first-class research output – humanities has rich and complex research data and research materials that are well suited to emerging publishing and metrics opportunities.

The keys to getting research data cited, discussed, counted, and closely associated with the totality of a researcher’s output (and thus profile) can include:

Tapping into new research data publishing paradigms, such as data journals. Scientific Data (Nature Publishing Group: nature.com/sdata/) was launched in May this year, and featured the dataset ‘Systematic global assessment of reef fish communities by the Reef Life Survey program’, coordinated by Australian researchers from the Institute for Marine and Antarctic Studies, University of Tasmania.

 Routinely linking data and publications. New products such as the Thomson Reuters Data Citation Index count citations to datasets and link them to their associated journal articles. A dataset from the Australian Antarctic Division (and also in Research Data Australia), ‘Boron in Antarctic granulite-facies rocks: under what conditions is boron retained in the middle crust?’ (researchdata.ands.org.au/boron-in-antarctic-granulite-facies-rocks-under-what-conditions-is-boron-retained-in-the-middle-crust) has been cited upwards of 58 times.

Making data as “open for business” as possible. A recent article by Heather Piwowar and Todd Vision indicated that publications from 2004 and 2005 were cited 30 per cent more often if their data were freely available. Also, every 100 papers with open data prompted 150 ‘data reuse papers’ within five years. (Piwowar HA, Vision TJ. (2013) Data reuse and the open data citation advantage. PeerJ 1:e175, dx.doi.org/10.7717/peerj.175).

Connecting identifiers for greater visibility in global scholarly information systems and to ensure all research outputs are correctly attributed. ORCID (orcid.org) enables researchers to connect their Scopus Author ID, Researcher ID, journal articles and datasets listed in Research Data Australia (researchdata.ands.org.au). ORCID can thus be a one-stop shop to view data in their full scholarly context. Adding a Digital Object Identifier to journal articles, data and now software has made these connections all the more powerful.

Practising data management basics that encourage discovery and reuse through metadata that fully describes reuse conditions including open licensing and ethics approvals. As the U.S. Geological Survey so succinctly put it in its document The Value of Metadata: Why Scientists Should Embrace Data Documentation: “Metadata has value for data users, data developers, and organizations. No dataset should be considered complete without accompanying metadata. Data without metadata is useless”.

In support of these desirable behaviours, over the past year ANDS has been conducting a series of interactive workshops around Australia to explore how researchers and their institutions can optimise the impact and influence of their research data products. Participants work through a series of case studies and presentations to determine the current extent of discoverability of their own data products, and then generate an action list of ways to extend their research profile using current and future research data products. These workshops are tailored to the needs of the institution and/or research group and are available either online or face-to-face (or both for distributed campuses). Email contact@ands.org.au to ask about a workshop at your institution.
Linking genes, proteins and disease: The power of data sharing

Tim Thwaites, Science in Public

When the Human Genome Project (HGP) first published the complete human genome in 2001, many expressed disappointment as to how few genes – only about 20,300 – specified the world’s most complicated animal. Even mice and mundane flowering plants boasted more.

More than a decade down the track, the misguided nature of that viewpoint has become apparent.

In fact, as genes are activated, regulated and used as templates – and the code they print off is modified in various ways – those 20,000 plus genes result in millions of different forms of proteins being made in different cells at different times.

Sequencing and documenting all that complexity produces enormous amounts of data, which are handled using powerful computers and the techniques of bioinformatics.

To cut it down to size, and make the data intelligible and useful to biologists and medical researchers, demands sophisticated software tools. With financial support and technical assistance from ANDS, Australian bioinformaticians have been at the forefront of developing such tools. Three examples are the Proteome Browser of the Human Chromosome 7 Proteomics Integration Project, the Proteomic-Genomic Nexus tool, and the Cancer Genomics Linkage Application.

The Human Proteome Project (HPP), the successor to the HGP, aims to observe all the proteins generated from the human genome.

To manage this huge task, the 24 different human chromosomes and the mitochondrion have been allocated to different countries. Australia and New Zealand are responsible for Chromosome 7, which includes genes relating to several diseases, such as colorectal cancer, in which researchers in this region are particularly interested.

With assistance from ANDS, Mr Anthony Beitz, General Manager of Monash University’s e-Research Centre, and Prof Ed Nice, Head of the Clinical Biomarker Discovery and Validation lab at Monash, led a project team to initially develop a web browser which searches a series of reputable international data repositories for any information on specific proteins linked to Chromosome 7. It ranks the reliability of the information using a traffic light system: green for good, through amber, to red for questionable.

The system has now been extended to all other chromosomes, and is applicable to any organism.

According to Nice, the openly accessible Proteome Browser not only allows researchers to find what is known, but alerts them to look for gene products or proteins that are “missing”. It even provides hints as to where they are likely to be. Finding linked proteins is the best evidence for, or validation of, specific genes.

“We are now looking at how the information can be displayed in other ways, for instance, the genes and proteins relevant to a particular disease,” says Nice.

The browser will be presented at the Annual Congress of the Human Proteome Organisation in Madrid this October.

More specifically focused is the Proteomic-Genomic Nexus, developed at the Systems Biology Initiative at the University of New South Wales. The tool allows researchers to validate their data on genes and proteins by linking results from genome sequencers, including messenger RNA (ribonucleic acid) to mass spectrometer data on proteins.

It has already been used in several interesting projects, including validating the genome of a strain of bacteria associated with Crohn’s disease and tracing the changes in proteins associated with the differentiation of bone cells.

“It allows us to ask fundamental biological questions about what is responsible for what,” says Prof Marc Wilkins, Director of the Systems Biology Initiative.

The Cancer Genomic Linkage Application lets bioinformaticians create and share workflows that provide clinical researchers with the capacity to integrate and interpret genomics data, and to search within the tumour genomes of their patients for the genes and mutations that drive cancer.

“It provides clinicians with access to resources which are most of the time in the realm of the bioinformatician,” says Dr Dominique Gorse, General Manager of QFAB Bioinformatics at the University of Queensland, where the software was developed.
Why isn't my inbox overflowing with requests for help with data sharing?

Amanda Nixon, Flinders University

Most of us are familiar with the benefits of sharing research data and will know it’s a good thing.

ANDS has created some excellent resources for Australian researchers showing exactly how to share data. Australian research institutions are for the most part ready to assist. Some research funders and journals require data sharing, and datasets may now be cited. So why isn’t my inbox overflowing with requests to help with data sharing?

In my role at Flinders University, I act as a data sharing advocate and spend a lot of time talking to researchers about their data. In those conversations, we cover the “pros” of data sharing from a researcher’s point of view:

» Data sharing treats the dataset as a research output (which is another way of recognising the hard work of creating it in the first place!)

» It formalises any informal data sharing researchers may already do, and provides the opportunity of recognising the use of the data through citation

» It may prevent duplication of a data collection

» It may enable comparative studies

» It’s good science – that is, it allows for greater transparency of research results

» It’s a good habit to get into because some research funders require it anyway, as do some journals.

Many researchers agree with at least some of the above in our conversations, and yet the number of datasets I deal with would be described as a steady trickle rather than an overwhelming demand.

Drawing from those conversations, I think it is not so much that there are startling “cons” for sharing data that are making researchers reluctant to move quickly, but there are barriers – real or perceived – preventing data being made available.

**Funders** – The funder of the research may not allow data sharing.

> This landscape is changing. Internationally, high profile funders, such as the Bill and Melinda Gates Foundation, make data sharing a requirement.

**Research project** – The project may be cross-institutional and responsibility for the data may not be clearly set out.

Researchers often have many competing priorities and sharing data does not feature highly.

> This can be solved by doing the groundwork to create a data management plan at the start of a project – and ensuring all collaborators support it.

**Researcher** – As collectors of the data, researchers often wish to gain as much value out of the dataset as possible before sharing it. In other cases, some researchers cannot see any value in making the data available because they feel no one would be interested.

> For confidential data, there may be value in investigating options to de-identify the data. For data requiring costly review, it may be a case of considering sources of funding to undertake the work. A case could be made, considering the value of sharing the data.

**Sensitivity of the data** – The data may need to remain closed because they are clinical, sensitive, confidential, personal or have potential for commercial use. Or they may require review, correction, deduplication or verification before being shared and the cost of doing this is prohibitive.

> People can misinterpret or misuse any information, including traditional research publications. This is not unique to research data. On the other hand, what if sharing the data led to a new, high profile collaboration? What if the data were cited by colleagues worldwide? What if, as a result of this, you were able to find your own data again after ten years?

Clearly in some cases, the nature of the data – the funding that enabled the collection or the conditions under which they were collected – means they cannot be shared under any circumstances. However, most other objections can be overcome with a good dose of data sharing advocacy and support as well as data management planning at the outset of the project.

I think the numbers in my inbox might be about to go up!

Amanda Nixon is Manager of eResearch@Flinders
After 30 years of research, with possibly the most significant work of his career yet to be completed and published, Prof Richard Maltby of Flinders University realised he needed a back-up plan. To preserve his legacy – namely his work investigating early American cinema – Maltby decided to share his research data, all 30 painstaking years of it.

In 1984, New York, Maltby gained access to a treasure trove of early Hollywood correspondence from the Motion Picture Producers and Distributors of America (MPPDA) and has spent the interim years preserving this information for current and future generations. Established in 1922, the MPPDA was the trade association for the US’s motion picture industry, later called the Motion Picture Association of America (MPAA).

In 1965, the MPAA copied the MPPDA’s General Correspondence files onto 16mm microfilm, and threw away the originals. The archive was then stored in the MPAA’s New York offices, where it was largely unavailable to researchers, with no inventory of its contents made.

The MPPDA’s General Correspondence files represent a rich source of information about the political history of the industry and a number of its key activities and decisions. They include extensive correspondence and other documentation relating to industry policy, public relations, trade practices, external censorship and self-regulation. The great majority of this material is not available from other sources.

For further study, Maltby was allowed to copy 12 of the archive’s microfilm reels, covering the period 1922–1939. Many of these reels have since been lost and Maltby’s copies are all that’s left of much of the archive’s material.

With the help of his wife, Dr Ruth Vasey – a Senior Lecturer in Flinders University’s Department of Screen and Media – Maltby built a catalogue of the archive and created the Motion Picture Producers and Distributors of America Digital Archive (MPPDA Digital Archive).

With the aid of grants from Flinders University, the Australian Research Council and ANDS, a frame-by-frame digital copy of the archive was made. This has led to the indexing of about 35,000 pages that have been further categorised into nearly 3500 records, each describing a document or cluster of related documents.

The archive also contains biographical information on the leading MPPDA characters and organisations, as well as information on 671 films.

The MPPDA Digital Archive has been described as “a model piece of database construction... It is an extraordinary achievement and a huge boon to moving image research, for the silent era and beyond.” ([thebioscope.net/2012/05/09/]

“The project was first to make [the data] useable for us,” admits Maltby. “But what the ANDS project has done is to make it available for any researcher who wants to use it.”

The MPPDA Digital Archive* (mppda.flinders.edu.au) is a public, user-friendly web portal. Users can view records in low resolution or download full versions by completing a simple, free online registration process.

There are currently more than 500 registered MPPDA Digital Archive users spread worldwide, with 175 universities and professional organisations represented.

Alongside research, the archive is also used for teaching purposes at several universities, such as DePaul and St Joseph’s in the US.

“I just hope everybody cites the material properly. In a sense, that’s the return I get,” says Maltby.

“One would hope that before too long... the kind of work that’s involved in producing humanities data in a form that other people can use, will receive acknowledgement as research in itself.”

Maltby says the question of ‘Do you keep the data to yourself or make it available to others?’ is an important one.

“Clearly individuals will make different decisions, but if we are engaged in a common enterprise – the creation of knowledge – then that should be what drives us to make the material available.

“In the past, the best way to make it available was by digesting information and publishing it in an accessible form, i.e. via books and articles in journals. Now we can do something else as well – we can make the source material available.

“While I’m still working on my book [based on the MPPDA data], I’m not locking up the material and saying no one can use it. Instead, I’m sharing and that’s a better way to practise.”

According to Maltby, humanities researchers need not worry about sharing data before publishing and having someone else beat them to the finish line.

“It’s not as if there’s a discovery, it’s not like racing someone to

Continued on page 8.
Researchers can produce a bigger impact from their research by sharing their data, but this requires more mature data management practices. Data are valuable. Data are a resource to be exploited. They can lead to new collaborations, new publications and new grants. Here are five ways to create value from data:

1. Shared open data means more citation: share data

The internet has changed how we do research. Researchers can easily find, extract, analyse, and share research data like never before. Such potential has found a voice in the ‘open data’ movement that is affecting both government and research data.

For government, sharing means less expense in fulfilling ‘Freedom of Information’ requests.

For a researcher, open data means better connections into the international research conversation, and leads to more citation. See research documenting the benefits of sharing data on citation rate at Piwowar HA, Vision TJ. (2013) Data reuse and the open data citation advantage. PeerJ 1:e175, (dx.doi.org/10.7717/peerj.175).

Sharing data can also help researchers reach out to industry through bearing gifts of knowledge.

Sharing data is of course subject to ethical, legal and cultural sensitivities, and while a work in progress, like software, data can be released progressively to show the movement along the research journey.

2. Scientific data are not something to be hidden away: open data

Opening data means making it useful to someone else. This means describing and documenting your methodology, process, and such minor things as your local abbreviations to allow other researchers to make sense of your data.

Data, alongside journal publications, conference papers, and other presentations are an important research output. But data are becoming the new way to separate a researcher from their peers.

Set your data free through opening and sharing them publicly (where appropriate; considering ethics and intellectual property) using services like figshare (figshare.com), ZENODO (zenodo.org) and the Australian Data Archive (ada.edu.au).

Publishing your data, particularly with a DOI [Digital Object Identifier, e.g. from DataCite: datacite.org] attached, gives them a chance to be cited, alongside your other written work.

3. Published data are a publication: publish data

Once your data are publicly available, and citable, they can justifiably sit alongside your journal articles and conference presentations as part of your research output.

You can include lists of your published data in your CV, on LinkedIn, on job and grant applications, and on institutional and personal web pages. Your data become another publication, which showcases your work.

4. Contribute a verse to the global research conversation: talk data

The internet enables access to other’s work and provides a platform for presenting your contribution.

As a researcher, you are a publisher as well as a retriever and reader. You are a producer and consumer. As Walt Whitman so eloquently said, “That the powerful play goes on, and you may contribute a verse”.

In the internet age, you have a voice for talking about your research. And by sharing your data, you will stand out from the crowd.

5. Applying for ARC grants now means having a data management plan: plan for data

In early 2014, the Australian Research Council (ARC) took further steps towards recognising the importance of data in the global research conversation.

The ARC requested in its Discovery Grant and Fellowship applications that researchers include an outline of their plans for managing generated data.

Particularly the ARC asked for an outline of plans for “the management of data produced as a result of the proposed research, including but not limited to storage, access and re-use arrangements.”

Getting into the habit of planning, managing and sharing data effectively will set a track record that is likely to impress when it comes time to be judged on grant applications.

In conclusion, opening, sharing, publishing, talking and planning for data creates value for researchers. Such actions will give new researchers in particular an edge over their peers, and a fast track to catch up with their professors.

Sharing data will lead to more citations, strengthen grant applications and enable researchers to contribute a new verse to the global research conversation.

A version of this article is due for publication in The Thesis Whisperer.
Sensitive data are frequently excluded from discussions about data publication and sharing in research and data management communities.

Sensitive data are data that can be used to identify an individual or object and place them at risk of discrimination or unwanted attention. Familiar examples include: patient health records, personal information alongside religious, cultural or political affiliations, and location of endangered species or conservation efforts.

In the past, these data have often been relegated to the ‘too hard basket’ under the assumption they cannot be shared legally or ethically, or it is too difficult to do safely.

These ‘too hard basket’ beliefs are now being challenged and frequently overcome.

Methods to ‘de-sensitise’ data are more widely understood and used. These include removing aspects of a dataset that can be used to identify individuals or objects and otherwise modifying original data so they are no longer sensitive.

The power of appropriate licensing is also better understood. Researchers are applying licences to their data to place conditions on its reuse and prevent misuse, and to protect themselves.

Several influential funding bodies and publishers are now requiring or incentivising researchers to publish and share their data – sensitive or not.

Australian researchers and government bodies gather comprehensive and high-quality data that contain sensitive information.

By their nature, sensitive data are amongst the most expensive and time-consuming to collect and manage. Greater exploitation of these data via sharing and reuse is thus consistently called upon by Government* for reasons of finance, expediency and participant fatigue. Benefits are also passed on to the original research team by way of collaborations, citations, and funding opportunities.

There is a widespread belief amongst researchers that they always require explicit approval from an institution’s Research Ethics Committee in order to share data about people. This is frequently the most common and pertinent issue for researchers who collect and use sensitive data where clear institutional guidance is not available or sought.

The best option is to request participants’ consent to share their deidentified data before the data are collected. If this can’t be done because the data were collected in the past, it is not necessarily the case that it can’t be shared at all.

It is all about how the data are treated for the purpose of ‘de-sensitisation’, how the data are licensed and in some cases, how access might be mediated.

In some cases the data can simply be licensed using one of AusGOAL’s Creative Commons licences. In others, an AusGOAL restrictive licence might be called for. Or, in the case of very sensitive data, the best strategy might be to publish a metadata description of the data only, and potential users can then contact the researcher and discuss use on a case-by-case basis.

In summary, sensitive data may be shared when certain conditions are met. And the benefits of doing so may well far outweigh the costs or concerns.

Data often considered too sensitive to share are a highly valuable resource in the Australian research community. This resource remains largely untapped due to a lack of clarity around – rather than true constraint by – legal and ethical obligations.

ANDS is currently consulting with legal and research ethics experts to produce a user-focused guide to the publication and sharing of sensitive data. The Sensitive Data Guide will contain clear advice about when and how such data can be shared, alongside case studies. The guide is due for public release via the ANDS website (ands.org.au) later this year.

Dr Sarah Olesen is an ANDS Senior Data Management Specialist and a Research Fellow at the Australian National University. Dr Greg Laughlin is ANDS’ Principal Policy Advisor.
Opening new ways to treat malaria

The starting point for the project was a stock of compounds released into the public domain in 2010 by GlaxoSmithKline (GSK). Another batch has since come from Pfizer. And, at present, a GSK laboratory is screening the latest compounds from the project for antimalarial activity.

Beyond the fact that OSM provides interesting work of humanitarian significance, there are probably several reasons for the involvement of industry personnel. "Not least is PR," says Todd. "They can also publish their work in scientific journals. And because interactions happen in real time, companies can show how good they are, how quickly they can solve things."

The benefits, however, do not come without cost – for instance, the time involved in organising the project. Todd has recently spent some months developing a system to label, reference and number the new compounds being developed. "It was a big job, but we had to do it, to allow people to navigate the project." But that time is not wasted. The outcome will be used in future open source projects.

In fact, OSM builds on Todd’s first open source project, which successfully developed a more acceptable form of the drug currently used to treat bilharzia, a nasty parasitic disease that affects hundreds of millions of people in Africa and the Middle East.

Apart from determining what components, rules and software a successful open source project needs, he learned a particularly valuable lesson in psychology. "To get people to contribute, you have to bring something to the table, do something yourself."

But for Todd, the time spent and lessons learnt in setting up open source projects has become a labour of love. "This is what science should be. This is what I signed up for. I find it very, very satisfying."

Also, he says, because the interactions are happening in real time, there is the excitement of being at the cutting edge of research as it happens.

Having pioneered the open source approach, Todd has built a reputation. Computer entrepreneur and philanthropist Bill Gates was so impressed with Open Source Malaria, he tweeted to his 16 million Twitter followers a link to a YouTube video in which Todd describes the project (youtube.com/watch?v=gCOokjOiVTo). And an article Todd wrote for Nature Chemistry describing how the bilharzia project was organised has become one of that journal’s most widely consulted.

Now he wants to go further. So far, the projects tackled using the open source approach have involved drugs from which the pharmaceutical industry would not expect to make big profits. Todd is now involved in international discussions about how to apply the approach to cancer, a condition for which development of effective drugs can be very profitable indeed.

Build your own virtual mausoleum

"I’ve been trying to sell the argument to people coming towards the end of their career, that they shouldn’t waste the work behind their publications. Now they have the chance to curate that work by putting it into a repository. Their data may also have a use for other people, perhaps a similar use to what they have for it.

“Leaving it on the hard drive for your grandchild to take to pieces after you’re dead is not a very good use for it. Doing something with it – putting it in a university repository for instance – is like building your own virtual mausoleum."

*The MPPDA collection is also available via Research Data Australia: researchdata.ands.org.au/motion-picture-producers-and-distributors-of-america-incorporated-database*
Data storage: A place for everything and everything in its place

Richard Ferrers, ANDS

Storage is an essential element of getting the best value out of research data, and ANDS is often asked for advice on this important topic.

The first of the four ANDS data transformations (unmanaged to managed) essentially assumes the data are going to be stored somewhere appropriate. And so ANDS cares about storage.

While not a service within the scope of ANDS, ANDS works closely with its research infrastructure partners whose aims do include storage. These include NeCTAR (tools), RDSI (storage), and state-based eResearch service providers, who often also operate RDSI funded storage nodes.

ANDS staff work closely with the state-based RDSI (Research Data Storage Infrastructure) nodes, such as Intersect in Sydney, eResearch SA in Adelaide, and QCIF (the Queensland Cyber Infrastructure Foundation) in Brisbane.

ANDS also works closely with research institutions, such as Monash University, Deakin University, the University of Technology, Sydney and the University of Western Australia – all of which have implemented data storage alongside metadata store (data description registry) projects. Links are available to these and other universities’ research data infrastructure projects from the ANDS Projects web page: projects.ands.org.au/getAllProjects.php?start=all.

Where researchers store data is dependent on several issues. Working data are often stored locally while collected and analysed. When results are published, then it is useful and sometimes required to submit data to a discipline or institutional repository, where they can be linked to a permanent identifier, like a DOI (Digital Object Identifier).

High value data, particularly data of national and long-term significance, may be stored in nationally-funded data centres, such as those operated by the nodes, which have high reliability and capacity.

The choice between data storage location is affected by several factors, including: discoverability; storage size; cost; convenience; access for research partners and collaborators; and access to bibliometrics and curation services. See more about this in the ANDS Guide to Storage (ands.org.au/guides/storage.html).

More and more institutions are also providing recommendations about types of storage researchers should use. See for instance Deakin’s advice on storage type advantages and limitations at: deakin.edu.au/research/eresearch/manage-data/store.

In the new highly connected digital era, online research storage is a critical component, providing an important foundation for sharing, preserving, collaborating over and extracting the most value from data for researchers.
In my last article I spoke of the significant changes to Machinery of Government and their impact on the Research Infrastructure team with whom we work.

I am very pleased to be able to report that the exceptional efforts of the team, now in the Department of Education, led to a major positive outcome in the May Budget statement. This is how it was reported to capability directors by the Department via Ditta Zizi, Branch Manager, Research and Higher Education Infrastructure:

"Tonight, as part of its budget, the Australian Government announced additional funding of $150 million in 2015–16, to continue the National Collaborative Research Infrastructure Strategy (NCRIS). The budget announcement is available at budget.gov.au or education.gov.au. This is a fantastic outcome and the funds will support the operation and maintenance of existing research infrastructure facilities covered by NCRIS… Finally, I would like to thank you all for your support in keeping the NCRIS flag flying! I would also like to thank my fantastic team, as without them this outcome would not have been possible."

This is a fantastic outcome and the funds will support the operation and maintenance of existing research infrastructure facilities covered by NCRIS. Finally, I would like to take this opportunity to thank you all for your support in keeping the NCRIS flag flying! I would also like to thank my fantastic team, as without them this outcome would not have been possible.

The months ahead represent an exciting period of hard work and opportunity for the ANDS team in vigorously pursuing the challenging goals that have already been established, as well as engaging constructively in planning for the future of research data infrastructure under "NCRIS reborn".

I would like to conclude by expressing my gratitude to [ANDS Technology Director] Andrew Treloar for taking over the reins of ANDS while [Executive Director] Ross Wilkinson is taking some exceptionally well-earned R & R cycling through France. Andrew has recently returned from the Research Data Alliance Third Plenary in Dublin for which he led the organising committee. I am reliably informed that the Plenary was a stunning success due in no small part to Andrew’s efforts. The international reputation of ANDS and Australian research data infrastructure has never been higher. Congratulations to all the ANDS staff who participated in this event.
Feeding the world? It's all about data!

Shannon Lindsay, ANDS

Science and shared data are the keys to food security for a projected population explosion of 9bn people by 2050, said Australia’s Chief Scientist Ian Chubb in his keynote speech at the Research Data Alliance Third Plenary in Dublin on 26 March 2014.

“[Australia’s] frequently articulated objective of being a [global] food bowl is an issue for science and it will generate data and the data will need to be shared,” said Prof Chubb.

The global sharing and collaboration over research data had changed research practice substantially in the last decade, noted Chubb, but this was essential: “Solutions to these problems cannot be achieved by one country or researcher alone. Global challenges will only be solved by global research endeavour.”

But he warned: “Countries cannot expect access to data without being contributors to the data pool.”

“Whether the global challenges are about managing our food and water assets, saving our biodiversity, predicting high impact weather disasters or addressing health issues, it’s all about data.

“We need national and international, intradisciplinary and interdisciplinary collaboration around data to put the jigsaw together.

“Our increasing connectedness as scientists and researchers will also allow us to advance the boundaries of discovering beyond anything that has ever been achieved before.”

Chubb said Australia has long recognised the importance of research data. “The establishment of the Australian National Data Service, which is the Australian non-government representative in the Research Data Alliance, formalises that recognition.”

The Alliance, he said, was established by “future planners [including ANDS] who realised we needed supporting infrastructure to enable collaboration, but that underlying that infrastructure must be a will to collaborate and a will to share and a forum to allow this all to happen.

“We’ve got to put aside the historical way we went about doing things – locking it [data] up, sticking it on arcane IP [intellectual property] rules, tying people in knots when they tried to collaborate...”

The collaborative response to global challenges isn't possible unless we get [research infrastructure] fundamentals right first, and one of those fundamentals is sharing high quality research data.”

Chubb’s comments were made before 500 international scientists, researchers and data management practitioners attending the Research Data Alliance Third Plenary to discuss global data sharing issues.

The Alliance was set up as a joint activity between Australia, the US and Europe, and is working to build the social and technical bridges that enable open sharing of data.

The Third Plenary was jointly organised between Australia (through ANDS) and Ireland (through the Digital Repository of Ireland), and representatives of both governments opened the conference, including the Australian Ambassador to Ireland, Her Excellency Dr Ruth Adler.

Key plenary sessions included examining the importance of data sharing for the development of agriculture globally, marine data management, and how to address the needs of the overall scientific community in collaboratively utilising large volumes of data.
Forthcoming events

Webinar: Dr Virginia Barbour presents
PLOS: open data, ORCIDs and Article Level Metrics

When: Thursday, 3 July, 12:30pm–1:30pm AEST
Where: online
What: The PLOS Medicine co-editor will discuss PLOS’ new data policy, the introduction of ORCID Identifiers and the development of PLOS article level metrics.
Register: www4.gotomeeting.com/register/175378023

Online: Research Data Management catch up

When: Thursday, 10 July, 3pm–4pm AEST
Where: online
What: A regular informal forum for colleagues in small to medium-sized or regional universities, working to enable stronger research data management practices, to network and talk through some common goals or challenges.
No need to register: www4.gotomeeting.com/join/273639063
Note: This is a monthly event. The next is Thursday, 14 August, 3pm–4pm AEST.

Workshops: Using Research Data to Raise Your Research Profile

When: Wednesday, 30 July, 9am–11am; repeated 2:30pm–4:30pm
Where: Morning – University of Tasmania, Sandy Bay and via videoconference from other UTAS campuses.
Afternoon – IMAS, Battery Point, TAS
What: This workshop will explore new opportunities and traditional and social media tools available to researchers to raise their academic profile using research data.

Webinar: TERN’s Dr Anita Smyth presents – ‘Can sharing research data raise your research profile?’

When: Thursday, 7 August, 12:30pm–1:30pm AEST
Where: online
What: Dr Smyth, researcher and Data Facilitator with the Terrestrial Ecosystem Research (TERN) Eco-informatics Facility, will talk about the growing body of evidence that suggests social media is an effective way to raise your research profile.
Register: www4.gotomeeting.com/register/231789071

Open Access Week 2014

When: 20–26 October 2014
What: A global event, now in its 7th year, promoting Open Access as a new norm in scholarship and research.
Participate: www.openaccessweek.org

eResearch Australasia 2014

When: 27–31 October 2014
Where: Melbourne
What: The eResearch Australasia conference brings together practitioners and researchers for a stimulating week to share ideas and exemplars on new information centric research capabilities. ANDS will see you there!
Register: conference.eresearch.edu.au/eres2014/registration/

For more information and the full list of events visit: ands.org.au/events/index.html

Get notified about our forthcoming newsletters via RSS feed: ands.org.au/newsletter
For more news, alerts, announcements and discussion subscribe to the ANDS General Google group by emailing: contact@ands.org.au
To learn about the upcoming ANDS events visit: ands.org.au/events
Join our conversation on Twitter: @andsdata

ANDS is supported by the Australian Government through the National Collaborative Research Infrastructure Strategy program and the Super Science Initiative.
This newsletter is designed by GRIT Brand + Online (gritcomms.com.au)

ANDS Project Partners:

With the exception of logos or where otherwise indicated, this work is licensed under the Creative Commons Australia Attribution 3.0 Licence.