Research Data Infrastructure
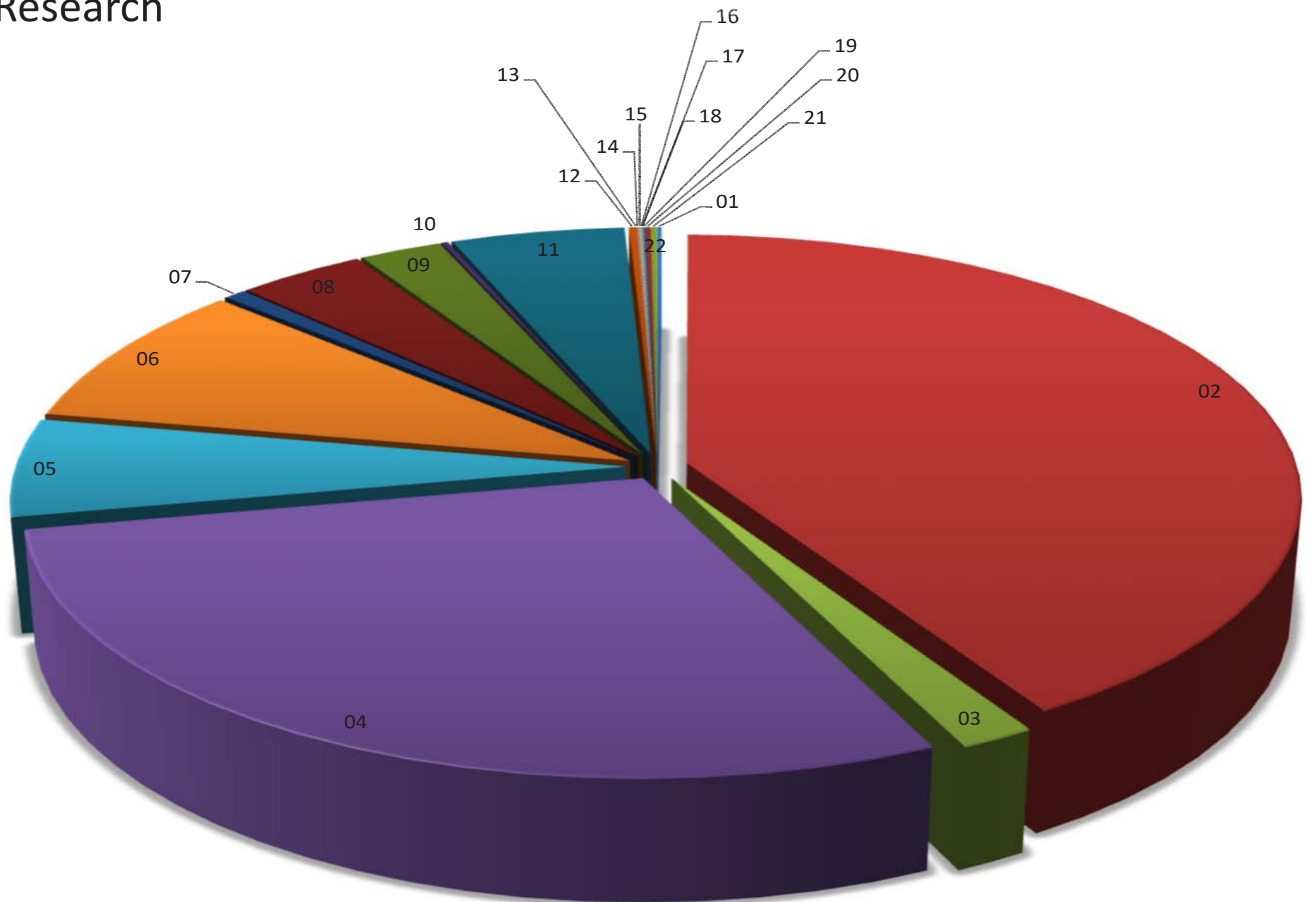
60PB of installed capacity - 42PB in use

Providing high capacity digital data storage for nationally significant collections, a merit allocation scheme, and collection curation and management resources.

Delivered by eight partners around Australia

	eRSA	Intersect	NCI	Pawsey	QCIF(BNE)	VicNode	QCIF(TSV)	TPAC
	1096	989	10296	17552	3851	4645	1149	2654
	3987	1546	11118	18746	13862	6085	2423	2395

	Primary Nodes					Additional Nodes	

Total: TB Ready		42231	Total: TB Approved		60162

A national view of research data collections

Data collections spanning over 22 Fields of Research

- 01 MATHEMATICAL SCIENCES - 0.13%
- 02 PHYSICAL SCIENCES - 41.22%
- 03 CHEMICAL SCIENCES - 1.76%
- 04 EARTH SCIENCES - 28.82%
- 05 ENVIRONMENTAL SCIENCES - 6.06%
- 06 BIOLOGICAL SCIENCES - 8.90%
- 07 AGRICULTURAL AND VETERINARY SCIENCES - 0.65%
- 08 INFORMATION AND COMPUTING SCIENCES - 3.83%
- 09 ENGINEERING - 2.46%
- 10 TECHNOLOGY - 0.19%
- 11 MEDICAL AND HEALTH SCIENCES - 5.18%
- 12 BUILT ENVIRONMENT AND DESIGN - 0.20%
- 13 EDUCATION - 0.02%
- 14 ECONOMICS - 0.07%
- 15 COMMERCE, MANAGEMENT, TOURISM AND SERVICES - 0.06%
- 16 STUDIES IN HUMAN SOCIETY - 0.03%
- 17 PSYCHOLOGY AND COGNITIVE SCIENCES - 0.02%
- 18 LAW AND LEGAL STUDIES - 0.01%
- 19 STUDIES IN CREATIVE ARTS AND WRITING - 0.07%
- 20 LANGUAGE, COMMUNICATION AND CULTURE - 0.13%
- 21 HISTORY AND ARCHAEOLOGY - 0.19%
- 22 PHILOSOPHY AND RELIGIOUS STUDIES - 0.00%

ardc
australian research data commons
Built from ands · nectar · RDS

RDS
Research Data Services

Research Community projects

10 priority research domains

The Research Data Services (RDS) project supported a number of research communities from around Australia. Services were provided through partnering Node Operators, which RDS enabled to carry out work and engage with researchers. Supported domains included:

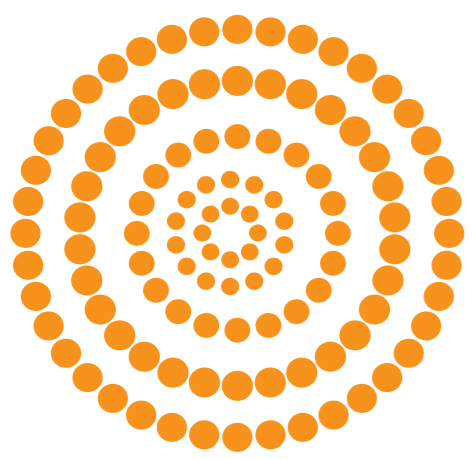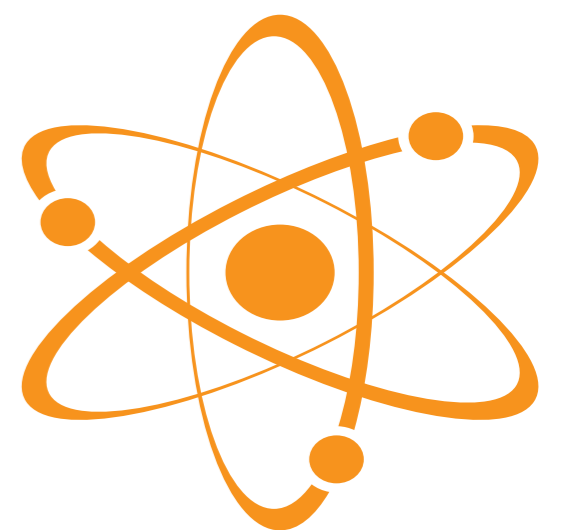|				
Earth Systems	Life Science	Medical and Health	Image Characterisation	Astronomy
Cultures and Communities	Terrestrial Systems	Marine	Geoscience	Geophysics
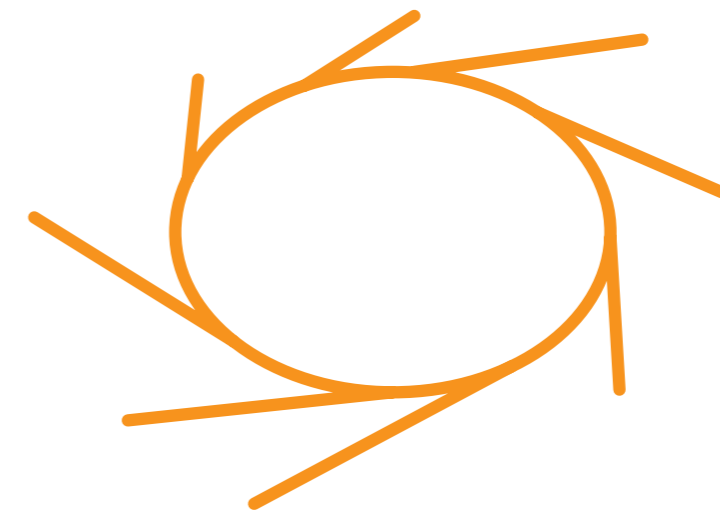
Flagship projects

Researtch Data Services engaged the community in system-wide projects to act as exemplars and technology drivers for additional data value generation.

Health Flagship - delivering an exemplar best-of-breed data curation and management desktop by leveraging the data management requirements of the BPA Omics Project

Moving Big Data 1 - optimising high-speed data shipping nationally and internationally both within the node community and to collaborating parties

Moving Big Data 2 - facilitating a national and federated view of collections, spanning infrastructures and enabling broader discoverability and re-use.

Leadership - concentrating on eResearch system-wide benefits such as user support, security, and architecture planning.

Omics Platform - cloud-based data services and tools for Australian Life Science Researchers to combine, analyse and interpret genomic (DNA), transcriptomic (RNA), proteomic (proteins) and metabolomic (small molecules) data.
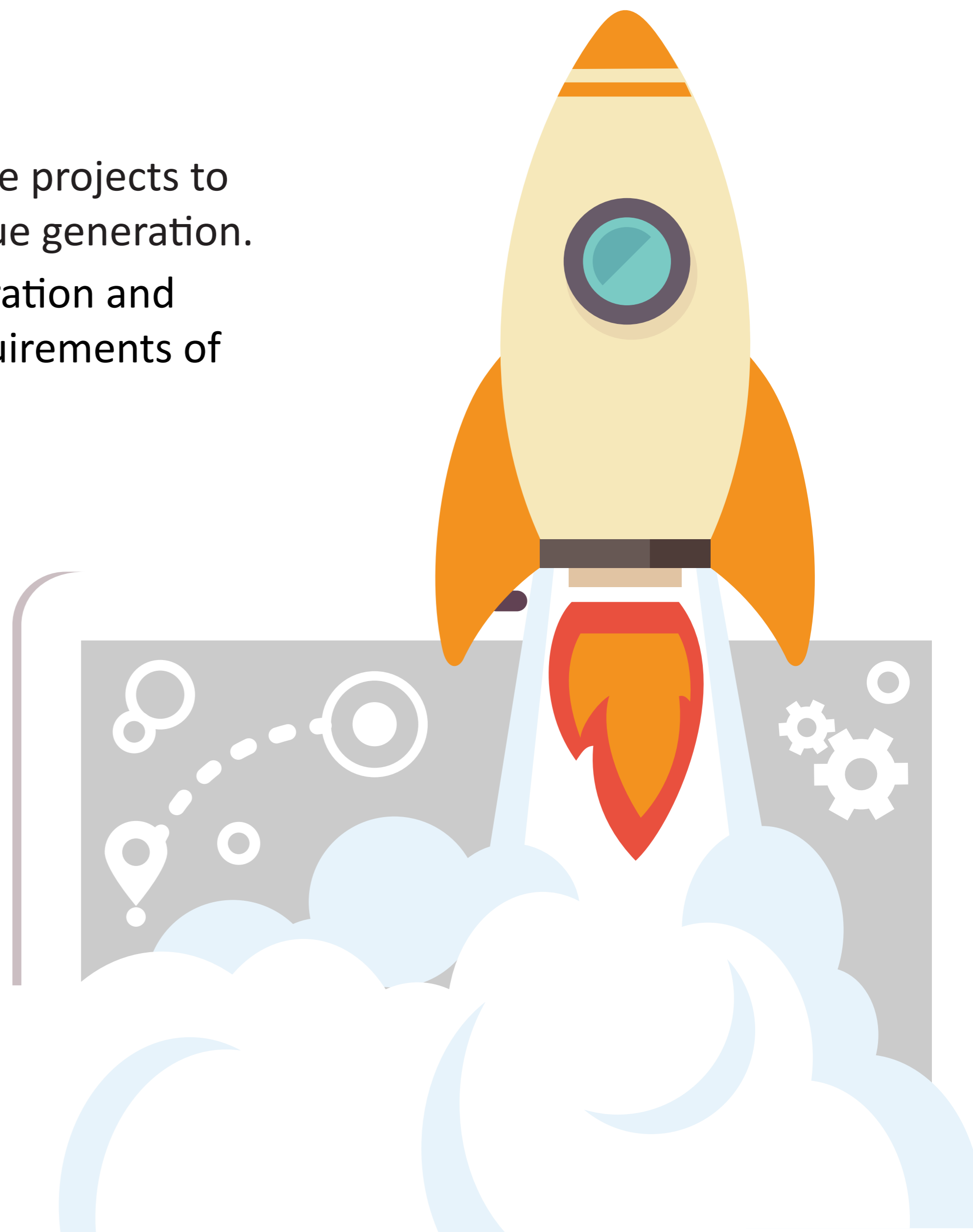
Virtual Laboratories

Domain focused

Digital interfaces, tools and data to online research communities

By logging into an online Virtual Laboratory interface, researchers can access both data and cutting edge tools from their own computer's web browser, instantly collaborating with colleagues around the world.

|						
Characterisation Virtual Laboratory	**Industrial Ecology** Virtual Laboratory	**Virtual Hazards, Impact and Risk** Virtual Laboratory	**Endocrine Genomics** Virtual Laboratory	**Virtual Geophysics** Virtual Laboratory	**Climate and Weather Science** Virtual Laboratory	**Genomics** Virtual Laboratory
Provides a remote desktop environment, run on the Nectar cloud, with access to a range of tools to process atom probe, neuroimaging, structural biology, X-ray and general imaging data.	Provides a platform for environmental footprint and life cycle sustainability assessments (LCSA)	A scientific workflow portal that provides natural hazard researchers with access to an integrated environment.	Provides clinical/biomedical data resources supporting a wide range of clinical research groups across Australia and globally.	A scientific workflow portal that provides Geophysicists with access to an integrated environment that exploits eResearch tools and Cloud computing technology.	Provides an integrated platform of tools and resources that assists scientific collaboration and information-sharing in Australian climate and weather research.	Provides a cloud-based suite of genomics analysis tools for research and training.
Marine Virtual Laboratory	**Alveo** Virtual Laboratory	**Biodiversity and Climate Change** Virtual Laboratory	**All-Sky Virtual Observatory** Virtual Laboratory	**Humanities Networked Infrastructure** Virtual Laboratory		**Microbial Genomics** Virtual Laboratory
Enables researchers interested in marine modelling access to an online portal that simplifies the process of marine modelling.	Provides human communication science researchers all over Australia with infrastructure to store data collections, analysis tools, and workflows.	A "one stop modelling shop" that simplifies the process of biodiversity-climate change modelling.	Provides a direct and vital link between the theoretical and observational aspects of data collection and analysis.	Combines data from many Australian cultural websites into the biggest humanities and creative arts database ever assembled in Australia.		Provides a computer image based around Galaxy with a defined set of microbial-specific tools and reference data

Value

Enabling innovation and collaboration

"The tools for automated handling and storage of data provided by Characterisation Virtual Laboratory [CVL] are a key enabler for researchers to derive the most benefit from these large datasets."
Dr Keith Schulze, Monash University.

Saving time, money and resources

"The speed at which we can now carry out our geophysical inversions [using the Virtual Geophysics Laboratory] was not possible before. It removes all of the pre-processing that used to take us days to complete."
Dr Carina Kemp,
Geoscience Australia.

Virtual Laboratories are generating a return on investment of up to **138 times** their cost, according to a report by Victoria University academics.
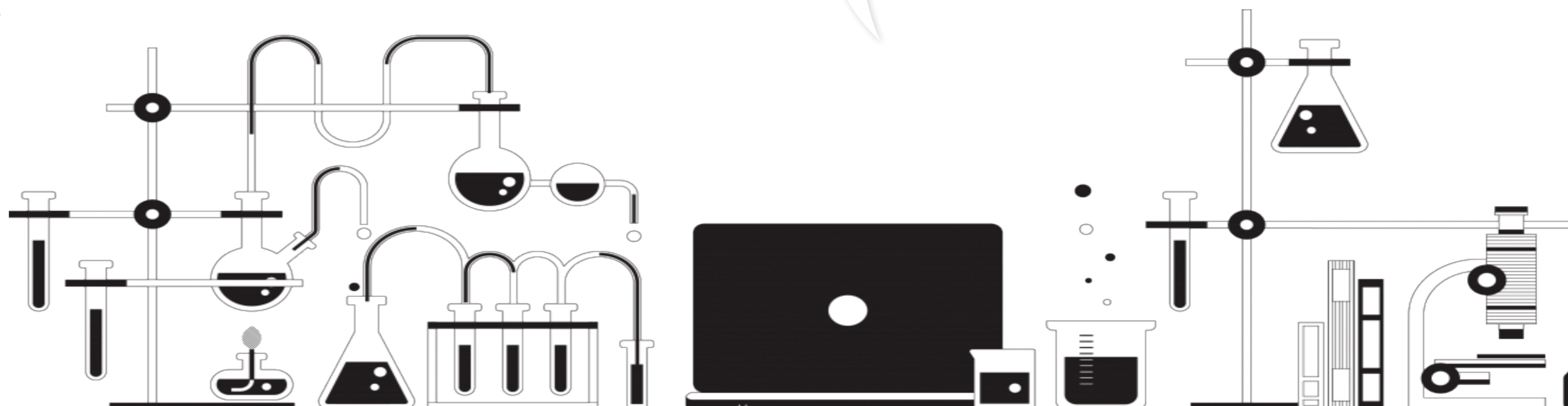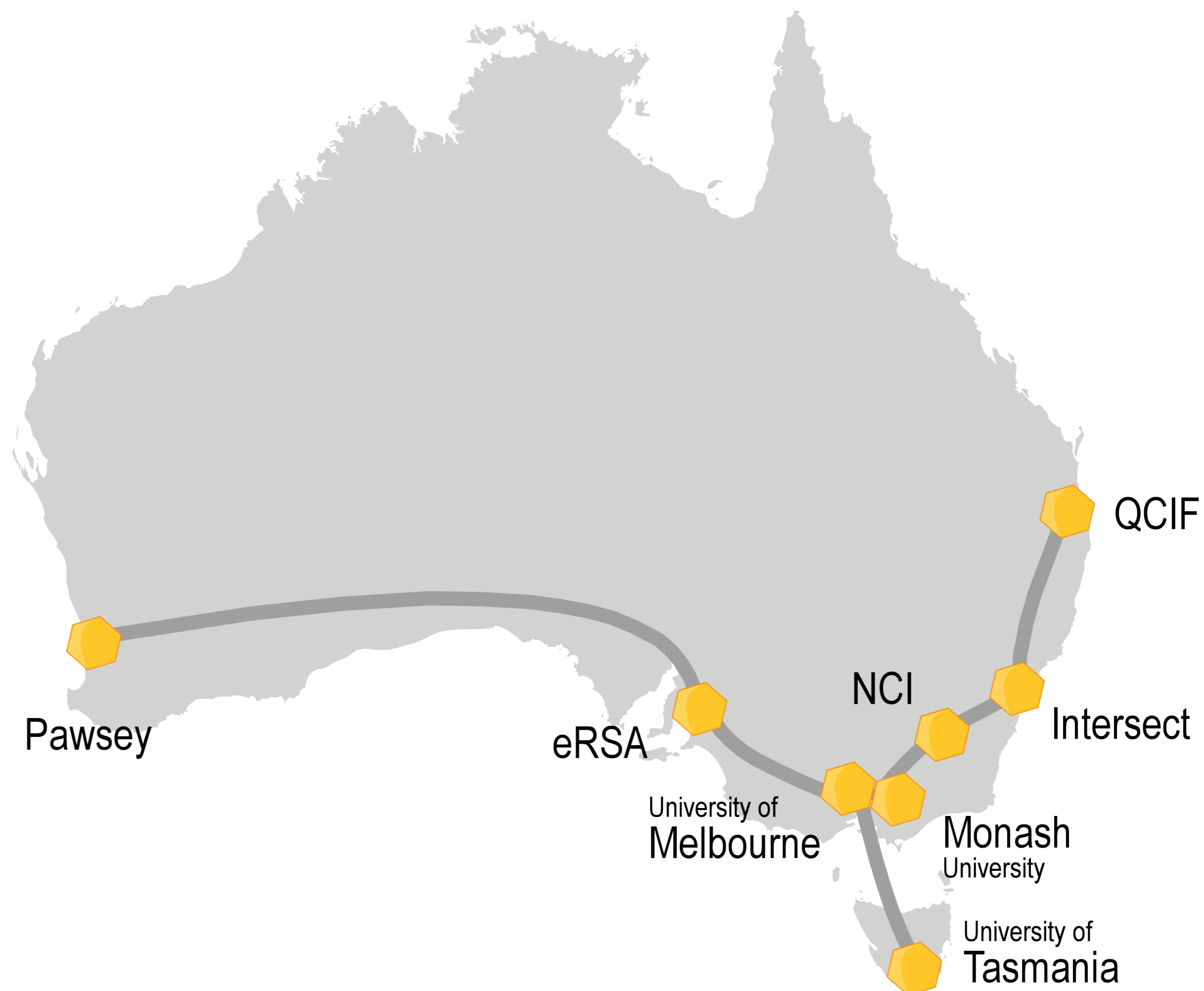
Research Cloud

A world first...

The Nectar Research Cloud is a partnership between eight institutions and research organisations who are deploying and operating Australia's first federated research cloud.

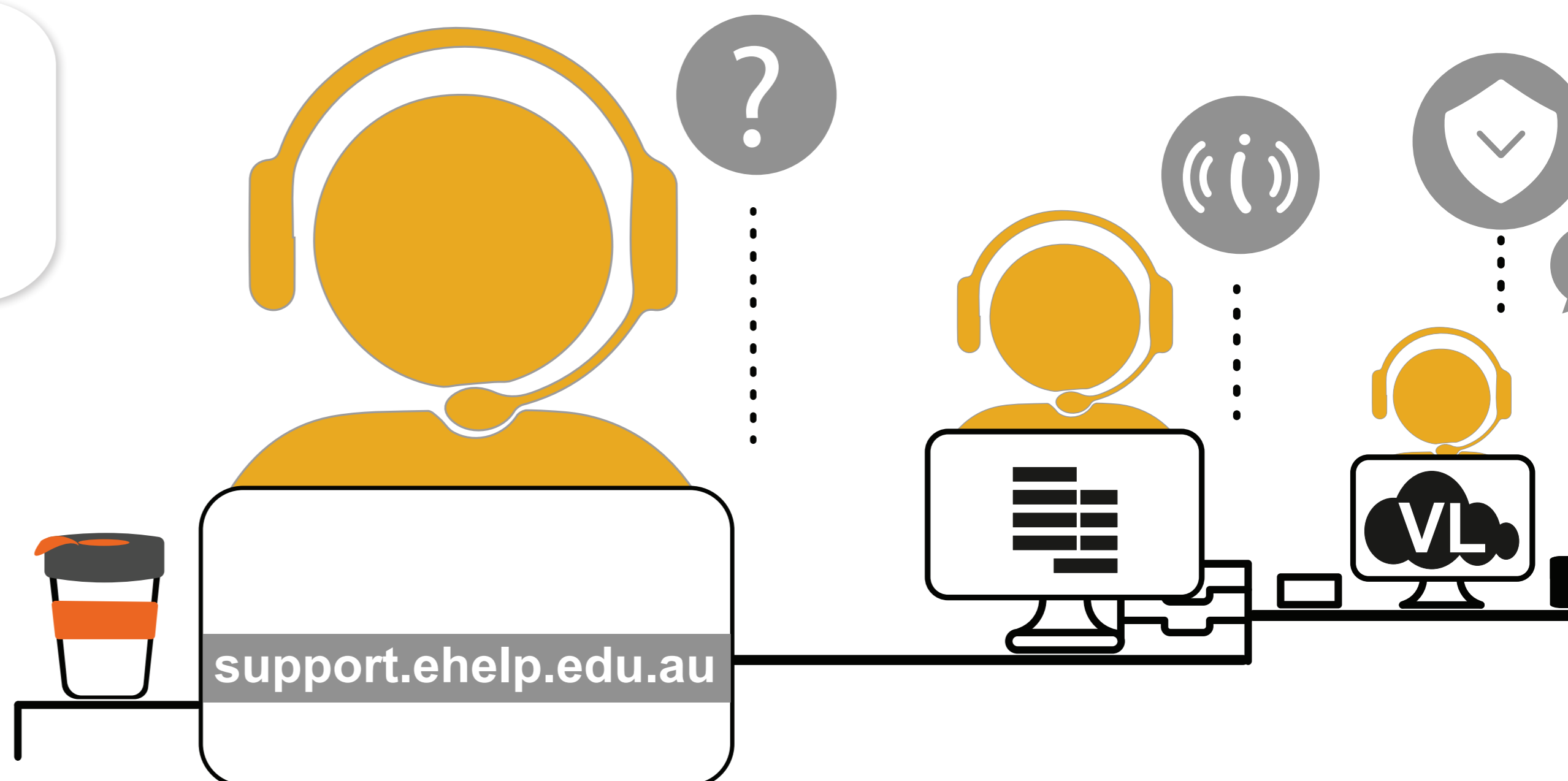Providing Australia's research community with:

- Computing infrastructure, software and services
- The ability to store, access, and run data, remotely, rapidly and autonomously.

Outcomes

- Reached **10,000 registered users** in March 2017.

- Over one million virtual machines were spun up to support research across the **22 fields of research** (all 2-digit FOR codes).

- Projects vary in size from a single Central Processing Unit (CPU) through to **1,000+ cores**.

- Users describe the resources they host in the cloud as *"incredibly useful"*, *"brilliant"*, *"the backbone of the project"*, *"a life-saver"*.

- Access to the cloud enhances innovation in research by allowing researchers to easily explore the potential of cloud computing. This is particularly important for early career researchers.

Pawsey
QCIF
eRSA
NCI
Intersect
University of Melbourne
Monash University
University of Tasmania

Users are supported by a comprehensive help desk portal, including telephone, email, and up-to-date documentation.

support.ehelp.edu.au

ardc
australian research data commons
Built from ands · nectar · RDS www.rds.edu.au

nectar

Policy

Key stakeholders

Australian Government
Department of the Prime Minster and Cabinet

Australian Government
Department of Education and Training

Australian Government
Australian Research Council

Australian Government
National Heath and Medical Research Council

100+ submissions to Government:

- Productivity Commission - Intellectual Property Arrangements
- Review on Research Policy and Funding Arrangements
- Defence Trade Control Act 2012
- Data Availability and Use Productivity Commission - Draft Report
- ARC's Research Engagement and Impact Assessment Consultation Paper
- Medical Research Future Fund (MRFF)
- Australian Government **Data Sharing and Release Act**

Guides and consultancies

MOU
with
NHMRC
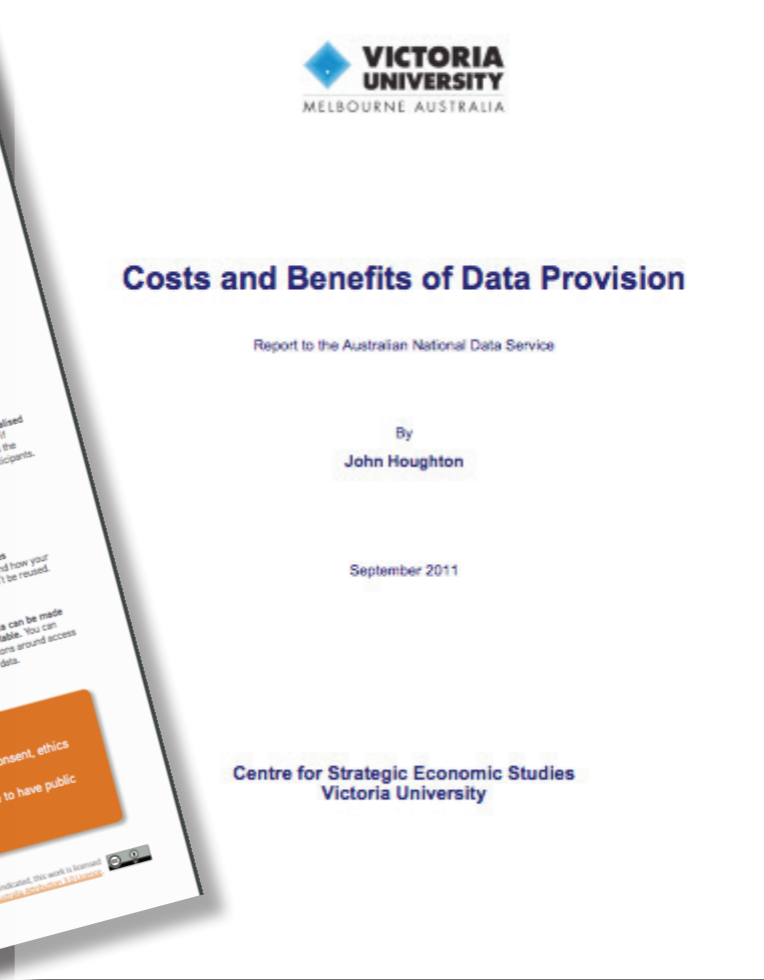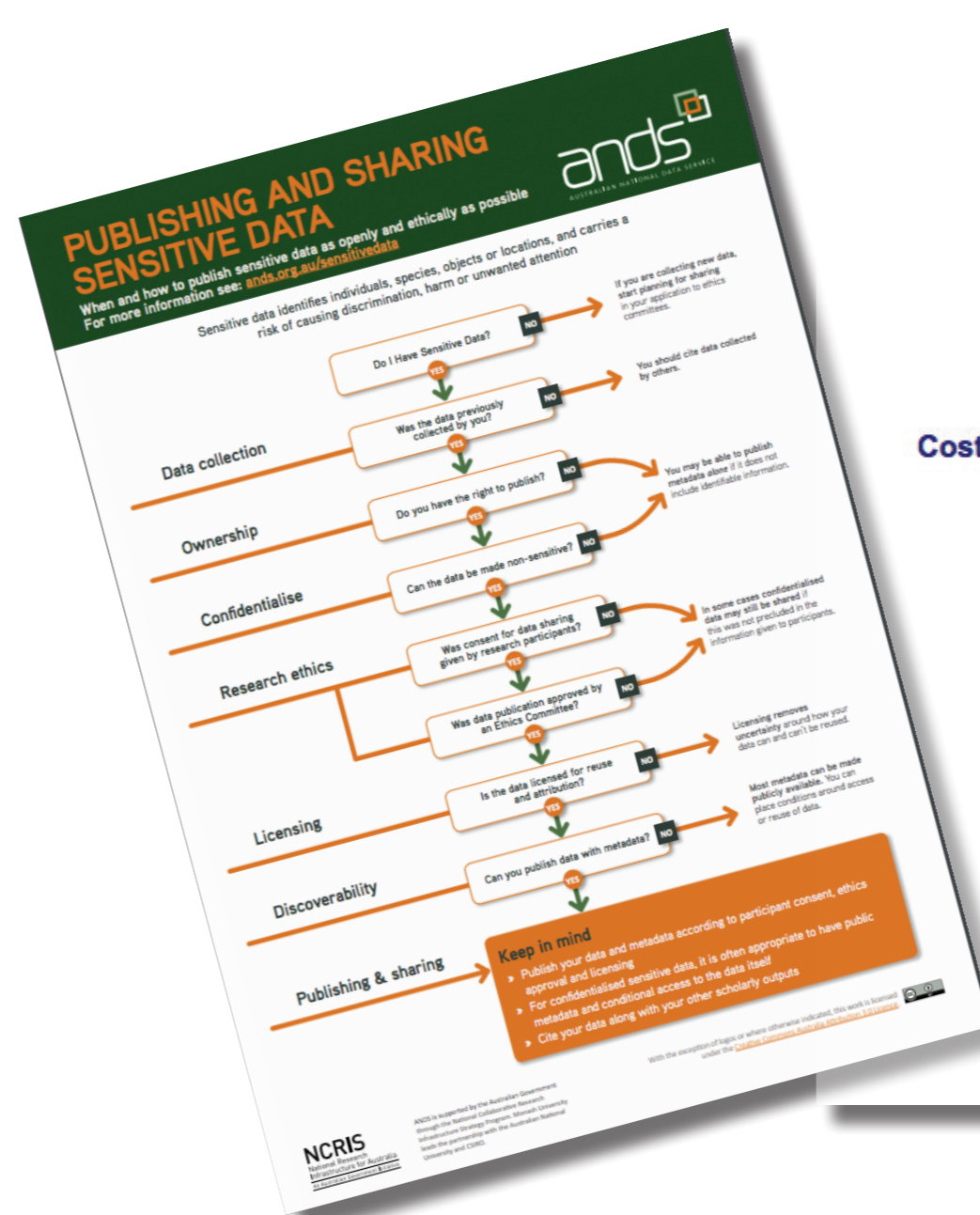
Invited input to national policy

Including:

- New (2018) Code for the Responsible Conduct of Research (ARC, NHMRC and Universities Australia)
- Management of Data and Information in Research - a data sharing guide supporting the New (2018) Code
- Australian Government Public Data Policy Statement (Prime Minister and Cabinet)
- Australian Government **Data Sharing and Release Act** (in preparation)

ardc
australian research data commons
Built from ands • nectar • RDS

ands
AUSTRALIAN NATIONAL DATA SERVICE

International

Aims

- Position Australia as a leader in international research data and methods.
- Partner with mature international infrastructure initiatives to provide better infrastructure and services to the Australian research community.
- Influence international policy discussions around the best use of e-Infrastructure to deliver research outcomes.
- Influence developing international infrastructure initiatives to align with technologies and approaches compatible with those adopted in Australia, and vice-versa.
- Learn from alternative national, regional, and domain specific infrastructure implementation approaches to inform the Australian approach and encourage sharing of best practice.

Staff are involved in

- Research Data Alliance - Council, Technical Advisory Board, Secretariat, Working Group and Interest Group chairs.
- Organisation for Economic Co-operation and Development (OECD) - Co-chairs of High Level Expert Groups on International Coordination of Data Infrastructures, Data Skills.
- Research software infrastructure international initiatives such as the International Coalition on Science Gateways, Science Gateways Community Institute (USA), Horizon 2020 (Europe), Software Sustainability Institute (UK) and Centre for Open Science.
- International infrastructure activities such as Openstack development.

With the USA and EU we initiated the Research Data Alliance

Outcomes

- Recognised international leadership in eResearch infrastructure
- Adoption of the Research Data Alliance outputs in Australia (23 Research Data Things, Dynamic Data Citation)
- Australian input into the OECD policy documents
- Australian leadership of the Research Data Alliance Interest and Working Groups

ardc
australian research data commons
Built from ands • nectar • RDS www.rds.edu.au

ands
AUSTRALIAN NATIONAL DATA SERVICE
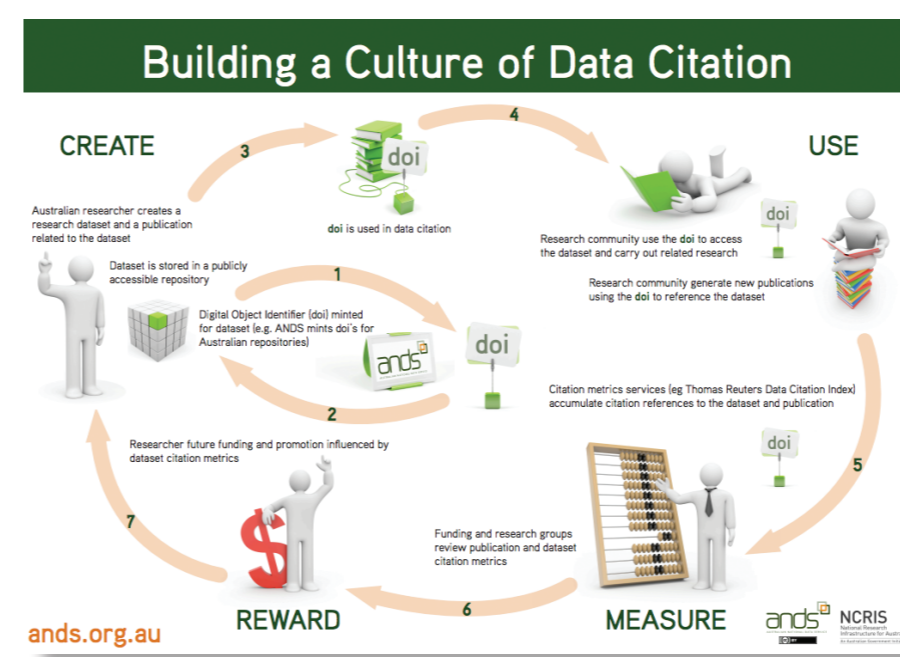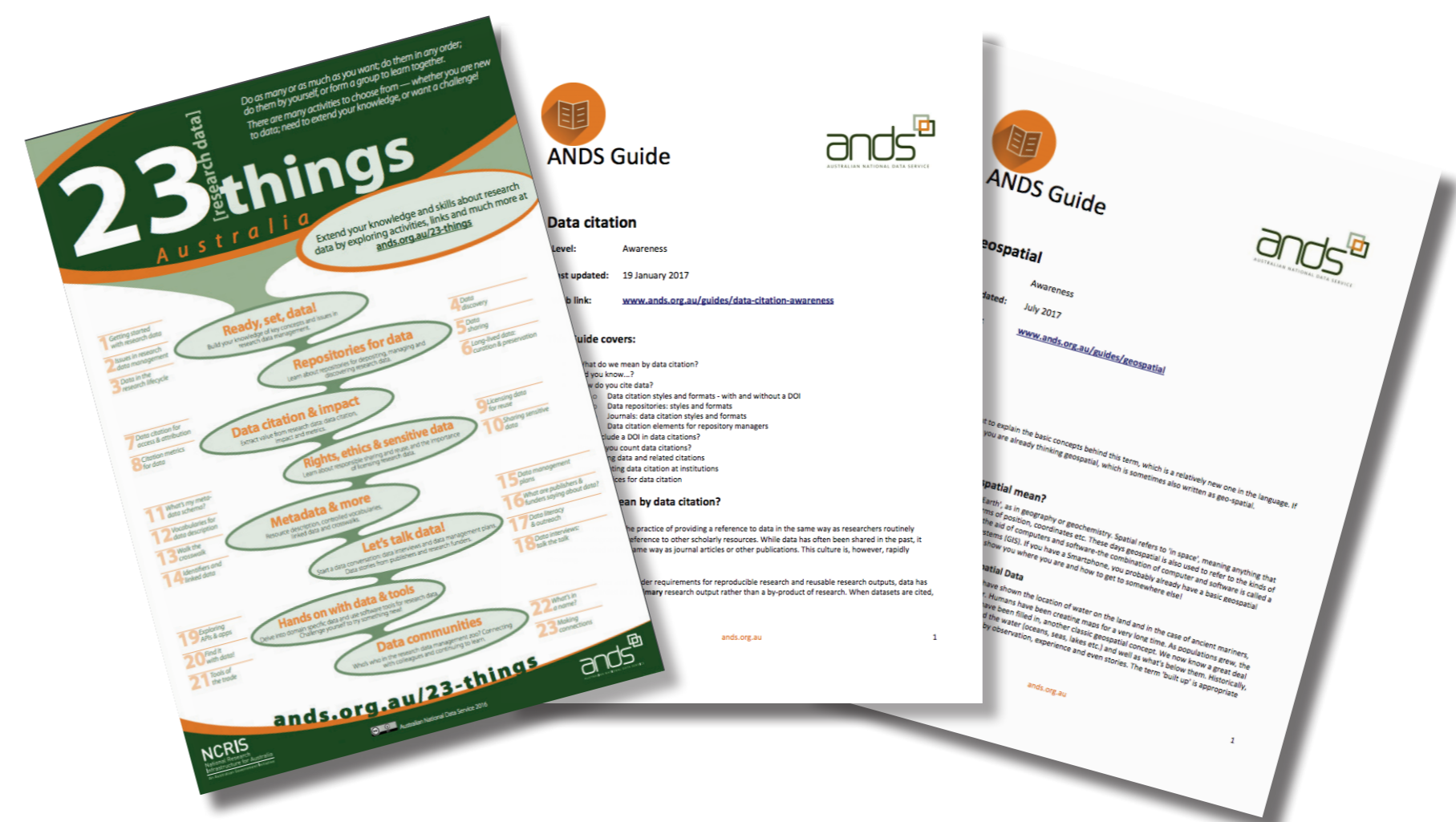
Skills

What we do

Deliver programs that:

- Develop a skilled workforce supporting data intensive research communities
  - » Research user communities (current and potential users of eResearch infrastructure)
  - » Communities that support and engage with researchers (librarians, eResearch support staff, software engineers, research office staff, data skills trainers)
- Build awareness of, and ability to exploit, digital data assets, methods and tools across the research sector
- Coordinate the skills and community building agenda within and across institutions, domains and communities

23 (Research Data) Things world renowned training

Guides, Reports, Resources, Website, Training

- 30+ guides and reports - thousands of downloads each year
- Rich web content - over 5,000 page visits per month
- Development of concept information and resources
- Intensive national skills program, delivered via:
  - » Webinars
  - » Face to face training
  - » Online resources

Outcomes

- Increased skills, knowledge and confidence around data management concepts, tools and practices across the Australian research sector
- Established active Communities of Practice amongst data librarians and repository managers
- Initiated Special Interest Groups to share best practice around data management plans, persistent identifiers, institutional infrastructure and more
- Established collaborative relationships with key stakeholders such as the Research Office, Office of Research Integrity and Ethics, Medical Research Institutes and the Australian Library and Information Association to deliver targeted training and resources
- Delivered a wealth of general and domain specific skills related resources and events aligned to NCRIS and national priority areas

ardc
australian research data commons
Built from **ands** • **nectar** • **RDS** www.rds.edu.au

ands
AUSTRALIAN NATIONAL DATA SERVICE

Data Discovery

Mint and support persistent identifiers

Enabling connected research

- DOI - Digital Object Identifier
    - » Identifies research outputs
- ISGN - International Geo Sample number
    - » Identifies physical samples collected during the course of research
- RAiD - Research Activity Identifier
    - » Identifies research project activities
- ORCID - Open Researcher and Contributor Identifier
    - » Identifies who is doing research

Explore research grants and projects

- Finds information on Australian grants and projects by subject, funder, funding scheme or managing institution
- Brings together research grant information from participating funders and project descriptions from institutions for discovery
- Provides links to related data descriptions in Research Data Australia.

Research Vocabularies Australia

Enabling the use of common language for concepts in datasets

- One stop shop for finding controlled vocabularies used in research
- Available at no cost for anyone to search and browse
- Australian research organisations can register for free and create, manage and publish their own vocabularies
- Makes it easy to reuse vocabularies across research communities.

Research Data Australia

Showcasing Australia's data to the world

- One stop shop: an online portal for finding research data
- Almost 100 contributors
- Available at no cost for anyone around the world to search for and browse research data
- Does not store the data itself, but aggregates data descriptions from all disciplines to enable searching
- Within Research Data Australia, you can also find information on data-related software and service.

139,000+ datasets in Research Data Australia

ardc
australian research data commons
Built from ands • nectar • RDS
www.rds.edu.au

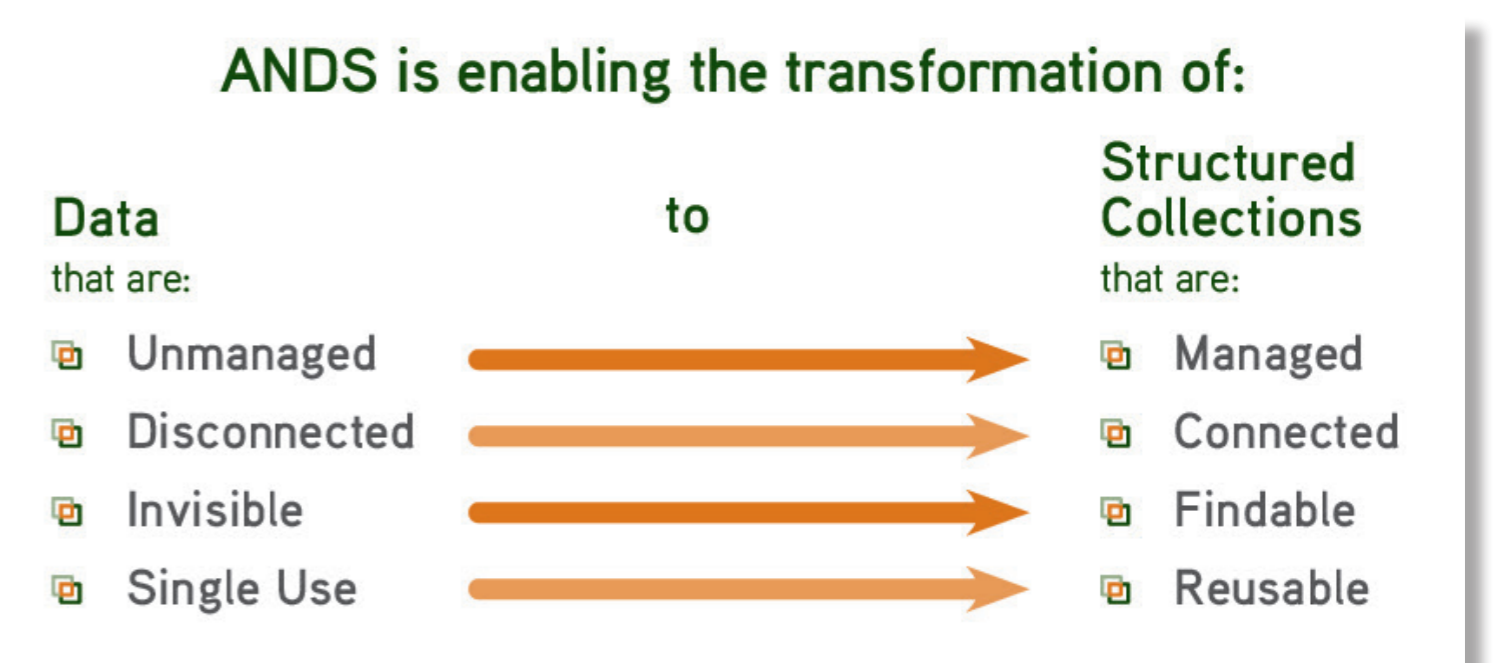**ands**
AUSTRALIAN NATIONAL DATA SERVICE

Data Management

Key stakeholders

- Universities and publicly funded research organisations
- Government departments and agencies
- Medical Research Institutes and Co-operative Research Centres
- NCRIS facilities

What we do

- Engage with stakeholders to build an institutional data management culture and framework
  - » Fund projects to build data management infrastructure, tools and capability
  - » Foster internal collaboration between key stakeholders within an institution such as researchers, library, research office, IT Department, DVCR
  - » Support training and skills development programs
- Encourage collaboration and communication across institutions and domains
  - » Co-ordinate tech talks and roundtable events
- Communicate and support uptake of FAIR principles amongst stakeholder groups

ANDS is enabling the transformation of:

Data that are:	to	Structured Collections that are:
Unmanaged	→	Managed
Disconnected	→	Connected
Invisible	→	Findable
Single Use	→	Reusable

Programs and projects

- The Open Data and High Value Collection projects enabled institutions to publicly showcase some 60 high profile datasets
- The Trusted Repositories and Outputs projects delivered exemplars of best practice in the management of repositories and research outputs
- The Metadata Stores program established a national network of institutional data registries that made data discoverable through Research Data Australia and beyond
- The Seeding the Commons program enabled institutions to develop data management policies, plans and procedures
- The National Collections program delivered enhanced access to significant data holdings in NCRIS facilities, national agencies and government departments

Findable

Accessible

Interoperable

Reusable

Outcomes

- Strong collaborative relationships established with key stakeholder groups
- Universities and other stakeholders received funding and practical support to:
  - » establish foundational data management infrastructure
  - » develop data policies, data management plans and data services (eg DOI)
- More Australian research data is now well managed, discoverable and reusable
- Data management plans, data citation and FAIR are becoming routine research practice
- ANDS has a reputation as being a world leader in research data management

Leading research data management in Australia

ardc
australian research data commons

Built from ands · nectar · RDS
www.rds.edu.au

ands
AUSTRALIAN NATIONAL DATA SERVICE