Australian Research Data Commons

ARDC Digital Research Skills Summit 2024

Day 2: Building Community Around Research Infrastructure

22 May 2024

Australian Government

NCRIS National Research Infrastructure for Australia

ARDC is enabled by NCRIS

Setting the Scene

Kathryn Unsworth- Manager, Skilled Workforce Development

Kathryn Unsworth

Manager, Skilled Workforce Development

Australian Research Data Commons

Currently leading the Skills & Workforce Development team. Kathryn drives a nationally coordinated approach to skills and training programs targeting key areas of the workforce that conduct, underpin, and enable data-intensive research.

Australian Research Data Commons
NCRIS National Research Infrastructure for Australia An Australian Government Initiative
ARDC is enabled by NCRIS

Logistics

- Welcome pack
  - Code of Conduct for ARDC activities
  - Photography/recording
- Nearest exits
- Restrooms
- Catering
- Speaker gifts - Fifteen Trees
- Meet our skills concierges
- Overview of today's program

Use the QR code to access your Welcome Pack

Make this space inclusive, respectful and safe. If you feel unsafe, tell us!

Australian Research Data Commons

NCRIS National Research Infrastructure for Australia
An Australian Government Initiative

ARDC is enabled by NCRIS

FIFTEEN
TREE5
www.15trees.com.au

ARDC Skills Concierges

Adeline Wong

Skills Development Lead
(Learning Design)

Amany Gouda-Vossos

Skills Development Lead
(People)

Ellen Lyrtzis

Skills Development Lead
(NCRIS)

Kathryn Unsworth

Manager,
Skilled Workforce
Development

Kit Greenhill

Skills Development Lead
(HASS and Indigenous)

Liz Stokes

Skills Development Lead
(Trainer and Research
Communities)

Meirian Lovelace-Tozer

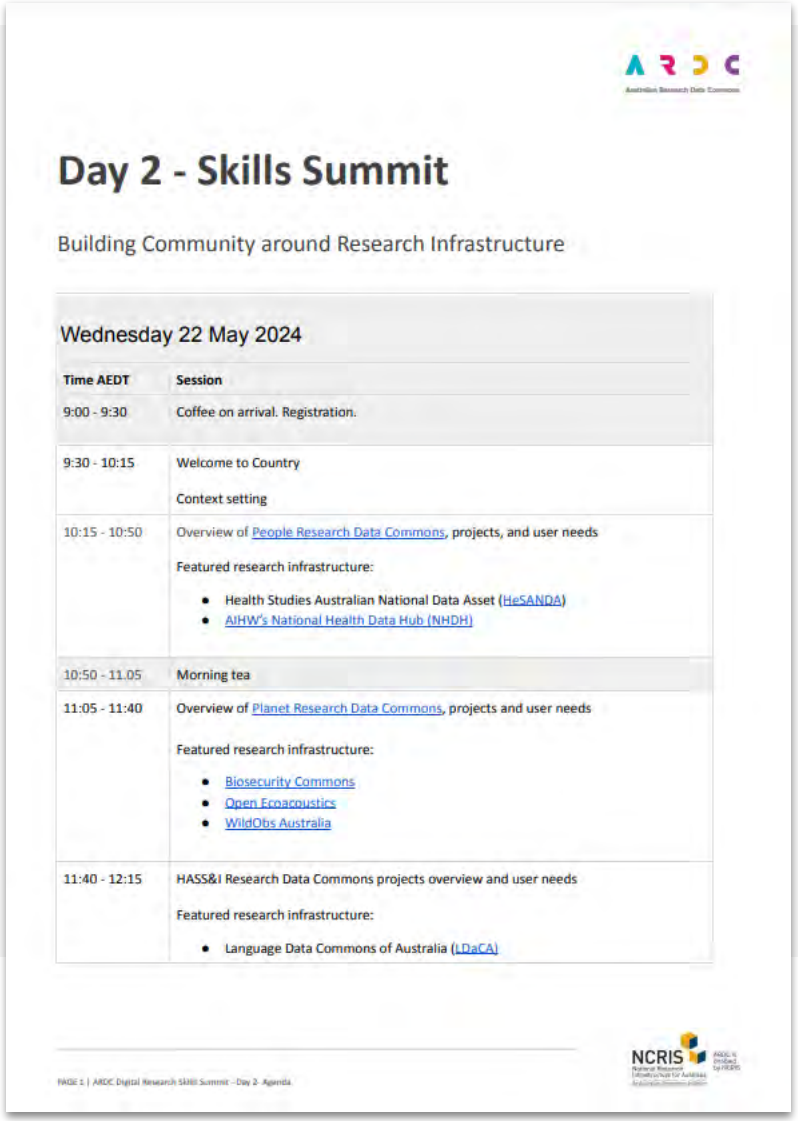
Skills Development Lead
(Services)

Rob Clemens

Skills Development Lead
(Planet)

Australian Research Data Commons
NCRIS National Research Infrastructure for Australia An Australian Government Initiative
ARDC is enabled by NCRIS

ARDC Digital Research Skills Summit - Program overview

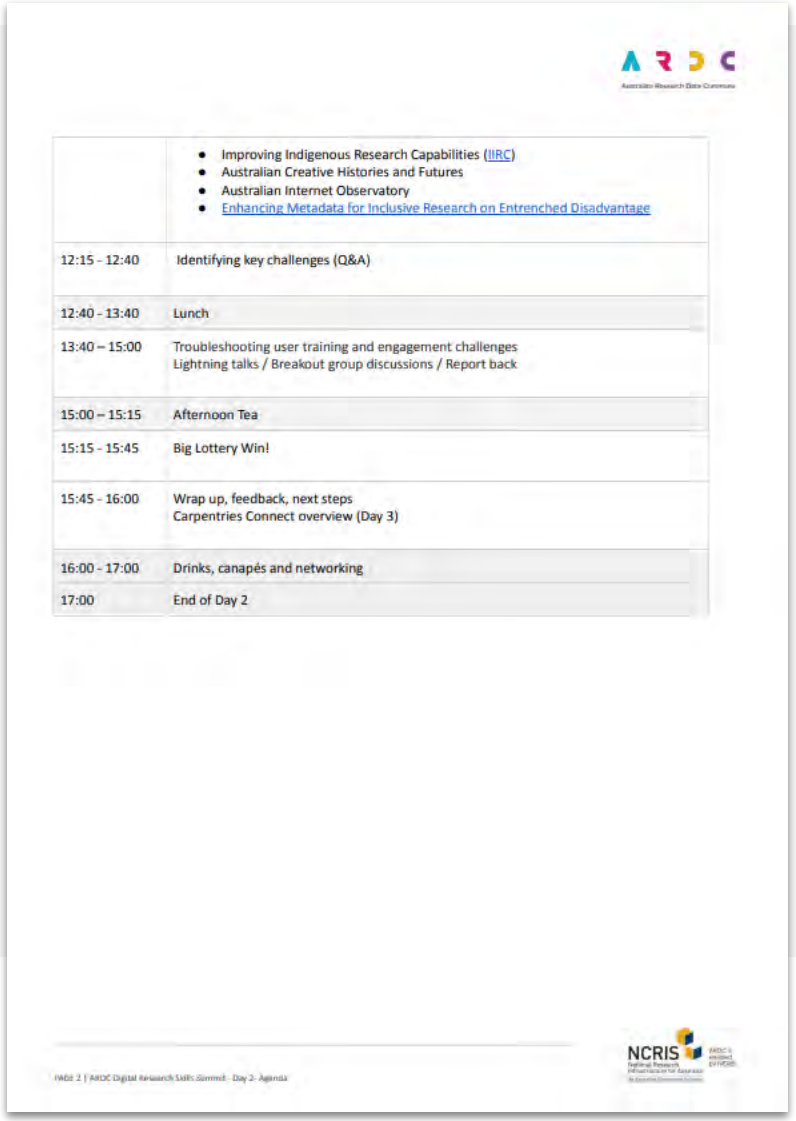


Day 2 - Skills Summit

Building Community around Research Infrastructure

Wednesday 22 May 2024

Time AEDT	Session
9:00 - 9:30	Coffee on arrival. Registration.
9:30 - 10:15	Welcome to Country Context setting
10:15 - 10:50	Overview of People Research Data Commons, projects, and user needs Featured research infrastructure: • Health Studies Australian National Data Asset (HeSANDA) • AIHW's National Health Data Hub (NHDH)
10:50 - 11.05	Morning tea
11:05 - 11:40	Overview of Planet Research Data Commons, projects and user needs Featured research infrastructure: • Biosecurity Commons • Open Ecoacoustics • WildObs Australia
11:40 - 12:15	HASS&I Research Data Commons projects overview and user needs Featured research infrastructure: • Language Data Commons of Australia (LDaCA)



	• Improving Indigenous Research Capabilities (IIRC) • Australian Creative Histories and Futures • Australian Internet Observatory • Enhancing Metadata for Inclusive Research on Entrenched Disadvantage
12:15 - 12:40	Identifying key challenges (Q&A)
12:40 - 13:40	Lunch
13:40 – 15:00	Troubleshooting user training and engagement challenges Lightning talks / Breakout group discussions / Report back
15:00 – 15:15	Afternoon Tea
15:15 - 15:45	Big Lottery Win!
15:45 - 16:00	Wrap up, feedback, next steps Carpentries Connect overview (Day 3)
16:00 - 17:00	Drinks, canapés and networking
17:00	End of Day 2

Australian Research Data Commons

NCRIS National Research Infrastructure for Australia An Australian Government Initiative

ARDC is enabled by NCRIS

Welcome to Country

Wurundjeri representative

Setting the Scene

Keith Russell - Director, Outreach, ARDC

Keith Russell

Director, Outreach

Australian Research Data Commons

With an international background in research data management and sharing, Keith works with the Skills, Engagements and Communications teams to ensure the ARDC reaches our stakeholders, partners and communities to grow the capability across the sector.

Australian Research Data Commons
NCRIS National Research Infrastructure for Australia An Australian Government Initiative
ARDC is enabled by NCRIS

Objective of the 2024 ARDC Digital Research Skills Summit

To discover how digital infrastructure providers and research communities are upskilling researchers in emerging research technologies.

Australian Research Data Commons
NCRIS National Research Infrastructure for Australia An Australian Government Initiative
ARDC is enabled by NCRIS

ARDC Leadership forum: Panel discussion recap

Australian Research Data Commons
NCRIS National Research Infrastructure for Australia An Australian Government Initiative
ARDC is enabled by NCRIS

Australian Research Data Commons

About the ARDC

The ARDC runs services and partnership programs that ensure Australian researchers are internationally competitive through high-quality data assets, platforms, infrastructure, policies, people and skills.

Australian Research Data Commons

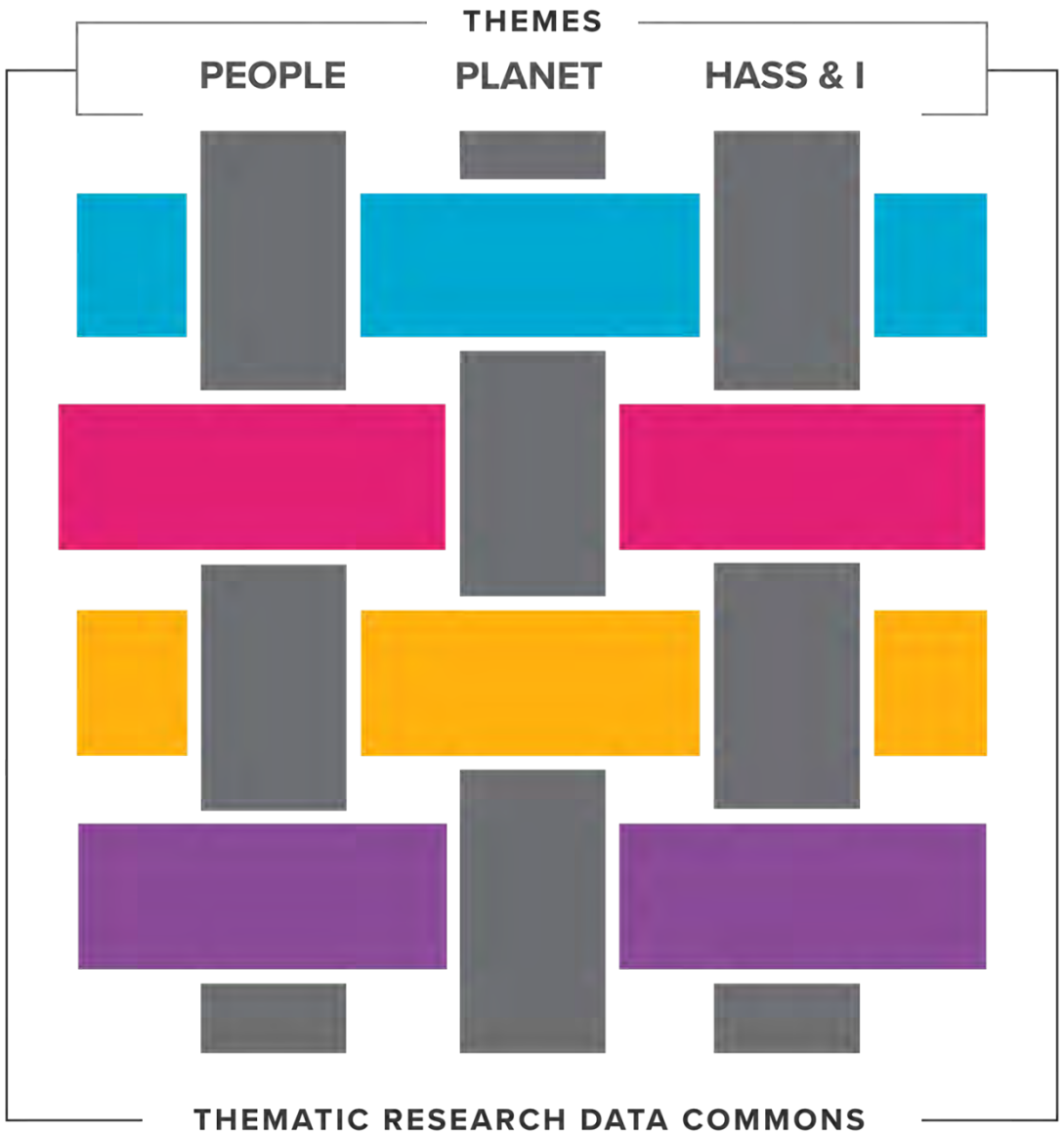
ARDC Strategy

Our Purpose

To provide Australian researchers with competitive advantage through data.

Our Mission

To accelerate research and innovation by driving excellence in the creation, analysis and retention of high-quality data assets.



Australian Research Data Commons

NCRIS National Research Infrastructure for Australia An Australian Government Initiative

ARDC is enabled by NCRIS

Australian Research Data Commons

What is a Research Data Commons?

A research data commons brings together people, skills, data, and related resources such as storage, compute, software, and models to enable researchers to conduct world class data-intensive research.

Australian Research Data Commons

NCRIS National Research Infrastructure for Australia
An Australian Government Initiative
ARDC is enabled by NCRIS

PEOPLE RESEARCH DATA COMMONS

A national-scale data infrastructure for health research and research translation

Image — Gorodenkoff - 236237514 /

Australian Research Data Commons
NCRIS National Research Infrastructure for Australia An Australian Government Initiative
ARDC is enabled by NCRIS

NATIONAL DATA CHALLENGES FOR HEALTH RESEARCH

What **data assets** do Australian health researchers need & how do they found out more about them?

What **technology** can researchers use to securely access & work on health & medical data?

How can health researchers **efficiently bring together data** from different sources?

What **cutting edge analytical tools** (like machine learning) can researchers use to work on their data?

Australian Research Data Commons
NCRIS National Research Infrastructure for Australia An Australian Government Initiative
ARDC is enabled by NCRIS

Australian Government
Australian Institute of Health and Welfare

National Health Data Hub (NHDH)

May 2024

AIHW

Stronger evidence, better decisions, improved health and welfare

The National Health Data Hub

NHDH is a centralised deidentified linkage system that aims to deliver:

- Better coverage and expansion of data
- Access to government and non-government researchers
- Reduced duplication of effort and costs
- Improved time frames for access to data
- More efficient and safer data sharing
- Improved timeliness of data
- Streamlined governance arrangements, including acceptance of programs of work
- Flexible and responsive to emerging issues
- Greater interoperability between State and Territory linkage nodes and existing linked data assets such as PLIDA and NDDA
- Enabling two-way data flows, and other data sharing initiatives (e.g. NMLK)
- Scalability and gradual expansion of content.

Principles of NHDH

The guiding principles for the development of the NHDH include:

- **Privacy by design** and embedded at the integration system, asset, and project level.
- **Minimise the need to share and use personal information**. Adoption of a distributed interoperable linkage model with state and territory data linkage units linking to a common spine (based on the Medicare Consumer Directory) and sharing through de-identified maps.
- **Share once and use many times**. Reuse existing pipelines of data supply with new authorisations.
- **Employ efficient methodologies and processes** that are fit-for-purpose in integration activities.

Data sources included in NHDH

National Aged Care Data Clearinghouse

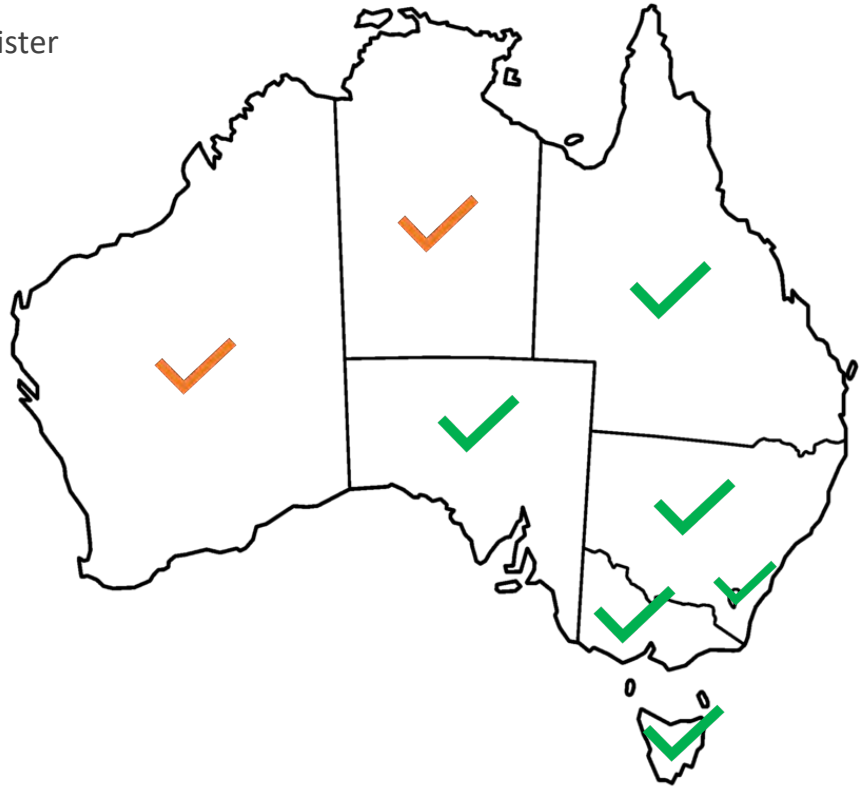
Australian Immunisation Register

Pharmaceutical Benefits Scheme

Hospital data

Medicare Benefits Schedule

National Death Index



Progressing towards the inclusion of:

- National disability data
- Cancer data
- COVID-19 cases from the National Notifiable Disease Surveillance System
- Australian and New Zealand Intensive Care Society data
- Mental health data
- Perinatal data

What the NHDH can and can't be used for

NHDH **can** be used for:

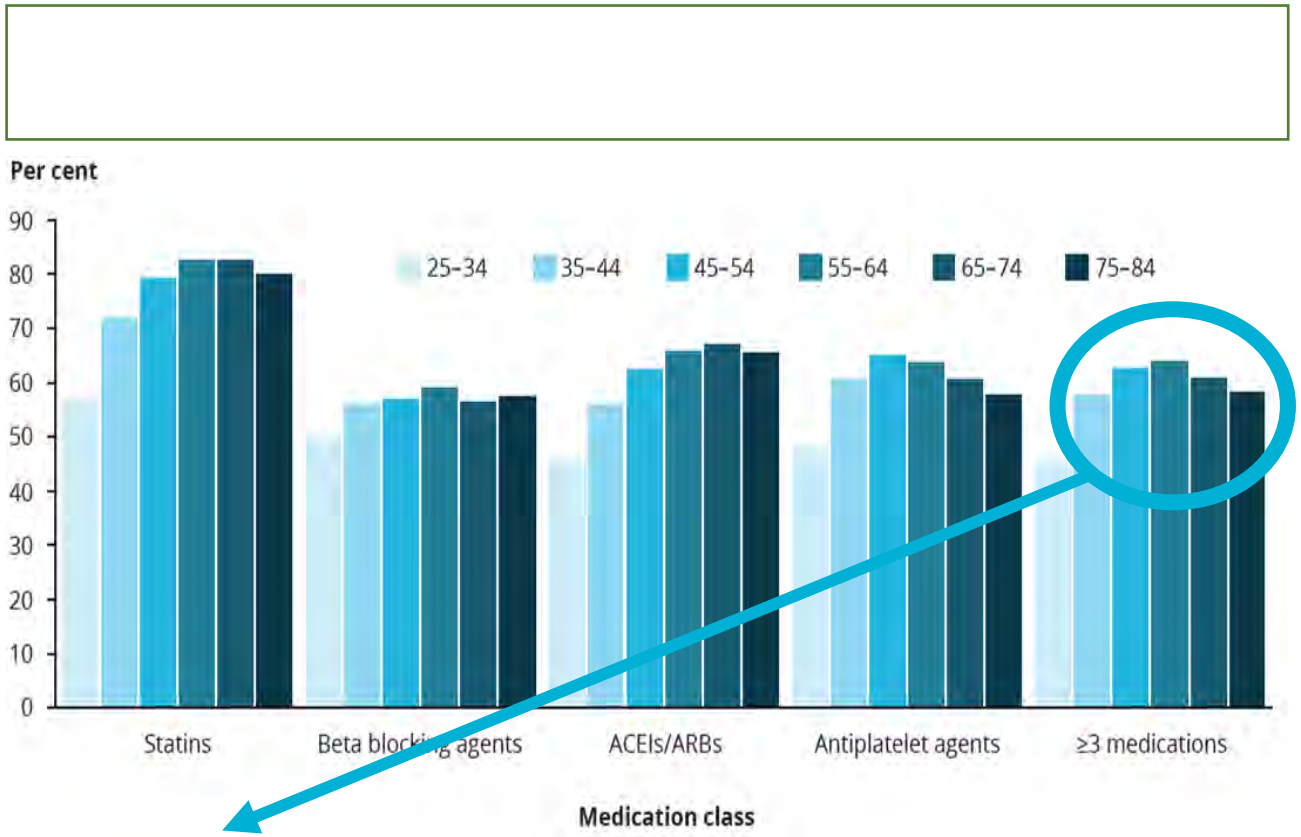
- Health research and statistical analyses that supports health service planning, monitoring and evaluation and health policy development
- Monitoring variations and patterns of population health outcomes to inform clinical practice review and service delivery for the purposes of ensuring safety and quality of care
- Performance and health outcomes reporting at a national level
- Design and development of performance and productivity measures

NHDH **can't** be used for:

- To identify and report on any individual
- To identify and report on any service provider or clinical practice
- To identify and report on individual diagnosis of a medical condition(s)
- Performance monitoring and reporting below the national level
- Author data insights, findings, and reports at an individual level
- Administrative and/or compliance reporting purposes

What general insights have been obtained from NHDH?

Are clinical guidelines for medications for CHD being followed?



Key Finding: Only 61% people with ACS were dispensed 3 or more of the recommended cardiovascular medicines within 40 days of leaving hospital

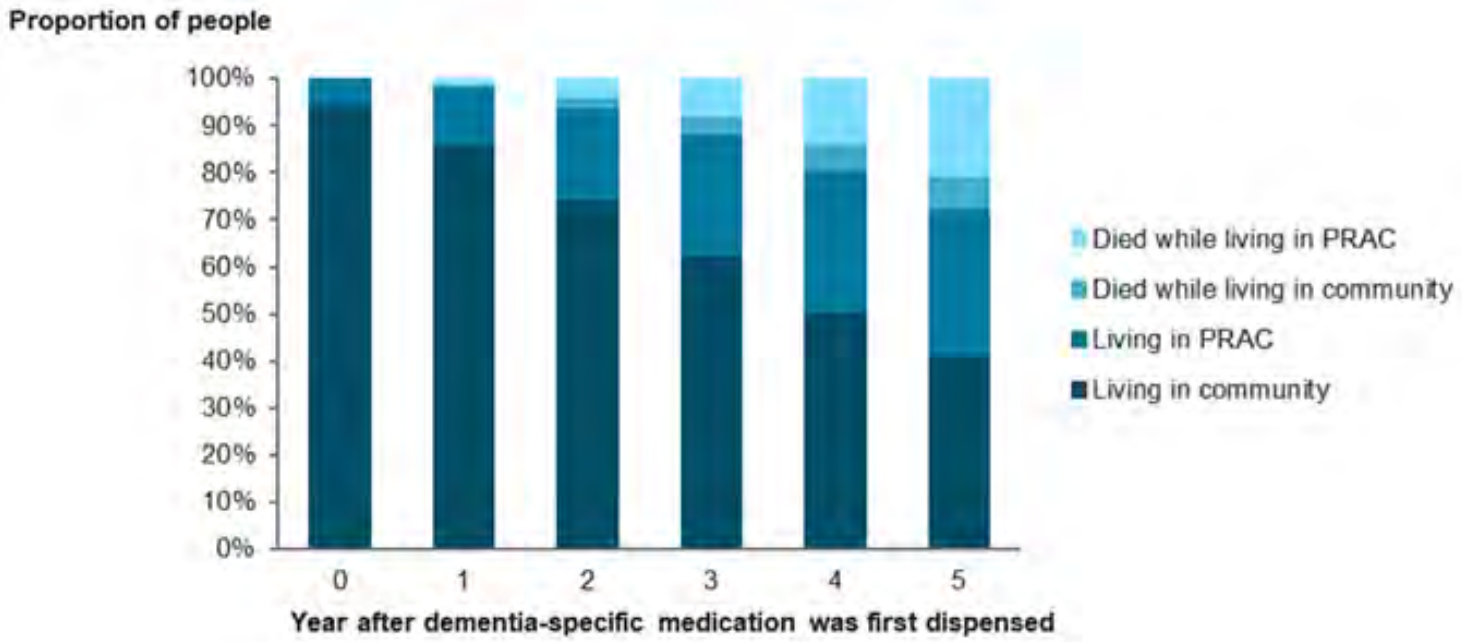
Other NHDH project examples:

- Health service use in the last year of life – overall population and people who died by suicide
- Health Outcomes Modelling and Evaluation (HOME) model
- Advanced understanding of health and aged care service use for people with dementia
- Hospital use and deaths for people hospitalised for family and domestic violence

Analysing Outcomes for populations of Interest using NHDH

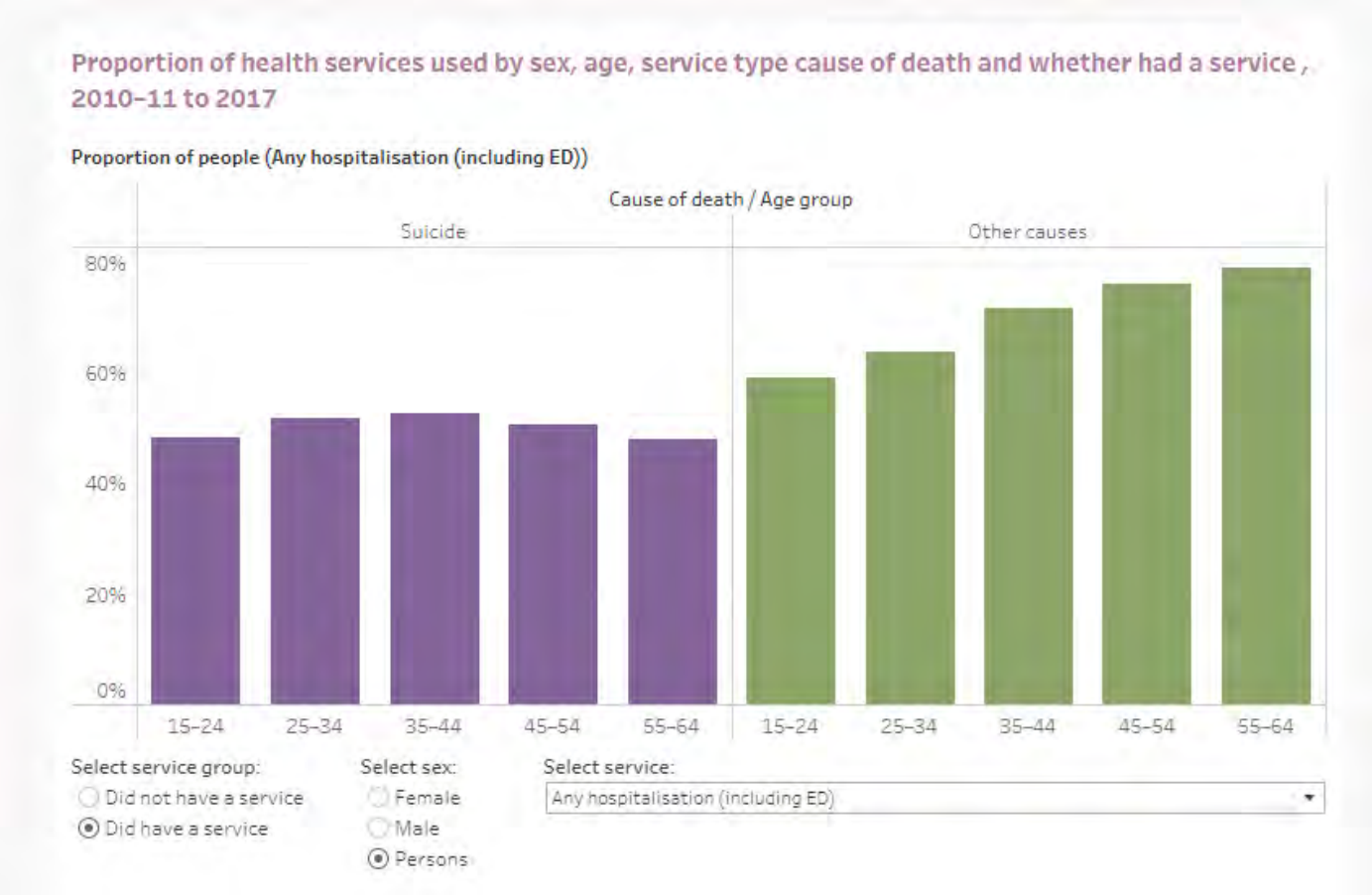
Overview of outcomes for people with younger onset dementia

Figure 1: Outcomes for people with younger onset dementia (ages 30–69), by year since dementia-specific medication was first dispensed in 2011–2012



Further insights from NHDH

People who died by suicide accessed fewer health services in their last year of life than those who died from other causes



Secure Environment for Analysing Data

NHDH is available via an AIHW managed instance of the Secure Environment for Analysing Data (SEAD)

Standard software in SEAD includes:

- R
- Python
- LibreOffice

Optional software includes:

- SAS
- Stata
- Databricks*

*Databricks may be required for large-scale data analysis projects

Users of NHDH

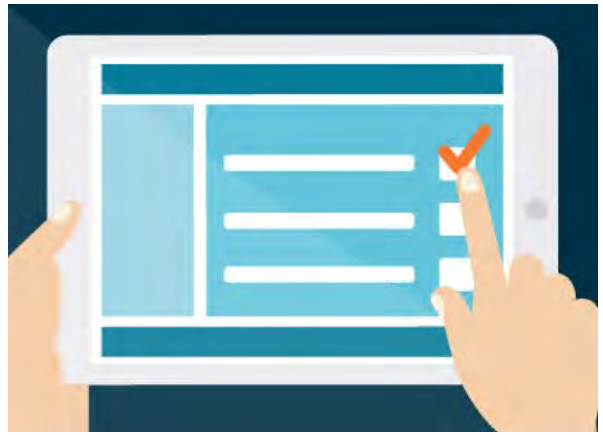
- **Analysts, Researchers and Academics**
  - Expertise in medical research, biostatistics, and epidemiology
- **Healthcare Policymakers**
  - Expertise in health policy and public administration
- **Public Health Officials**
  - Expertise in public health and preventative medicine

Skills for proficient use of the NHDH

Users of the NHDH require a sound knowledge and skillset for the following:

- **Data literacy**
  - Understand complex datasets, statistical methods, and data visualisation
- **Data privacy and security**
  - Ability to maintain security and privacy of sensitive data
- **Research and Analytic skills**
  - Skills in formulating research designs and employing appropriate research methodologies
- **Policy interpretation**
  - Ability to interpret policies

Gathering user input and feedback for NHDH



Previous Practices:

- The LINDAHR Project

Current Practices:

- Track utilisation and outcomes of NHDH projects
- Analyse trends and identify areas for improvement
- Feedback forums to discuss limitations, strengths and direction

Planned Practices via User Feedback Surveys:

- Gather insights on strengths, areas for improvement and support needs annually
- Inform Enhancements to data access and usability

QUESTIONS

HeSANDA and Health Data Australia

22 May 2024

Mr Nemanja Zivanov - Associate Research Fellow and Project Manager for the Mental Health Node

Ms Katie Ozdowska - UoM MISCH Clinical Trials Node Manager and Project Manager for MACH

Mr Mat Ishac - Senior Project Officer at Monash University and Project Manager for Monash Partners

MACH Melbourne Academic Centre for Health

Mental Health Node

MONASH University

ARDC Australian Research Data Commons

NCRIS National Research Infrastructure for Australia An Australian Government Initiative

ARDC is enabled by NCRIS

What is Health Data Australia?

Platform, services, and tools.

MACH Melbourne Academic Centre for Health

Mental Health Node

MONASH University

ARDC

NCRIS National Research Infrastructure for Australia

Australian Research Data Commons

Health Data Australia

Launched in July 2023

Discover

Request Access

Metadata

Partner Organisations

MACH Melbourne Academic Centre for Health

Mental Health Node

MONASH University

ARDC

NCRIS

PEOPLE RESEARCH DATA COMMONS

A national-scale data infrastructure for health research and research translation

Image – Chinnapong 325872880 / bigstockphoto.com

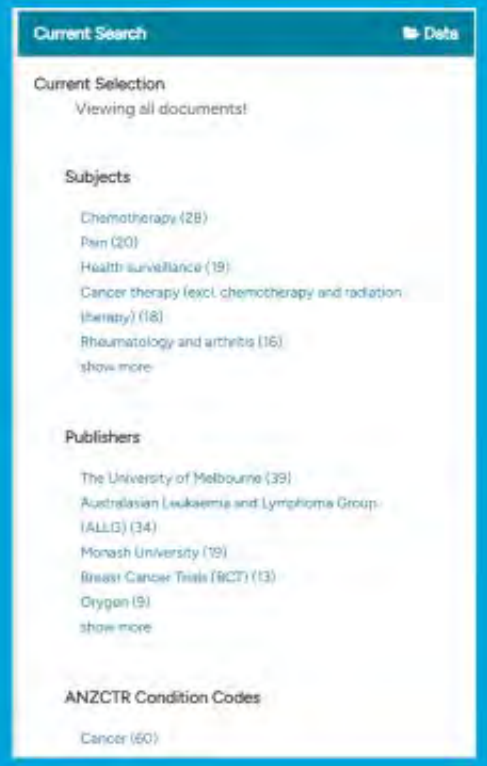
MACH Melbourne Academic Centre for Health

Mental Health Node

MONASH University

ARDC

NCRIS



User-friendly Interface

ARDC
Australian Research Data Commons

Health Data Australia

Extensive Metadata Repository with Continuous Updates

MACH Melbourne Academic Centre for Health

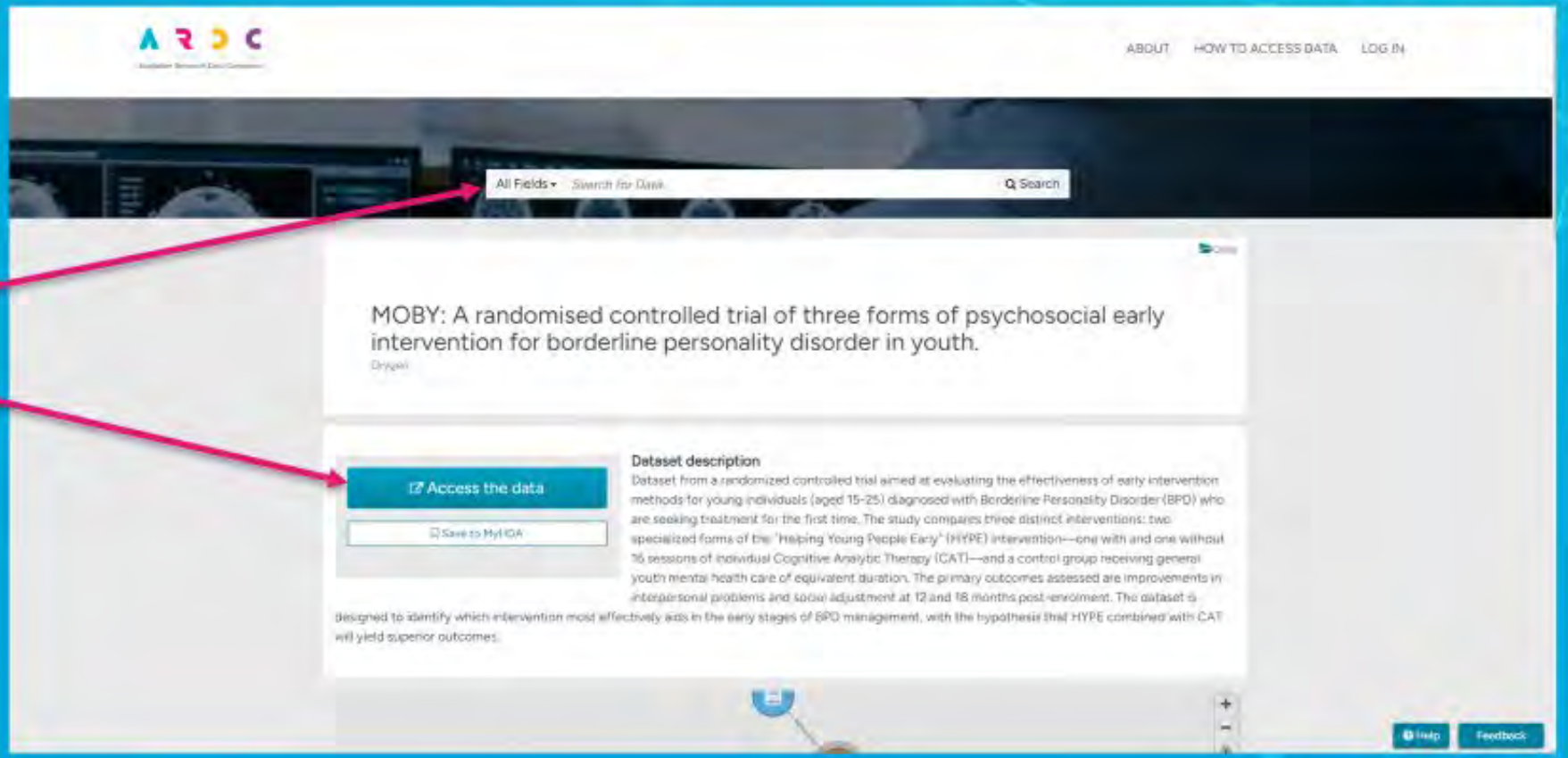
Mental Health Node

MONASH University

ARDC Australian Research Data Commons

NCRIS National Research Infrastructure for Australia

- Data Discovery
- Data Access Request
- Data Integration



MACH Melbourne Academic Centre for Health

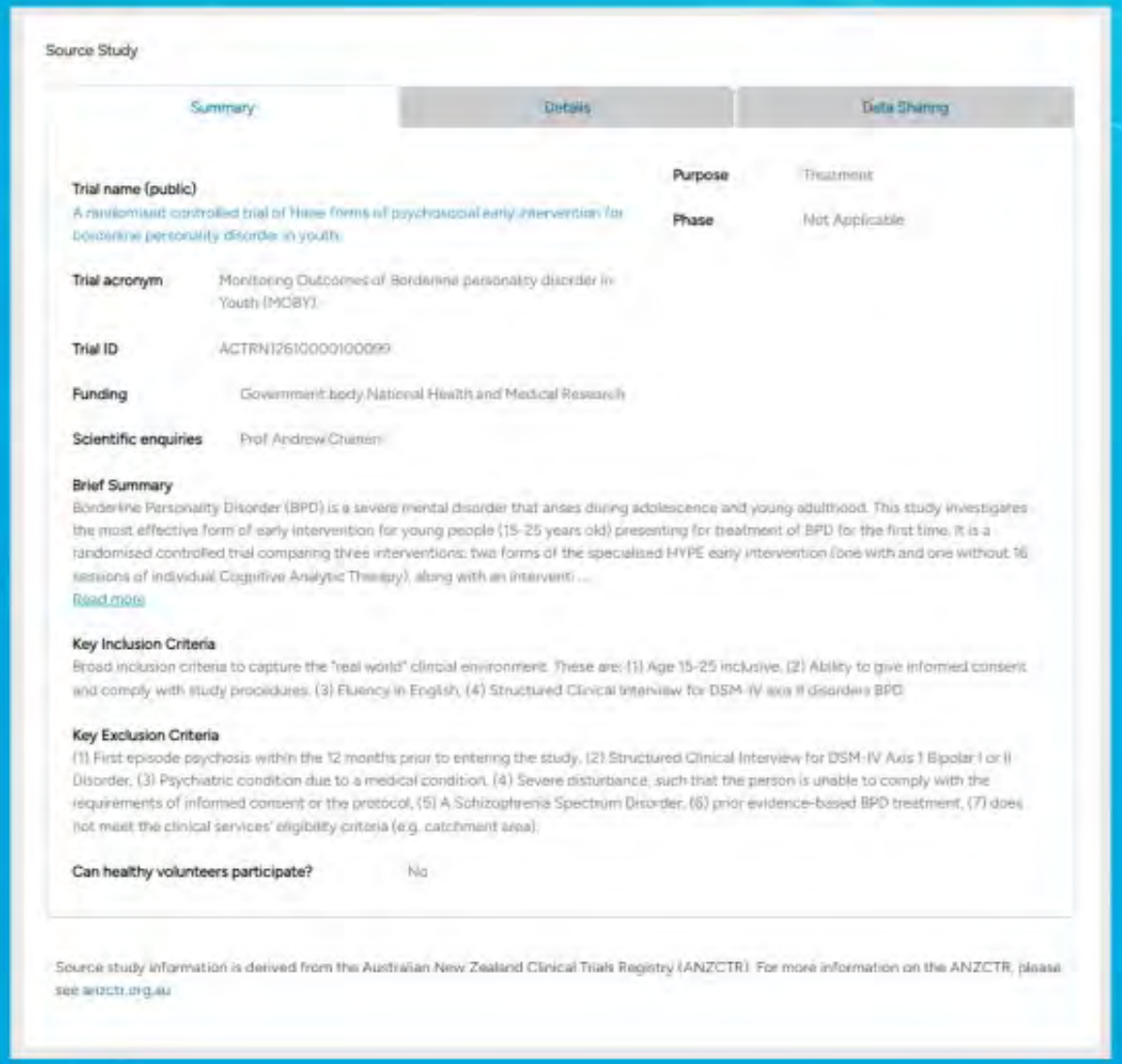
Mental Health Node

MONASH University

ARDC

NCRIS

ANZCTR API



Source Study

Summary | Details | Data Sharing

Trial name (public)
A randomised controlled trial of three forms of psychosocial early intervention for borderline personality disorder in youth

Trial acronym Monitoring Outcomes of Borderline personality disorder in Youth (MOBY)

Trial ID ACTRN12610000100099

Funding Government body National Health and Medical Research

Scientific enquiries Prof Andrew Chanen

Purpose Treatment

Phase Not Applicable

Brief Summary
Borderline Personality Disorder (BPD) is a severe mental disorder that arises during adolescence and young adulthood. This study investigates the most effective form of early intervention for young people (15-25 years old) presenting for treatment of BPD for the first time. It is a randomised controlled trial comparing three interventions: two forms of the specialised HYPE early intervention (one with and one without 16 sessions of individual Cognitive Analytic Therapy), along with an intervent ...
Read more

Key Inclusion Criteria
Broad inclusion criteria to capture the "real world" clinical environment. These are: (1) Age 15-25 inclusive, (2) Ability to give informed consent and comply with study procedures, (3) Fluency in English, (4) Structured Clinical Interview for DSM-IV axis II disorders BPD

Key Exclusion Criteria
(1) First episode psychosis within the 12 months prior to entering the study, (2) Structured Clinical Interview for DSM-IV Axis 1 Bipolar I or II Disorder, (3) Psychiatric condition due to a medical condition, (4) Severe disturbance, such that the person is unable to comply with the requirements of informed consent or the protocol, (5) A Schizophrenia Spectrum Disorder, (6) prior evidence-based BPD treatment, (7) does not meet the clinical services' eligibility criteria (e.g. catchment area).

Can healthy volunteers participate? No

Source study information is derived from the Australian New Zealand Clinical Trials Registry (ANZCTR). For more information on the ANZCTR, please see anzctr.org.au

MACH Melbourne Academic Centre for Health

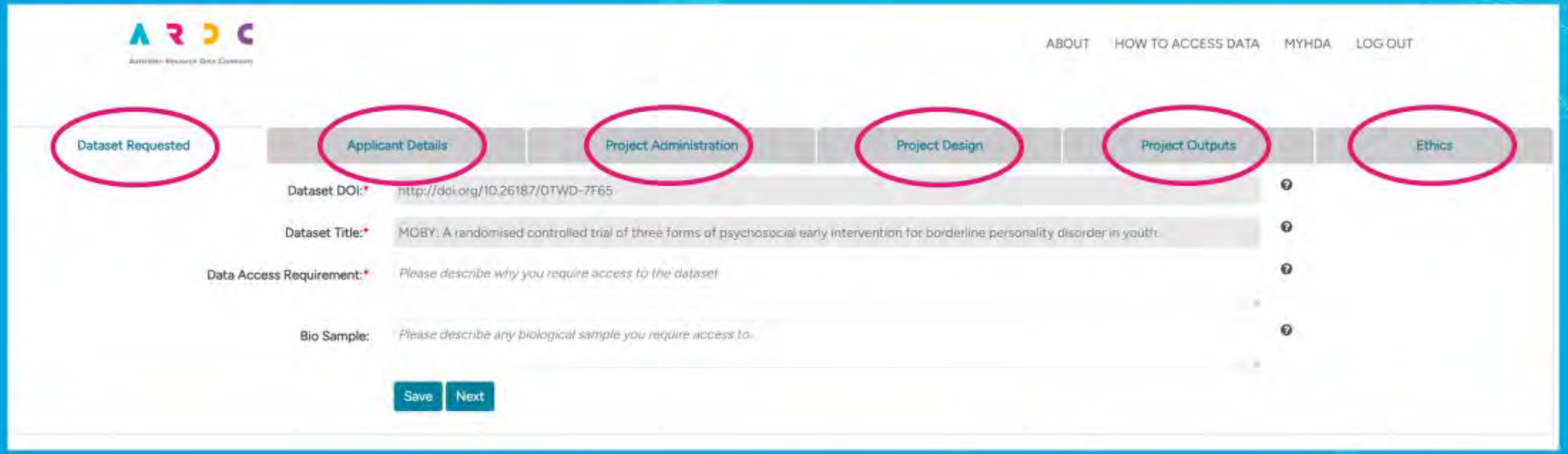
Mental Health Node

MONASH University

ARDC

NCRIS

EOI Form



MACH Melbourne Academic Centre for Health

Mental Health Node

MONASH University

NCRIS

Northern Australia Node (NA Node)

Queensland Node

Sydney Health Partners Node

National Cancer Cooperative Trials Groups Node (NCCTG)

West Australian Health Translation Network Node (WAHTN Node)

Mental Health Node (MHN)

South Australia Node (SA Node)

Melbourne Academic Centre for Health (MACH Node)

Monash and Partners Node

MACH Melbourne Academic Centre for Health

Mental Health Node

MONASH University

ARDC

NCRIS

Impact on health research

Accelerated Research

Making health data more accessible to researchers across the country, while ensuring strict compliance with privacy laws and ethical standards.

Collaboration Opportunities

Encouraging collaborative research efforts among universities, hospitals, and other health research entities.

Improving Research Efficiency

Reducing redundancies in data collection and increasing the cost-effectiveness of health research.

Supporting Innovation

Enabling innovative research into public health issues and medical treatments by providing rich, cohesive data resources.

MACH Melbourne Academic Centre for Health

Mental Health Node

MONASH University

ARDC

NCRIS

HeSANDA User Groups

The HeSANDA Network

9 HeSANDA nodes which represent 72 health research organisations across Australia

Melbourne Academic Centre for Health (MACH) Clinical Trials Consortium Node
Administered by: The University of Melbourne

Mental Health Node
Administered by: Deakin University

Monash and Partners Node
Administered by: Monash University

National Cancer Cooperative Trials Groups Node
Administered by: Australasian Leukaemia and Lymphoma Group (ALLG)

Northern Australia Node
Administered by: Menzies School of Health Research

Queensland Node
Administered by: Health Translation Queensland in collaboration with CSIRO's Australian e-Health Research Centre and Queensland Cyber Infrastructure Foundation (QCIF).

South Australia Node
Administered by: South Australian Health and Medical Research Institute (SAHMRI)

Sydney Health Partners Node
Administered by: NHMRC Clinical Trials Centre at The University of Sydney

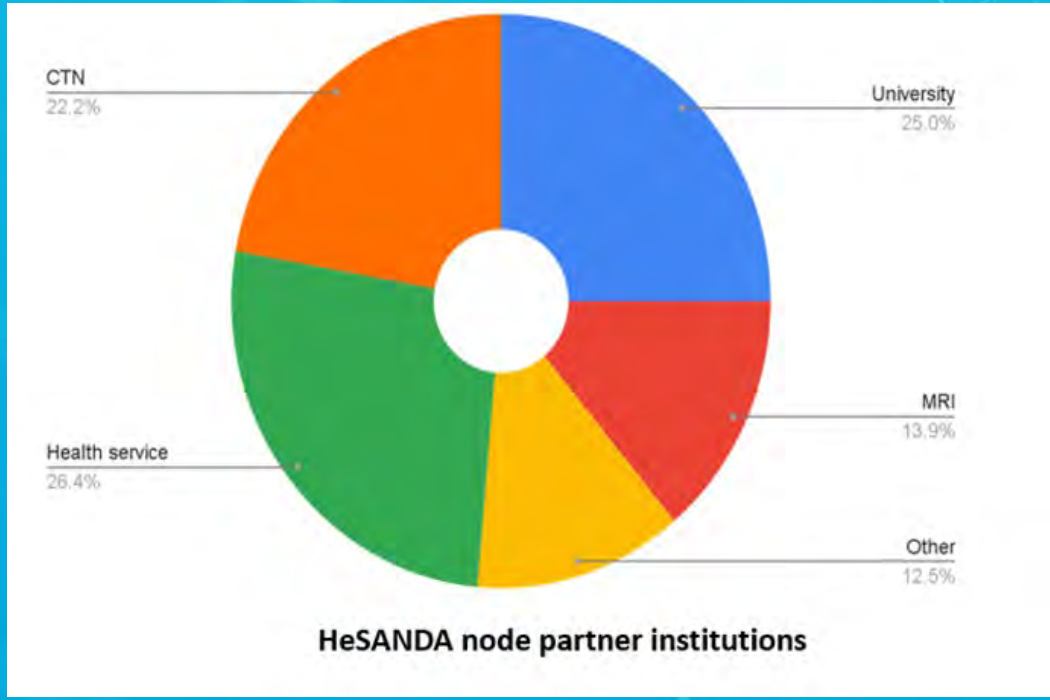
Western Australia Node
Administered by: Curtin University

HeSANDA node partners and the ARDC HeSANDA program staff at the event. Image: ARDC / Andy Steven

MACH Melbourne Academic Centre for Health
Mental Health Node
MONASH University
ARDC Australian Research Data Commons
NCRIS National Research Infrastructure for Australia An Australian Government Initiative
ARDC is enabled by NCRIS

National Structure Breakdown

16 clinical trials networks
18 universities
10 medical research institutes
19 health service operators
9 others - infrastructure services



HeSANDA node partner institutions

MACH Partners

Austin Health
Bionics Institute
Centre for Eye Research Australia
La Trobe University
Mercy Health
murdoch children's research institute
NARI
Northern Health
Olivia Newton-John Cancer Research Institute
Peter Mac
St Vincent's Hospital Melbourne
SVI St Vincent's Institute Medical Research
The Florey
The Royal Children's Hospital Melbourne
The Royal Melbourne Hospital
The Royal Victorian Eye and Ear Hospital
The Women's
The University of Melbourne
WEHI
Western Health

MACH Melbourne Academic Centre for Health
Mental Health Node
Monash University
ARDC Australian Research Data Commons
NCRIS National Research Infrastructure for Australia
ARDC is enabled by NCRIS

Profile of HeSANDA & Health Data Australia user groups

Primary Researcher – Data contributor

- Gain increased recognition and visibility for their research
- Extend the life of their research
- Meet responsibilities to share data from funders/journals
- Facilitate new collaborations to expand research impact

User Profile

Secondary Researcher

- Inspired by new research questions, or the opportunity to perform meta analysis
- Understands the opportunity to build richer datasets
- Interested in pursuing new collaboartions
- Conscious of reducing research duplication
- Needs access to detailed datasets including demographic information, clinical outcomes and treatment variables in order to conduct further robust statistical analyses

Pre-existing user skills & knowledge

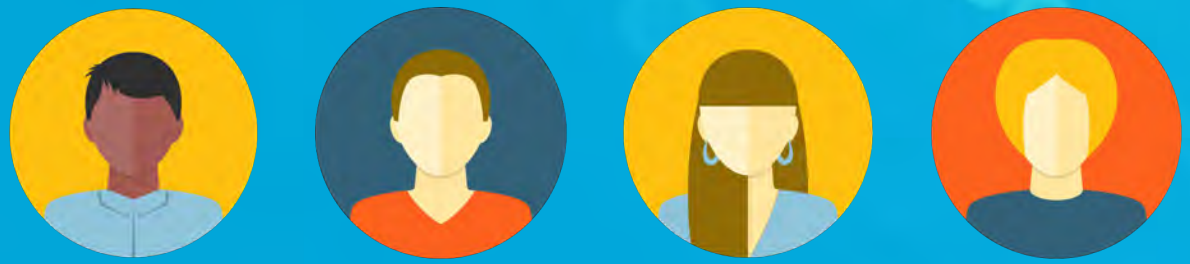
Health Data Australia catalogue

- Technically literate
- Familiar with the data access request requirements

Skills for secondary analysis

- Statistical analysis
- Data management
- Epidemiology & Biostatistics
- Medical & clinical knowledge
- Familiar with regulatory guidelines / ethical considerations
- Research methodology
- Literature review & synthesis

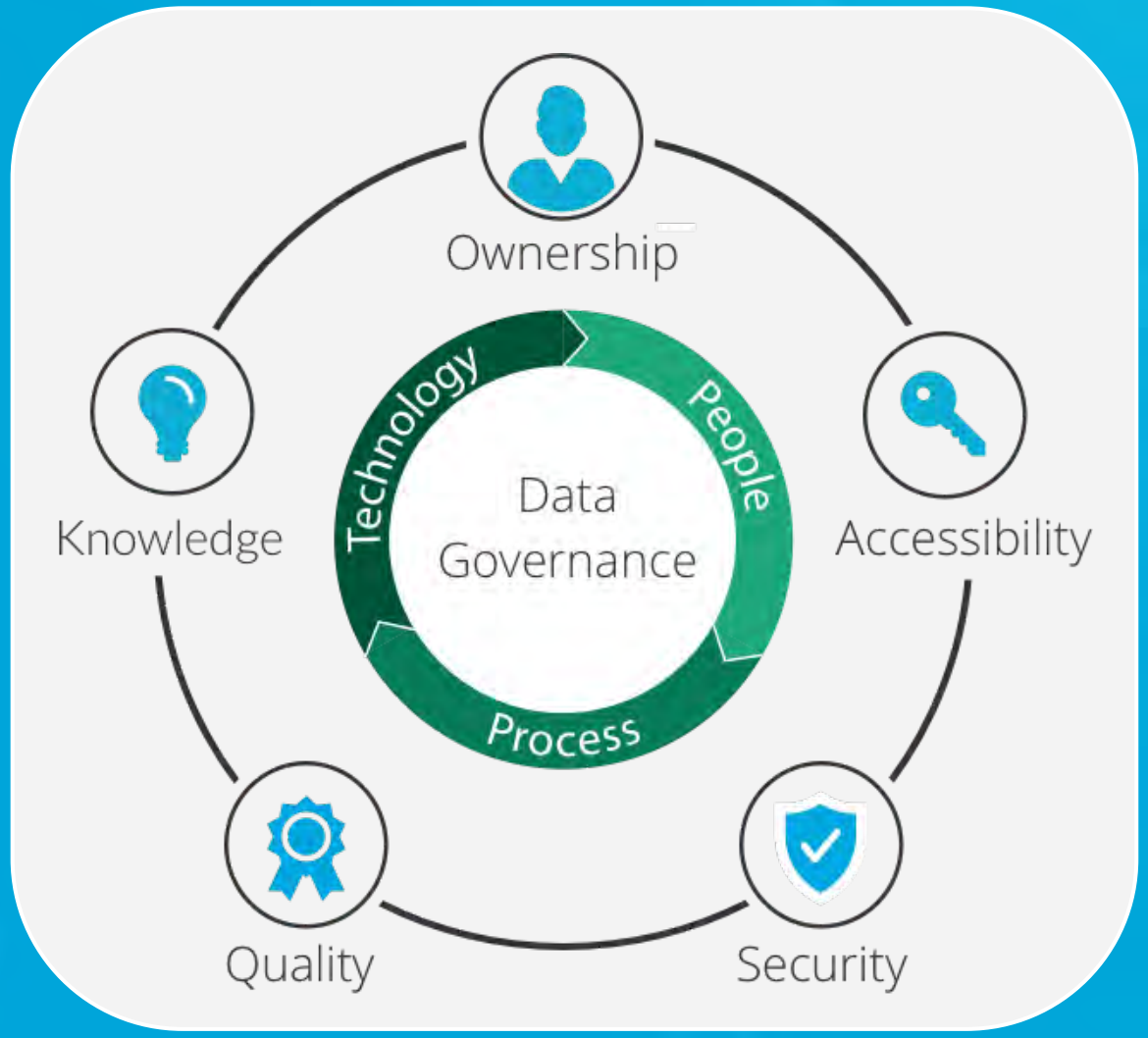
The primary researcher is responsible for…



- Review the request in accordance with your governance policies &procedures
- Approve or reject the request
- Establish terms of a data sharing agreement
- Track progress in Health Data Australia

To effectively leverage the Health Data Australia Catalogue:

- Data Governance and Policy
- Data Security Awareness
- Risk Management
- Metadata management



- Ensures Data Quality and Integrity
- Enhances Data Security and Privacy
- Facilitates Data Sharing and Collaboration
- Supports Compliance and Risk Management

MACH Melbourne Academic Centre for Health
Mental Health Node
MONASH University
Australian Research Data Commons
NCRIS National Research Infrastructure for Australia An Australian Government Initiative
ARDC is enabled by NCRIS

Engagement Plan

Outreach and Direct Engagement

- Stakeholder consultations
- Community engagement activities

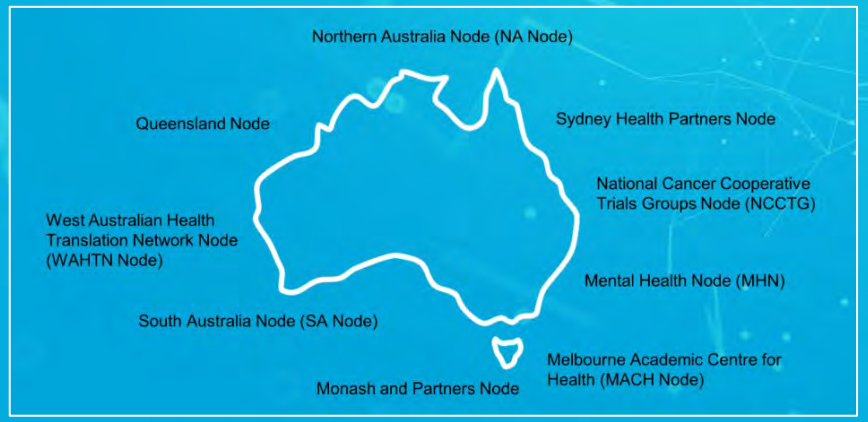
Infrastructure resources

- Guiding principles for infrastructure development

Outputs

- Consent template, SOPs & guides for data sharing
- Resources for Primary Researchers
- Resources for Secondary Researchers

MACH Melbourne Academic Centre for Health
Mental Health Node
MONASH University
ARDC Australian Research Data Commons
NCRIS National Research Infrastructure for Australia An Australian Government Initiative
ARDC is enabled by NCRIS



DOI minting

Data Access Requests

Communications

Documentation / Forms

Add your project name

PLANET RESEARCH DATA COMMONS

National-scale data infrastructure for environmental research and decision making

Image — themorningglory - 391783296 /

Australian Research Data Commons
NCRIS National Research Infrastructure for Australia An Australian Government Initiative
ARDC is enabled by NCRIS

1st Planet RDC 10 minute talk

Biosecurity Commons

A cloud-based decision-support platform for modelling and analysing biosecurity risk and response.

Presented By: James Camac

Led by - University of Melbourne

Partners

Centre of Excellence for Biosecurity Risk Analysis (University of Melbourne)

Queensland Department of Agriculture and Fisheries

Australian Government Department of Agriculture, Fisheries and Forestry

AUSVEG

The Atlas of Living Australia

Australian Research Data Commons

NCRIS National Research Infrastructure for Australia An Australian Government Initiative

ARDC is enabled by NCRIS

2nd Planet RDC 10 minute talk

WildObs Australia

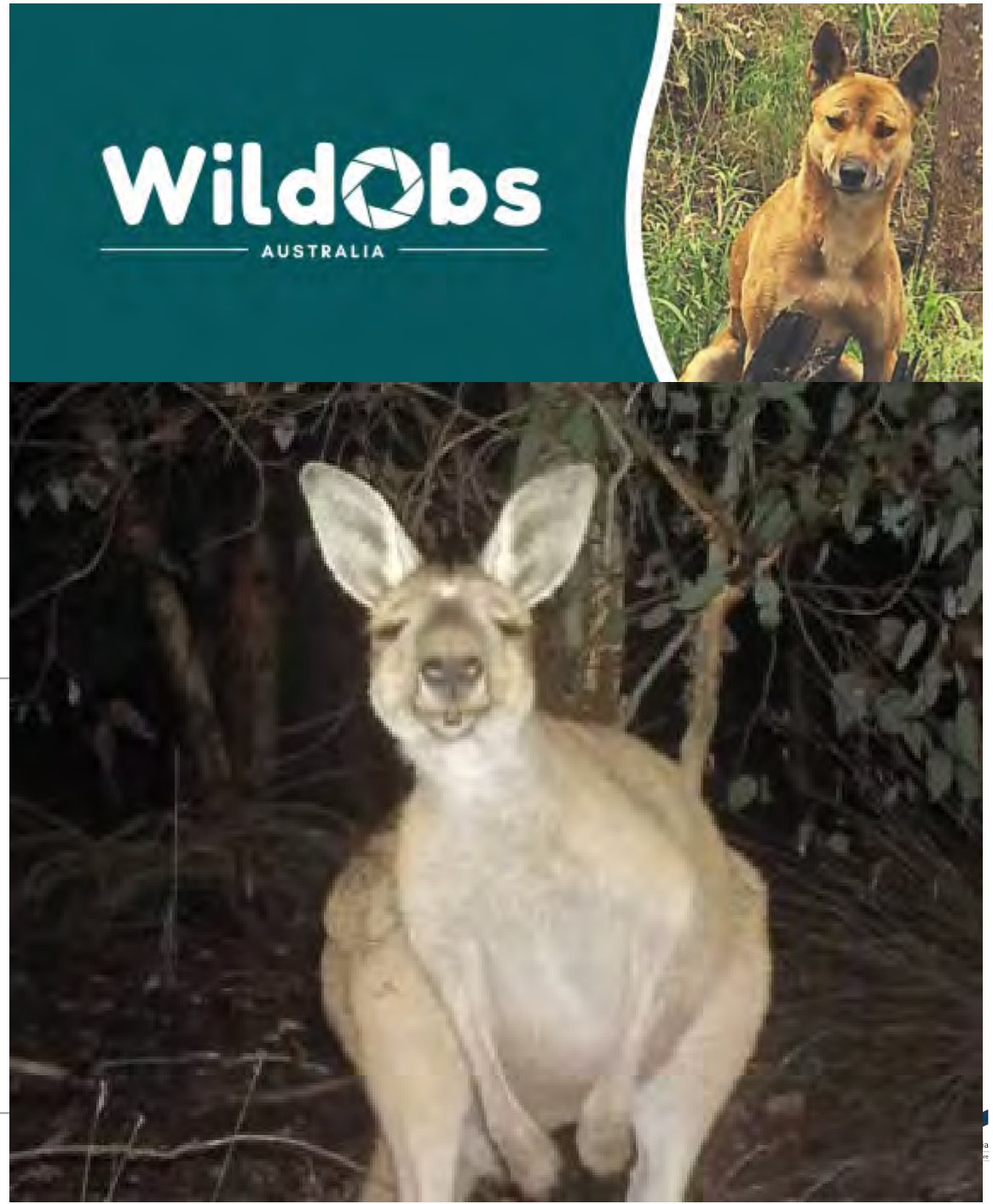
Continental-scale Wildlife Observatory for camera trap data

Presented By: Matthew Luskin

Led by - QCIF

Partners

University of Queensland
TERN
Queensland Department of Environment, Science and Innovation
Atlas of Living Australia
University of Tasmania
DCCEEW
Australian Museum
QUT
Museums Victoria
University of Sydney
Bush Heritage Australia
BirdLife Australia
WWF



ARDC is enabled by NCRIS

3rd Planet RDC 10 minute talk

Open EcoAcoustics

Continental-scale ecological monitoring and research

Presented By: Professor Paul Roe

Led by - QUT

Partners

- Atlas of Living Australia
- Queensland Department of Environment, Science and Innovation
- BirdLife Australia
- James Cook University
- The University of Queensland
- Bush Heritage Australia
- Google
- NSW Department of Primary Industries
- ACT Government, Environment, Heritage and Water Division
- University of New England
- Museums Victoria
- University of Queensland
- Australian Museum
- Australian Wildlife Conservancy
- TERN

ARDC is enabled by NCRIS

Biosecurity Commons

Dr James Camac

Chief Investigator CEBRA,
University of Melbourne, &
Project Manager
Biosecurity Commons

James is a Chief Investigator within the Centre of Excellence for Biosecurity Risk Analysis (CEBRA) with a PhD in ecology from the University of Melbourne. James is responsible for ensuring Biosecurity Commons develops cutting-edge risk analytics that meet the needs of biosecurity decision-makers across all levels of government and industry.

Australian Research Data Commons

NCRIS National Research Infrastructure for Australia An Australian Government Initiative

ARDC is enabled by NCRIS

Biosecurity
COMMONS

Dr James Camac

Project Manager

- biosecuritycommons.org.au
- Twitter: @biosec_commons
- LinkedIn: Biosecurity Commons Australia
- contact@biosecuritycommons.org.au

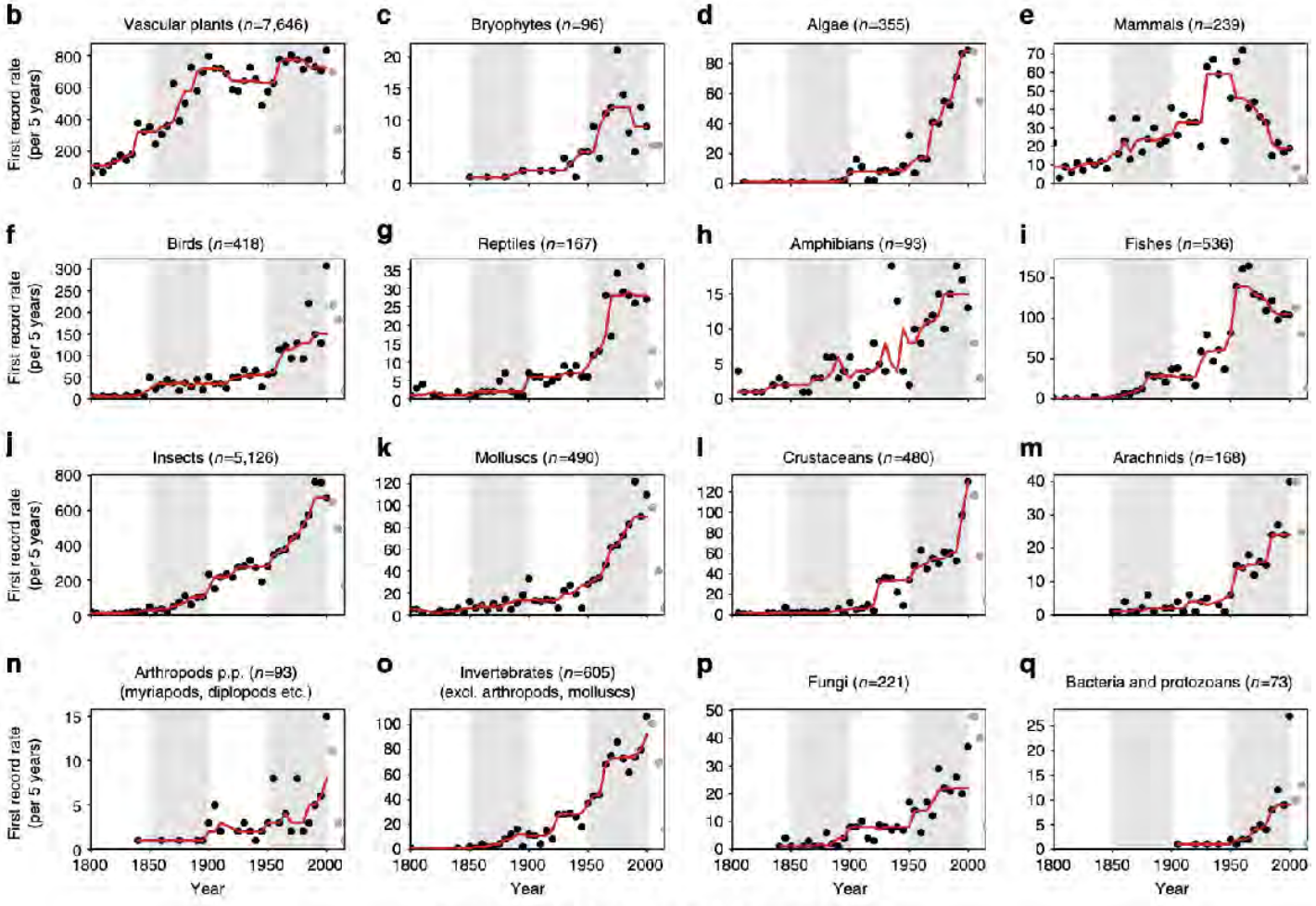
ARDC Australian Research Data Commons

NCRIS National Research Infrastructure for Australia
An Australian Government Initiative

ARDC is enabled by NCRIS

Biosecurity COMMONS

Biosecurity risk is increasing globally!



Seebens *et al.* (2017) Nature Communications

- Increasing global human movement
- Increasing trade
- Changes in species distributions

Australian Research Data Commons
NCRIS National Research Infrastructure for Australia
An Australian Government Initiative

Biosecurity COMMONS

AUS Pest Check®

Atlas of Living Australia
ala.org.au
The ALA Biosecurity Alerts Service

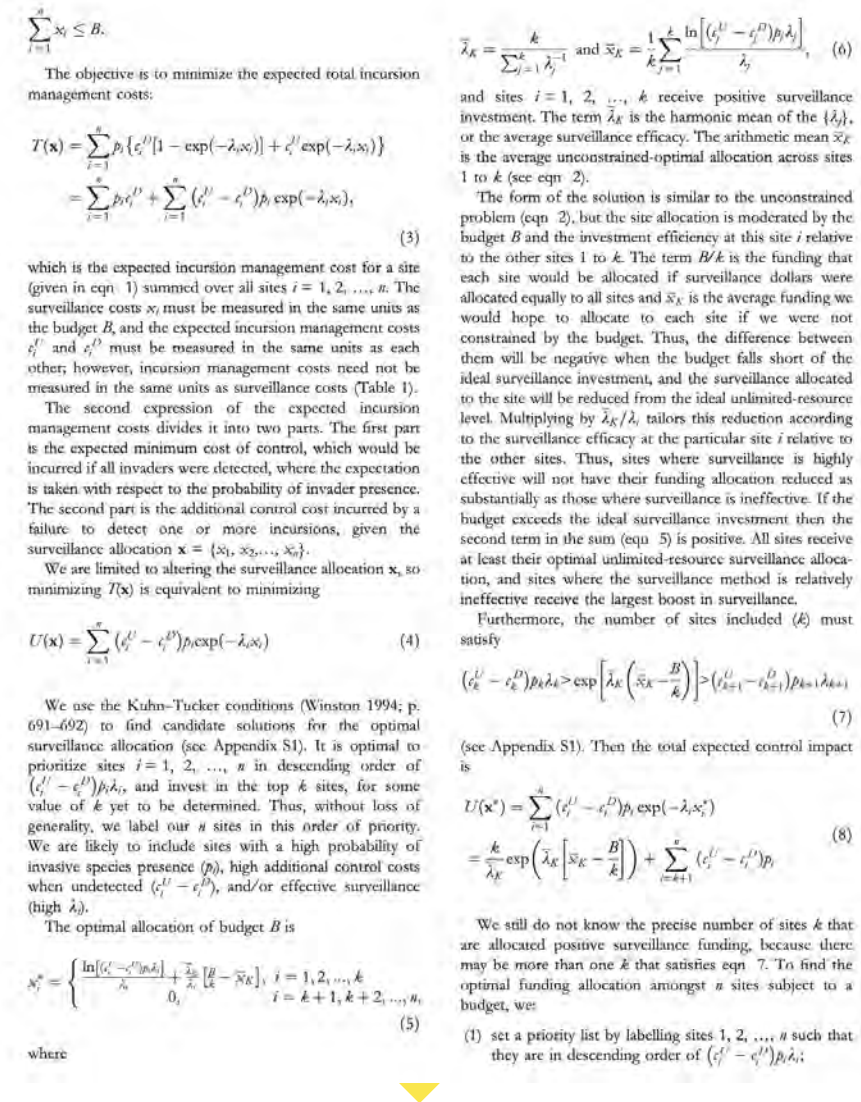
Data & Science informed risk management

- Better data harnessing capabilities
- Better screening/surveillance technology
- Better community & stakeholder awareness

Australian Research Data Commons

NCRIS National Research Infrastructure for Aus.

Biosecurity COMMONS



$$\sum_{i=1}^{n} x_i \le B.$$

The objective is to minimize the expected total incursion management costs:

$$T(\mathbf{x}) = \sum_{i=1}^{n} p_i \left\{ c_i^U [1 - \exp(-\lambda_i x_i)] + c_i^D \exp(-\lambda_i x_i) \right\}$$
$$= \sum_{i=1}^{n} p_i c_i^D + \sum_{i=1}^{n} (c_i^U - c_i^D) p_i \exp(-\lambda_i x_i), \quad (3)$$

which is the expected incursion management cost for a site (given in eqn 1) summed over all sites $i = 1, 2, \ldots, n$. The surveillance costs x_i must be measured in the same units as the budget B, and the expected incursion management costs c_i^U and c_i^D must be measured in the same units as each other; however, incursion management costs need not be measured in the same units as surveillance costs (Table 1).

The second expression of the expected incursion management costs divides it into two parts. The first part is the expected minimum cost of control, which would be incurred if all invaders were detected, where the expectation is taken with respect to the probability of invader presence. The second part is the additional control cost incurred by a failure to detect one or more incursions, given the surveillance allocation $\mathbf{x} = \{x_1, x_2, \ldots, x_n\}$.

We are limited to altering the surveillance allocation $\mathbf{x}$, so minimizing $T(\mathbf{x})$ is equivalent to minimizing

$$U(\mathbf{x}) = \sum_{i=1}^{n} (c_i^U - c_i^D) p_i \exp(-\lambda_i x_i) \quad (4)$$

We use the Kuhn–Tucker conditions (Winston 1994; p. 691–692) to find candidate solutions for the optimal surveillance allocation (see Appendix S1). It is optimal to prioritize sites $i = 1, 2, \ldots, n$ in descending order of $(c_i^U - c_i^D) p_i \lambda_i$, and invest in the top k sites, for some value of k yet to be determined. Thus, without loss of generality, we label our n sites in this order of priority. We are likely to include sites with a high probability of invasive species presence (p_i), high additional control costs when undetected ($c_i^U - c_i^D$), and/or effective surveillance (high λ_i).

The optimal allocation of budget B is

$$x_i^* = \begin{cases} \frac{\ln[(c_i^U - c_i^D) p_i \lambda_i]}{\lambda_i} + \frac{\bar{\lambda}_K}{\lambda_i} \left[\frac{B}{k} - \bar{x}_K \right], & i = 1, 2, \ldots, k \\ 0, & i = k+1, k+2, \ldots, n, \end{cases} \quad (5)$$

where

$$\bar{\lambda}_K = \frac{k}{\sum_{j=1}^{k} \lambda_j^{-1}} \text{ and } \bar{x}_K = \frac{1}{k} \sum_{j=1}^{k} \frac{\ln\left[(c_j^U - c_j^D) p_j \lambda_j\right]}{\lambda_j}, \quad (6)$$

and sites $i = 1, 2, \ldots, k$ receive positive surveillance investment. The term $\bar{\lambda}_K$ is the harmonic mean of the $\{\lambda_j\}$, or the average surveillance efficacy. The arithmetic mean $\bar{x}_K$ is the average unconstrained-optimal allocation across sites 1 to k (see eqn 2).

The form of the solution is similar to the unconstrained problem (eqn 2), but the site allocation is moderated by the budget B and the investment efficiency at this site i relative to the other sites 1 to k. The term B/k is the funding that each site would be allocated if surveillance dollars were allocated equally to all sites and $\bar{x}_K$ is the average funding we would hope to allocate to each site if we were not constrained by the budget. Thus, the difference between them will be negative when the budget falls short of the ideal surveillance investment, and the surveillance allocated to the site will be reduced from the ideal unlimited-resource level. Multiplying by $\bar{\lambda}_K / \lambda_i$ tailors this reduction according to the surveillance efficacy at the particular site i relative to the other sites. Thus, sites where surveillance is highly effective will not have their funding allocation reduced as substantially as those where surveillance is ineffective. If the budget exceeds the ideal surveillance investment then the second term in the sum (eqn 5) is positive. All sites receive at least their optimal unlimited-resource surveillance allocation, and sites where the surveillance method is relatively ineffective receive the largest boost in surveillance.

Furthermore, the number of sites included (k) must satisfy

$$(c_k^U - c_k^D) p_k \lambda_k > \exp\left[\bar{\lambda}_K \left(\bar{x}_K - \frac{B}{k} \right) \right] > (c_{k+1}^U - c_{k+1}^D) p_{k+1} \lambda_{k+1} \quad (7)$$

(see Appendix S1). Then the total expected control impact is

$$U(\mathbf{x}^*) = \sum_{i=1}^{n} (c_i^U - c_i^D) p_i \exp(-\lambda_i x_i^*)$$
$$= \frac{k}{\lambda_K} \exp\left(\bar{\lambda}_K \left[\bar{x}_K - \frac{B}{k} \right] \right) + \sum_{i=k+1}^{n} (c_i^U - c_i^D) p_i \quad (8)$$

We still do not know the precise number of sites k that are allocated positive surveillance funding, because there may be more than one k that satisfies eqn 7. To find the optimal funding allocation amongst n sites subject to a budget, we:

(1) set a priority list by labelling sites 1, 2, ..., n such that they are in descending order of $(c_i^U - c_i^D) p_i \lambda_i$;

What is missing?

A standardized system for developing risk analytics for decision-making

- Limited access to cutting edge tools & data
- National inconsistencies in application
- Limited sharing of risk analytics
- Sub-optimal outcomes

Australian Research Data Commons

NCRIS National Research Infrastructure for Australia. An Australian Government Initiative

ARDC is enabled by NCRIS

Biosecurity COMMONS

What is Biosecurity Commons

A cloud-based decision-support platform for modelling and analysing biosecurity risk and response.

cebra
Centre of Excellence for Biosecurity Risk Analysis

THE UNIVERSITY OF MELBOURNE

Griffith UNIVERSITY

Australian Government
Department of Agriculture, Fisheries and Forestry

ARDC
Australian Research Data Commons

NCRIS
National Research Infrastructure for Australia
An Australian Government Initiative

Atlas of Living Australia
ala.org.au

Queensland Government

EcoCommons

AUSVEG

Forest & Wood Products Australia

Biosecurity COMMONS

Scientific Workflows

- Published methods
- Translated into easy-to-use tools
- Linked workflows (outputs → inputs)
- 60k+ of curated and user-imported datasets

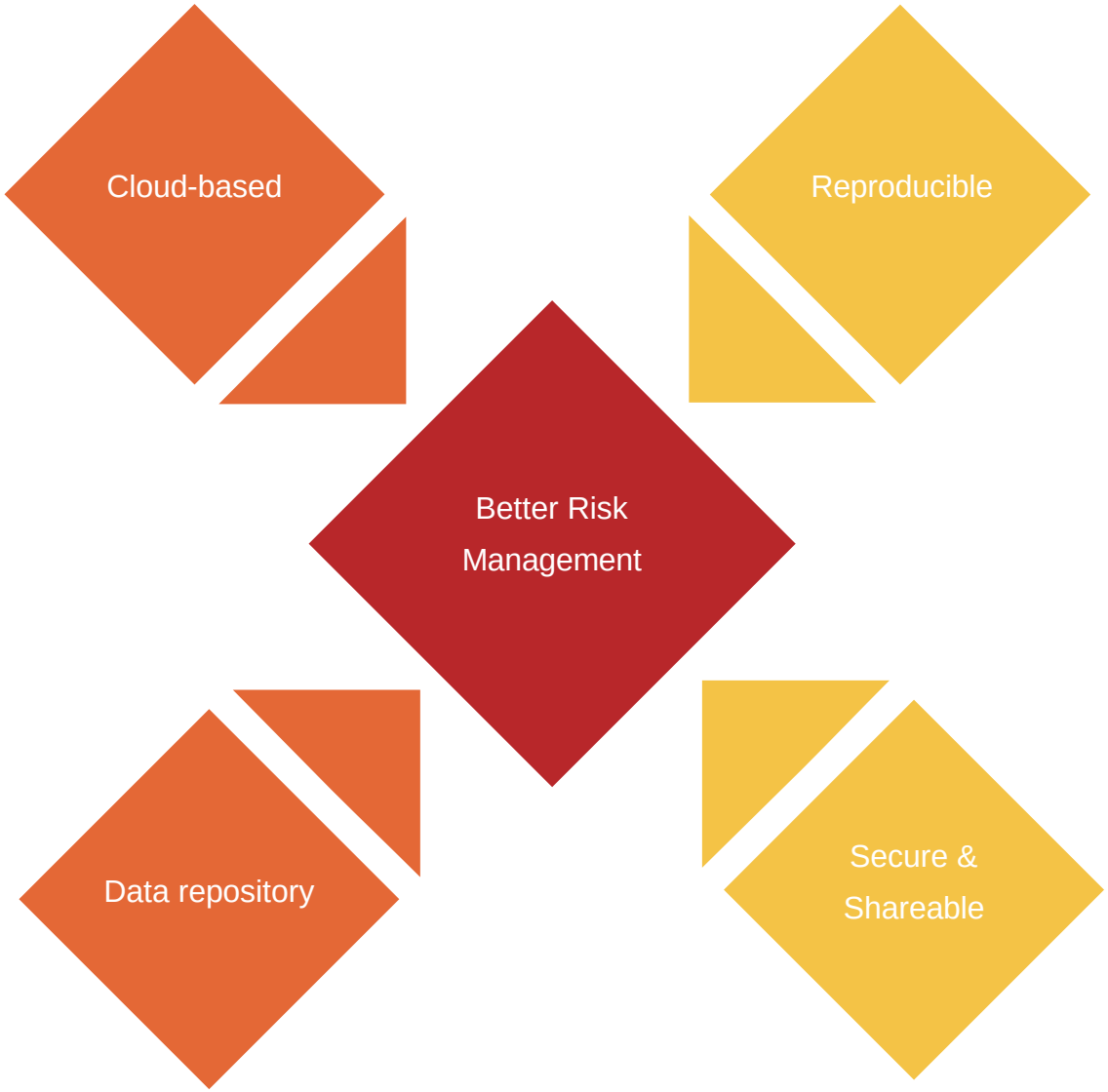
Australian Research Data Commons

NCRIS National Research Infrastructure for Australia

Biosecurity COMMONS

Key benefits

- Easier access to data & cutting-edge analytics
- Standardized & Reproducible!
- Fosters collaboration & sharing
- Reduces redundancy among organizations
- Opportunities for **nationally endorsed** frameworks for risk analytics & how they inform decision-making



Australian Research Data Commons
NCRIS National Research Infrastructure for Aus.
An Australian Government Initiative

Biosecurity COMMONS

Who are our intended users?

Governments

Biosecurity operations & policy

Industry & environmental groups

Risk management

Asset protection

Universities & research centers

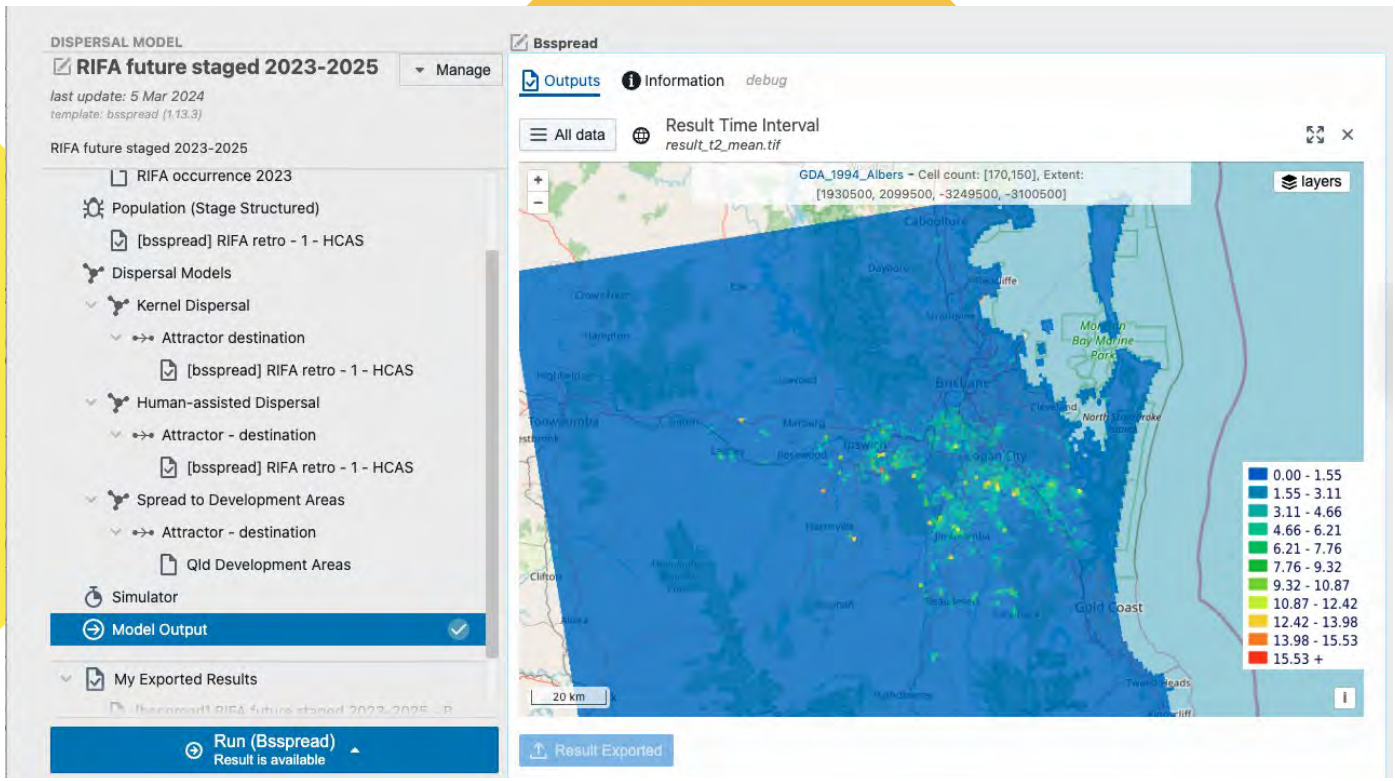
Tool developers

Training next generation

Australian Research Data Commons

NCRIS National Research Infrastructure for Aus. An Australian Government Initia...

Biosecurity COMMONS



Case study: Predicting the spread of Red Imported Fire Ant in and around Brisbane

How are we engaging?

Multi-prong approach:

Online user requirement workshops

(*develop what users need!*)

Collaborate to develop '*real world*' case studies

(Governments & industry)

Conferences, live demonstrations, training workshops

(Highlight capabilities to wide audience)

YouTube & Social media

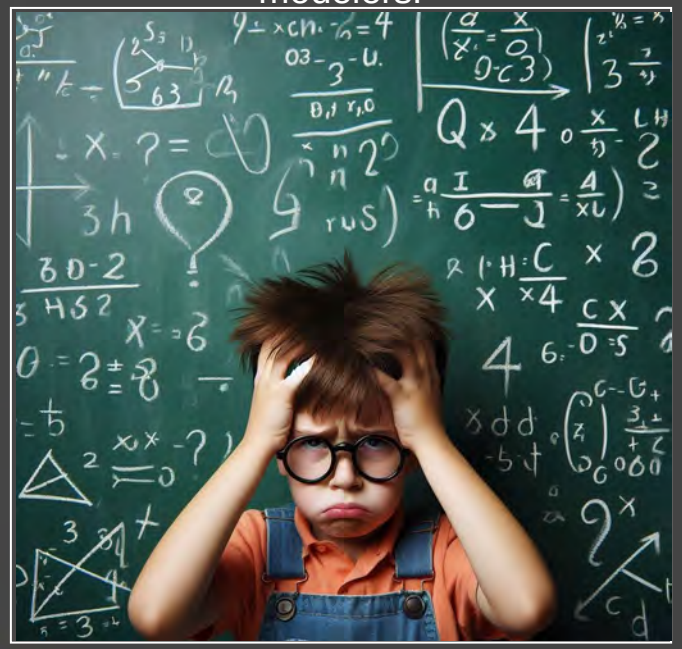
(Platform socialization & training videos modules)

Australian Research Data Commons

NCRIS National Research Infrastructure for Aus.

Challenges ahead

Risk analytics is not just for modelers!



Cognitive biases

Time poor individuals

High staff turnover in some organisations!

Data sharing & collaboration

Many data owners, different interests/objectives, historical distrust.

Sustainable funding

Who pays?

What is the value proposition?

Biosecurity COMMONS

Thank you

Biosecurity Commons Project Team

biosecuritycommons.org.au

Australian Research Data Commons

NCRIS National Research Infrastructure for Australia
An Australian Government Initiative

ARDC is enabled by NCRIS

The Wildlife Observatory of Australia

See recorded video of this presentation

Dr Matthew Luskin

Project Manager WildObs,
Senior Research Fellow

University of Queensland

Matthew is the Director of the Wildlife Observatory of Australia (WildObs), which is an ARDC Planet supported project in partnership with TERN and QCIF. WildObs is a platform for sharing, processing and analysing wildlife data collected using camera traps. Matthew is also a Senior Lecturer and Chief Investigator at the University of Queensland where he heads the Ecological Cascades Lab that examines food web ecology, plant–animal interactions, and conservation science.

Australian Research Data Commons

NCRIS National Research Infrastructure for Australia An Australian Government Initiative

ARDC is enabled by NCRIS

Open EcoAcoustics

See recorded video of this presentation

Professor Paul Roe

Head of School, Computer Science, Faculty of Science, QUT, & Project Lead of Open EcoAcoustics

Professor Paul Roe received his PhD from the University of Glasgow. At QUT, he founded and led the Microsoft-QUT eResearch Centre, and he has published over 200 papers, received over $10M in competitive research funding. Recently, he has created the world's first acoustic observatory to acoustically monitor biodiversity across Australia.

Australian Research Data Commons

NCRIS National Research Infrastructure for Australia An Australian Government Initiative

ARDC is enabled by NCRIS

HASS and Indigenous Research Data Commons

What we learned about skills needs in our co-design process

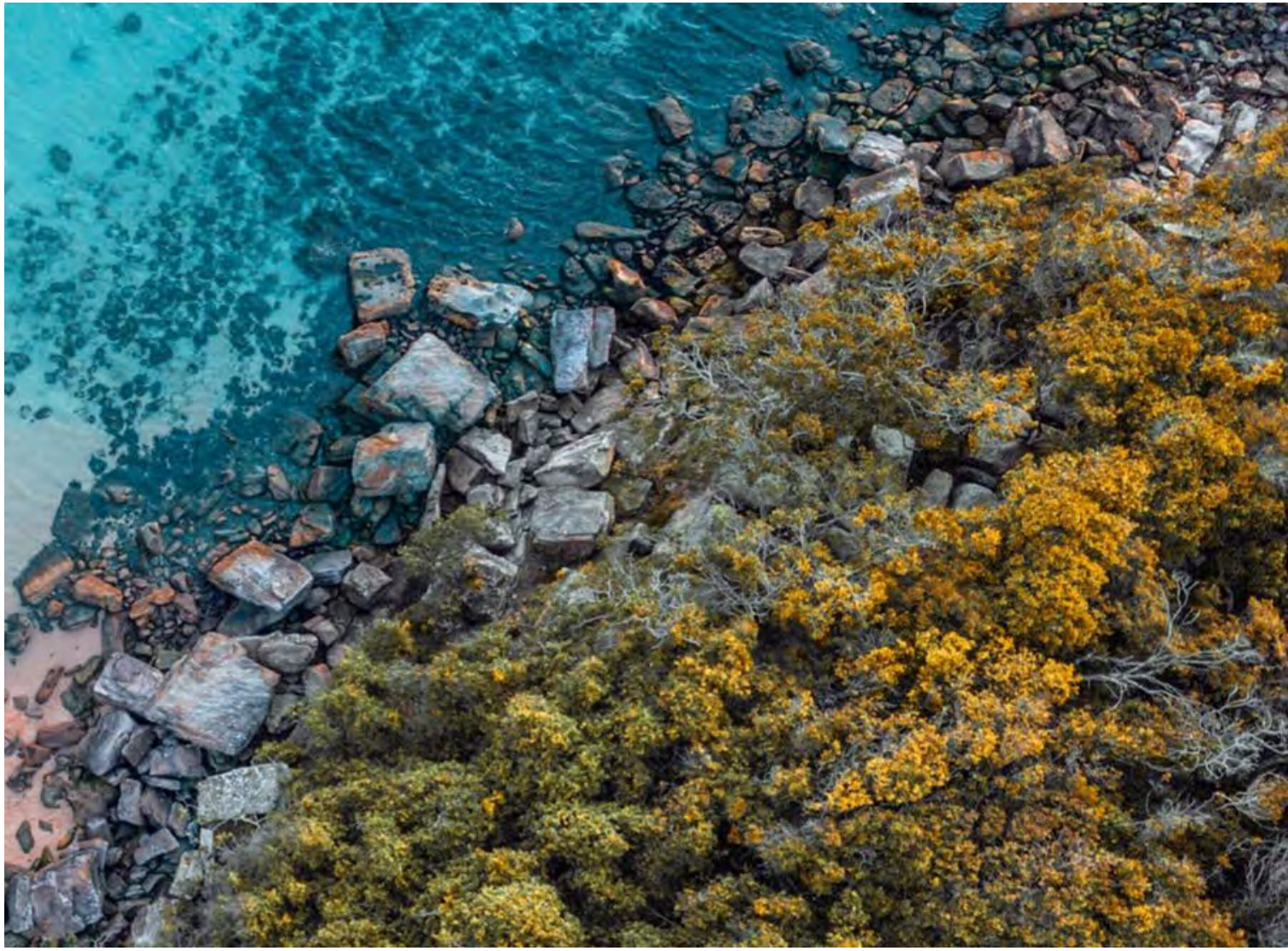
22 May 2024

PRESENTED BY
Jenny Fewster, Director
Kit Greenhill, Skills Development Lead

Australian Research Data Commons
NCRIS National Research Infrastructure for Australia An Australian Government Initiative
ARDC is enabled by NCRIS

ACKNOWLEDGEMENT OF COUNTRY

We acknowledge and celebrate the First Australians on whose traditional lands we meet, and we pay our respect to their elders past, present and emerging.



Australian Research Data Commons
NCRIS National Research Infrastructure for Australia An Australian Government Initiative
ARDC is enabled by NCRIS

HUMANITIES, ARTS, SOCIAL SCIENCES (HASS) AND INDIGENOUS RESEARCH DATA COMMONS

National-scale data infrastructure for HASS and Indigenous research and decision making

Image — Mario Purisic - jG1z5o7NCq4 /

Australian Research Data Commons

NCRIS National Research Infrastructure for Australia
An Australian Government Initiative

ARDC is enabled by NCRIS

1. The HASS & Indigenous Research Data Commons

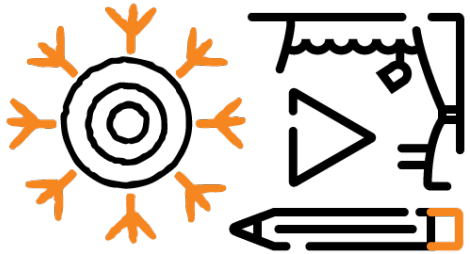
HUMANITIES, ARTS, SOCIAL SCIENCES (HASS) AND INDIGENOUS RESEARCH DATA COMMONS

Improving Indigenous Research Capabilities

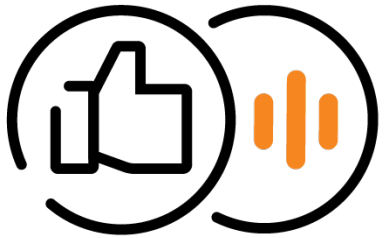
ARDC Community Data Lab

Language Data Commons of Australia

Social Sciences



Creative Arts



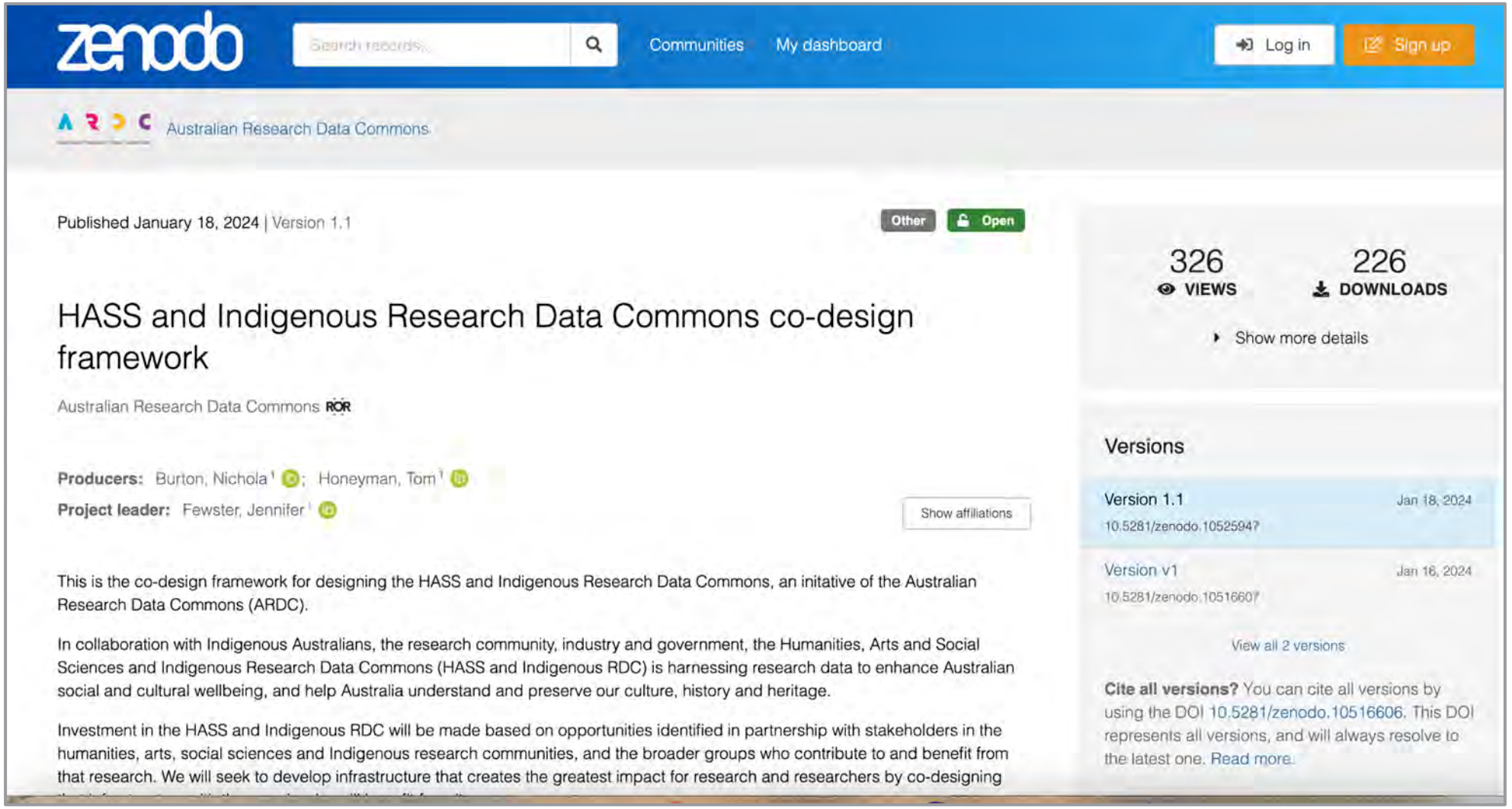
Media(ted) Data

Australian Research Data Commons

NCRIS National Research Infrastructure for Australia An Australian Government Initiative

ARDC is enabled by NCRIS

2. Co-design process



Burton, N., Honeyman, T., & Fewster, J. (2024). *HASS and Indigenous Research Data Commons co-design framework*. https://doi.org/10.5281/zenodo.10525947

What kind of challenge/opportunity?

Must be:

- Data/digital research infrastructure
- National in scope
- Collaborative

And there must be partner organisations in a position to help develop the solution (including providing co-investment)

What kind of challenge/opportunity?

Prioritised:

- Breadth of impact
- Importance (aligned with policy priorities)
- Urgency

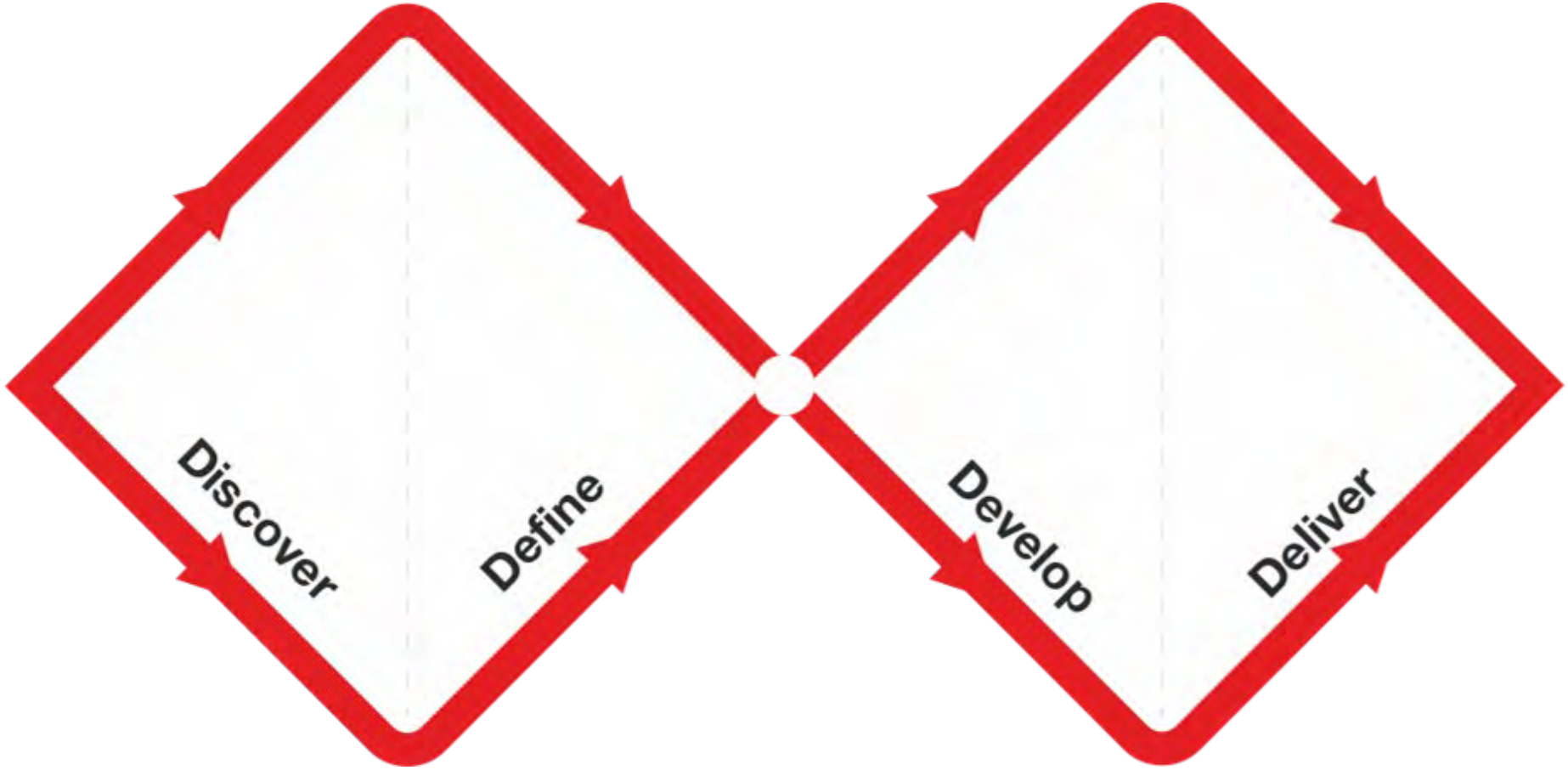
What we did

- Five focus areas Feb - May 2024
- One more area in July 2024.
- For each area:
  - One invited planning workshop - drafting problem
  - Two online workshops a fortnight apart
  - Workshop report & activity proposal released for feedback
  - Finalised agreed projects

Who we asked

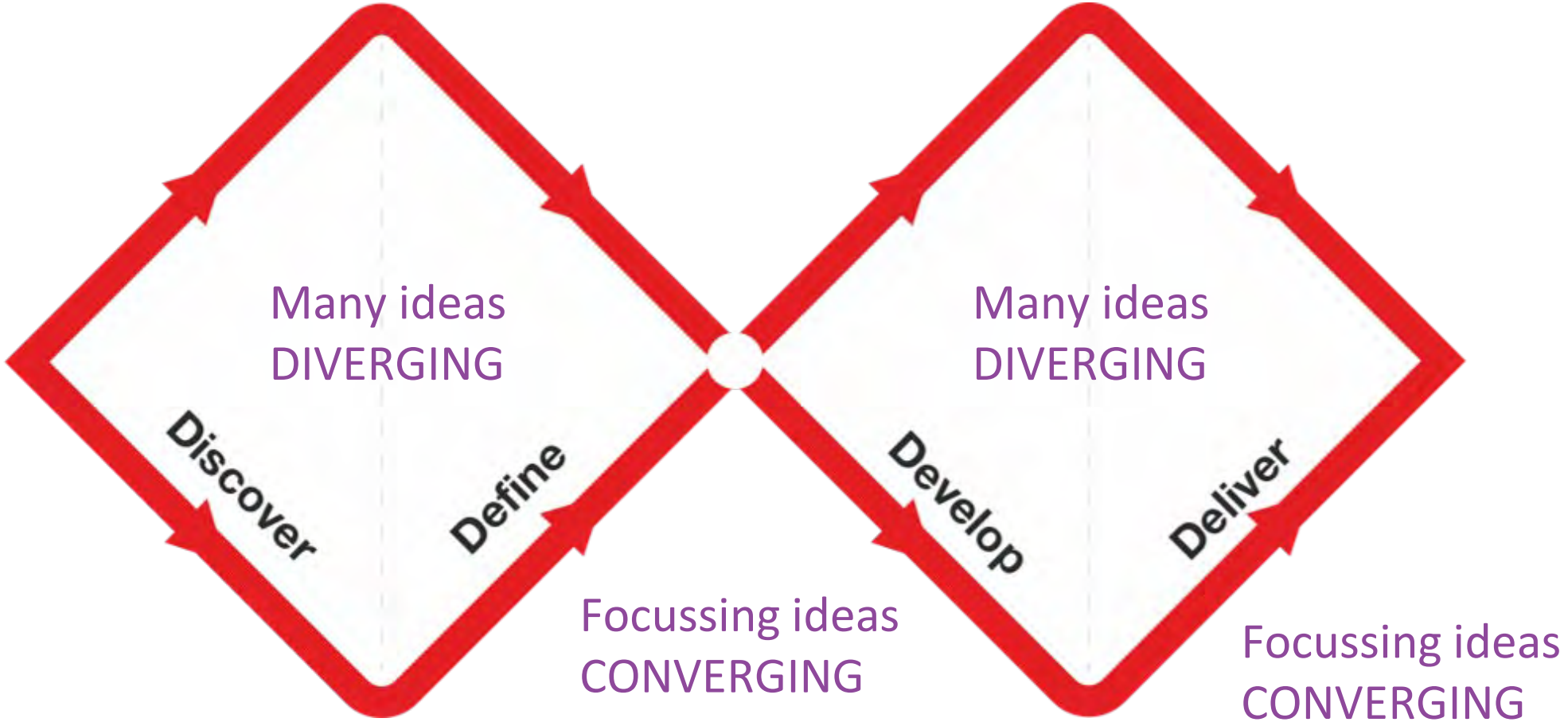
- Wide range of stakeholders, including:
  - Experienced researchers.
  - Those with insight into patterns of research practice (e.g. peak bodies).
  - Those experienced in research infrastructure.

Double Diamond Design, Design Council (UK)



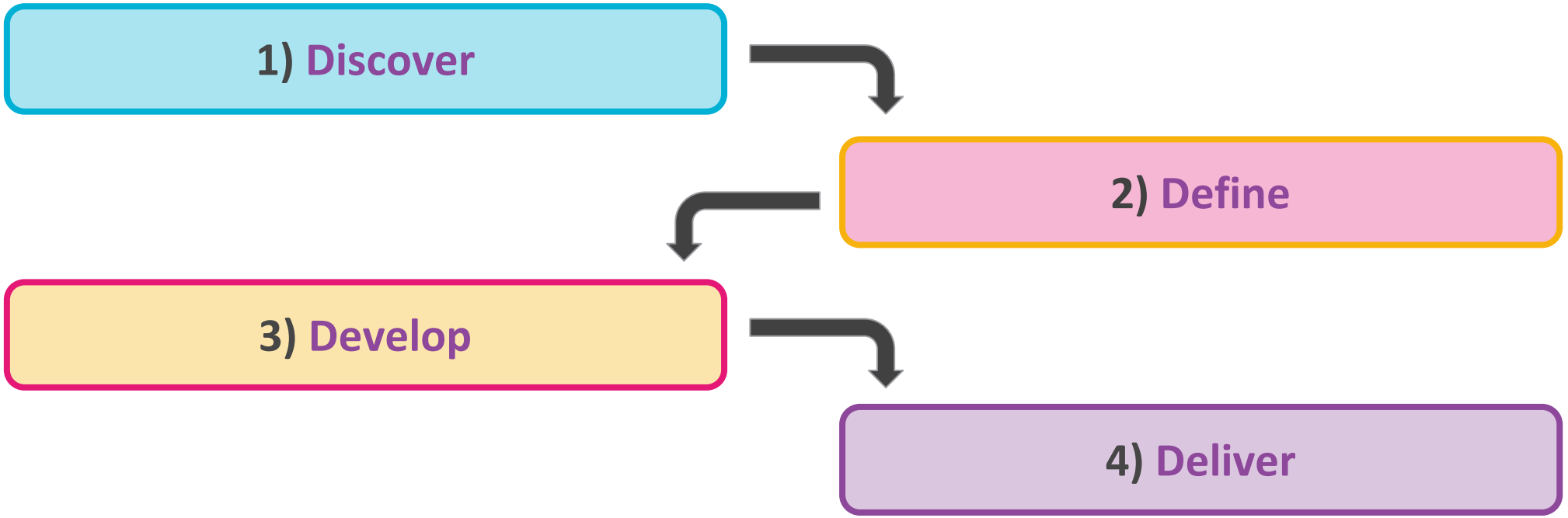
The Double Diamond by the Design Council CC BY 4.0

Double Diamond Design, Design Council (UK)



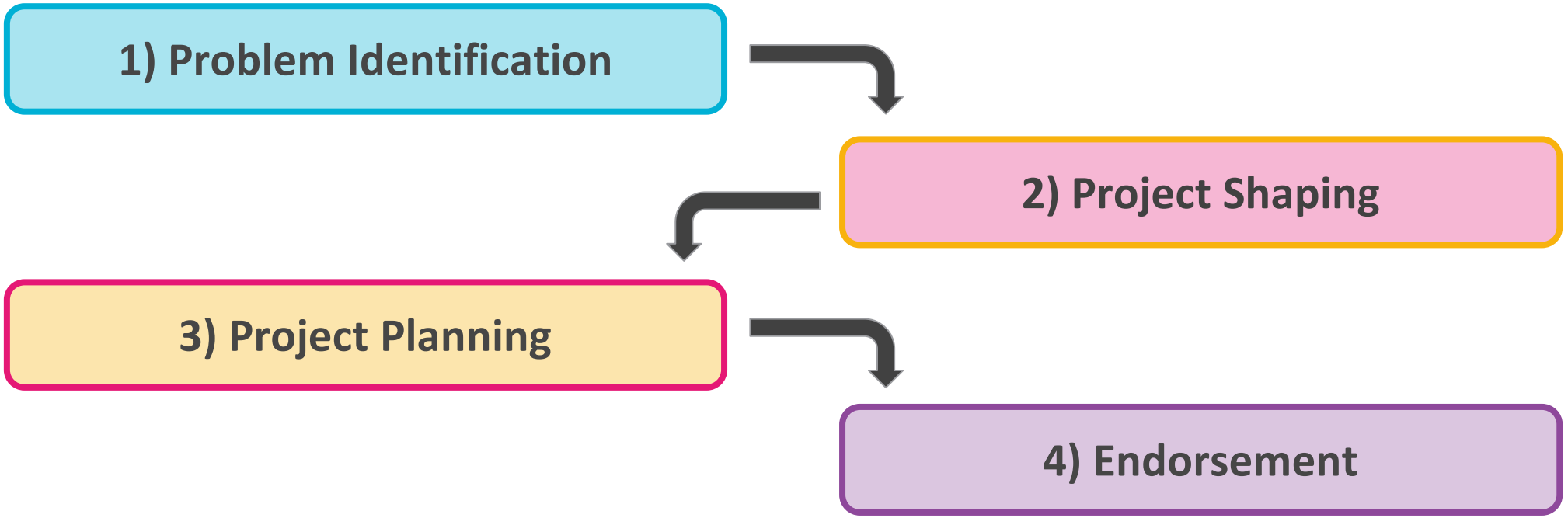
The Double Diamond by the Design Council CC BY 4.0

HASS and Indigenous RDC Co-Design Framework



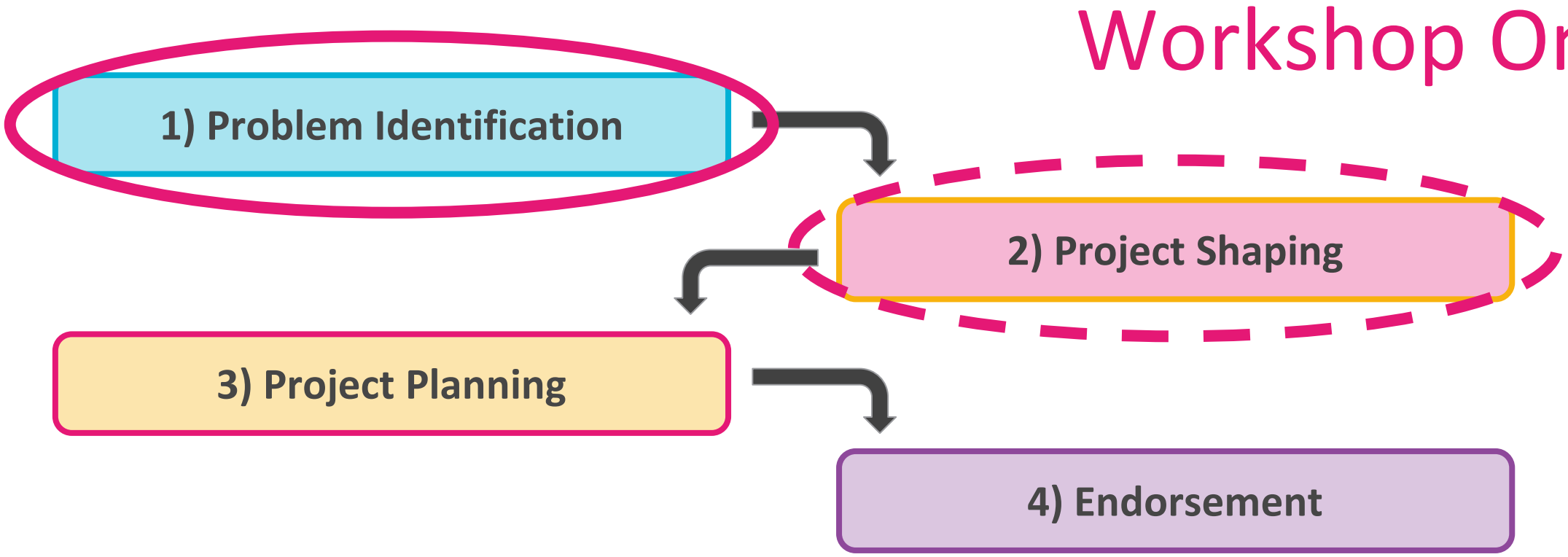
Australian Research Data Commons
NCRIS National Research Infrastructure for Australia An Australian Government Initiative
ARDC is enabled by NCRIS

HASS and Indigenous RDC Co-Design Framework



Australian Research Data Commons
NCRIS National Research Infrastructure for Australia An Australian Government Initiative
ARDC is enabled by NCRIS

HASS and Indigenous RDC Co-Design Framework



Australian Research Data Commons
NCRIS National Research Infrastructure for Australia An Australian Government Initiative
ARDC is enabled by NCRIS

Workshop One

To understand the problem and the outcomes we want to create

What we want to achieve.

Stayed away from details of **how**.

Australian Research Data Commons
NCRIS National Research Infrastructure for Australia An Australian Government Initiative
ARDC is enabled by NCRIS

Workshop One

“Blue Sky” thinking. Many divergent ideas.

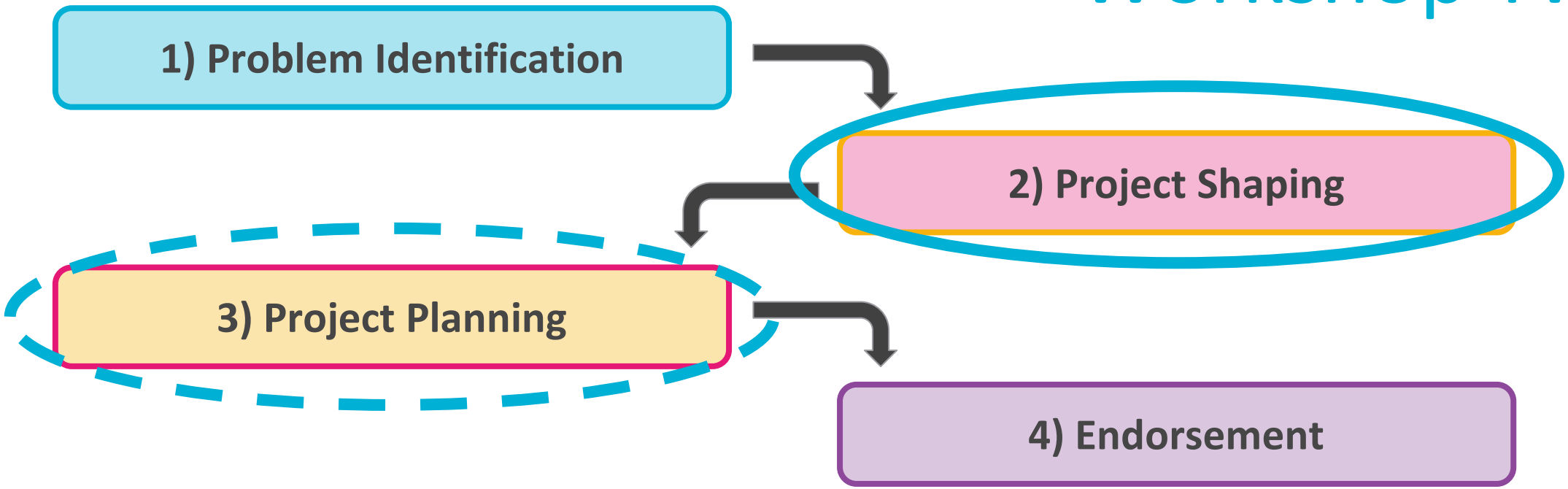
Cannot address all suggestions.

But:

- identified whether on the right track
- could adapt problem statement early
- understand context for what is co-created

HASS and Indigenous RDC Co-Design Framework

Workshop Two



Australian Research Data Commons
NCRIS National Research Infrastructure for Australia An Australian Government Initiative
ARDC is enabled by NCRIS

Workshop Two

How to achieve outcomes:

- specific solutions
- what would be practical
- what needs to be considered
- who would be involved
- measures of success.

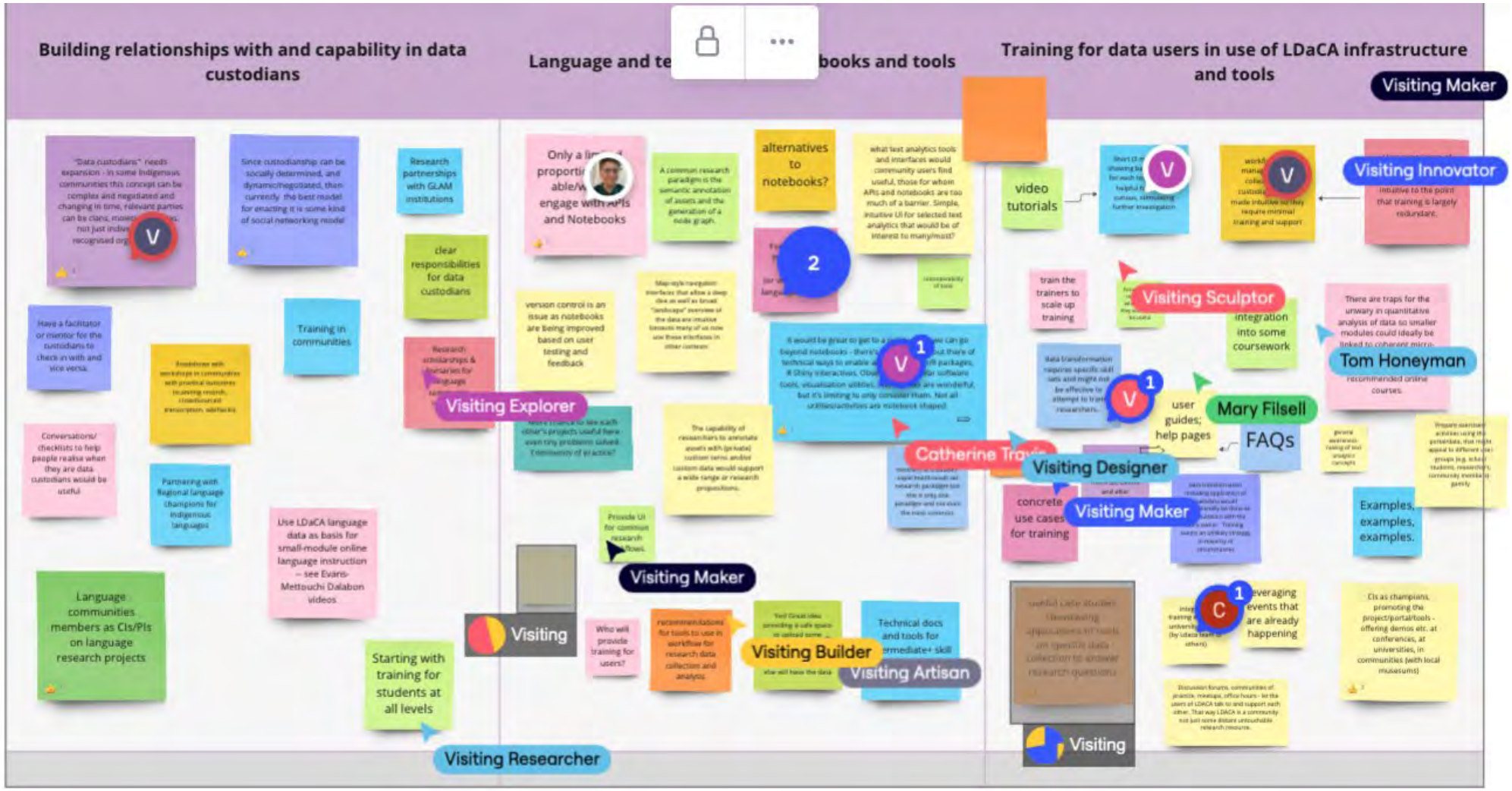
3. What we found out about skills needs

Sifting out skills

About 3000 data points.

Not all about skills.

LDaCA Workshop Two. Skills section of Miro board



Australian Research Data Commons
NCRIS National Research Infrastructure for Australia An Australian Government Initiative
ARDC is enabled by NCRIS

Sifting out skills

Nichola Burton, program manager, codified all submissions.

Codes different for each focus area, but included:

- Increased technical skills
- Skills and training
- Increase capability

Most skills feedback about:

HOW

WHO

WITH

Australian Research Data Commons
NCRIS National Research Infrastructure for Australia An Australian Government Initiative
ARDC is enabled by NCRIS

Culturally appropriate, online open resources (e.g., how-to videos, social media posts, Web pages) by & for Indigenous people involved in recording (collecting), using & sharing digital information. These resources must be done by Indigenous people with the support of technologists, but not led by (non-Indigenous) technologists.

Anonymous Participant, Improving Indigenous Research Capability *Workshop One*.

Australian Research Data Commons
NCRIS National Research Infrastructure for Australia An Australian Government Initiative
ARDC is enabled by NCRIS

HOW

Culturally appropriate, online open resources (e.g., how-to videos, social media posts, Web pages) by & for Indigenous people involved in recording (collecting), using & sharing digital information. These resources must be done by Indigenous people with the support of technologists, but not led by (non-Indigenous) technologists.

Anonymous Participant, Improving Indigenous Research Capability *Workshop One*.

Australian Research Data Commons
NCRIS National Research Infrastructure for Australia An Australian Government Initiative
ARDC is enabled by NCRIS

WHO

Culturally appropriate, online open resources (e.g., how-to videos, social media posts, Web pages) by & for Indigenous people involved in recording (collecting), using & sharing digital information. These resources must be done by Indigenous people with the support of technologists, but not led by (non-Indigenous) technologists.

Anonymous Participant, Improving Indigenous Research Capability *Workshop One*.

Australian Research Data Commons
NCRIS National Research Infrastructure for Australia An Australian Government Initiative
ARDC is enabled by NCRIS

WITH

Culturally appropriate, online open resources (e.g., how-to videos, social media posts, Web pages) by & for Indigenous people involved in recording (collecting), using & sharing digital information. These resources must be done by Indigenous people with the support of technologists, but not led by (non-Indigenous) technologists.

Anonymous Participant, Improving Indigenous Research Capability *Workshop One*.

Australian Research Data Commons
NCRIS National Research Infrastructure for Australia An Australian Government Initiative
ARDC is enabled by NCRIS

Most skills feedback about:

Australian Research Data Commons
NCRIS National Research Infrastructure for Australia An Australian Government Initiative
ARDC is enabled by NCRIS

Most skills feedback about:

HOW

Australian Research Data Commons
NCRIS National Research Infrastructure for Australia An Australian Government Initiative
ARDC is enabled by NCRIS

Most skills feedback about:

HOW

- preferred methods of delivery

Australian Research Data Commons
NCRIS National Research Infrastructure for Australia An Australian Government Initiative
ARDC is enabled by NCRIS

Most skills feedback about:

HOW

- preferred methods of delivery

WHO

Most skills feedback about:

HOW

- preferred methods of delivery

WHO

- who needs skills

Most skills feedback about:

HOW

- preferred methods of delivery

WHO

- who needs skills
- who should support skills acquisition

Most skills feedback about:

HOW

- preferred methods of delivery

WHO

- who needs skills
- who should support skills acquisition

WITH

Australian Research Data Commons
NCRIS National Research Infrastructure for Australia An Australian Government Initiative
ARDC is enabled by NCRIS

Most skills feedback about:

HOW

- preferred methods of delivery

WHO

- who needs skills
- who should support skills acquisition

WITH

- necessary conditions for skills acquisition

Not so much:

WHAT

- Topics and learning outcomes.

Most “what” will not be developed until proposals are finalised.

Basics are probably similar.

Intermediate and advanced will look different.

Australian Research Data Commons
NCRIS National Research Infrastructure for Australia An Australian Government Initiative
ARDC is enabled by NCRIS

HOW

Fit-for-purpose cookbooks & 'how to' guides.

It should be as easy to catalogue & share correctly tagged Indigenous data as make damper

Anonymous Participant, Improving Indigenous Research Capability *Workshop One*.

Australian Research Data Commons
NCRIS National Research Infrastructure for Australia An Australian Government Initiative
ARDC is enabled by NCRIS

HOW

DIY Resources

- How-to guides
- FAQs and help pages
- Self-help knowledge base
- Frameworks, guidelines and policies
- Online guides all in a single place like Libguides
- Lists of definitions of terms

HOW

Examples

- Exemplars
- Case Studies
- Pre-written code in notebooks
- Cookbooks
- Workflows
- Examples, Examples, Examples
- Checklists

HOW

Doing and making

- Data sprints
- Summer/Winter Schools
- Hackathons
- Gamified activities
- Workshops at conferences
- Workshops at events which are not obviously core audiences

HOW

Peer to peer

- Roundtables
- Peer-led teaching
- Carpentries-style events
- Student internships
- Mentoring for HDRs
- “Snowball training” - people do training then share with communities

HOW

Face to face

- Learning on Country
- Face-to-face sessions
- Office hours

HOW

Online

- Remote learning so people can stay on Country
- Webinars
- Short movies
- Open/interactive textbooks

HOW

Formal education

- Micro-courses and micro-credentials
- Formal PhD education
- Integration into undergraduate and postgraduate courses
- High school courses
- New university degrees in subjects like Indigenous Data Sovereignty
- Basic VET courses

HOW

Researchers are aware of many modalities

Ask preference.

HOW

Facilitate peer learning

HOW

If doing synchronous training, consider also producing a DIY resource, favouring examples.

HOW

If producing a DIY resource, consider synchronous training to match.

HOW

Formal education may have a coursework role to support the digital skills of potential researchers.

WHO

Who does this well? Look outside the sector -- incl STEM [and what STEM could stand to gain from creative arts]

Anonymous Participant, Creative Arts *Workshop Two*.

Australian Research Data Commons
NCRIS National Research Infrastructure for Australia An Australian Government Initiative
ARDC is enabled by NCRIS

WHO

Learners:

Researchers

- Different career levels

WHO

Learners:

University administration

- Ethics committees
- University research offices
- eResearch units

WHO

Learners:

Students

- Phd
- Undergraduate
- High School students

WHO

Learners:

Levels

- People at different stages of data lifecycle (e.g. creators vs. custodians vs. users)
- People at different levels of understanding (both basic and advanced training)

WHO

Learners:

Diversity

- Multilingual
- People with disabilities
- People from diverse cultural backgrounds
- Some may not have Internet access

WHO

Learners:

Indigenous

- Communities where the data originates (by and for)
- Prescribed Body Corporates and Local Aboriginal Land Councils
- Health workers, educators, community support workers

Australian Research Data Commons
NCRIS National Research Infrastructure for Australia An Australian Government Initiative
ARDC is enabled by NCRIS

WHO

Trainers

Australian Research Data Commons
NCRIS National Research Infrastructure for Australia An Australian Government Initiative
ARDC is enabled by NCRIS

A summit of Community workers first up. Collect ideas of what does and does not work. Perhaps consult education experts who work in Communities for their best learning design methods . Indigenous teachers, learning designers and teaching assistants.

Anonymous Participant, Improving Indigenous Research Capability *Workshop Two*.

Australian Research Data Commons
NCRIS National Research Infrastructure for Australia An Australian Government Initiative
ARDC is enabled by NCRIS

WHO

Trainers:

- Libraries and councils to target regional areas
- Communities where the data originates (by and for)
- “Bridging people” - who talk both tech and research
- Colleagues, not vendors, to train researchers

WHO

Researchers are not the only people who need digital skills training if more data-informed research is to happen in Australia.

THE RIGHT CONDITIONS

Important to link this [outreach] with training - connect with disciplinary associations, summer schools etc. so that it is more integrated with scholarly culture and not seen as 'infrastructure'

Anonymous Participant, Media(ted) Data *Workshop Two*.

WITH

Relationships

- Build relationships so right people connect to mentor each other
- Communities of practice
- Online forums to allow communication
- Better collaboration between tool-makers and researchers
- National communities of trainers

WITH

Community control

- Support by technologists but Indigenous communities determine
- Forums to find out from researchers what training they want

WITH

Improve tools

- Easier to use interfaces so training and guides are not necessary
- Allow dataset downloads so people can play
- Sandboxes for experimenting (e.g. with AI)

WITH

Clarification

- Define target audiences
- Information sessions with elders to explain what training is and the benefit to community
- Identify who is responsible for training which people

WITH

Ethics

- Ethical frameworks to guide any training
- Informed by cultural protocols
- Respect, relevance, reciprocity and responsibility as principles for working with Indigenous communities.

WITH

University systems

- Value training as a scholarly activity
- Universities formally valuing and rewarding researchers who have basic data literacy
- Clarification of university role to ensure researchers have basic digital literacy skills
- Sharing materials between universities

WITH

Fit learning journey

- Develop training early so it does not happen as afterthought
- Formal map of all training available to researchers
- Resources that can be updated and consulted long-term
- Remuneration for those being trained
- Jobs at end for people who undertake training
- Research software and management career paths (Retention issues for post-docs)

Need to try to embed training and capacity-building into other training. Community are overloaded with training, much of which leads no-where - to qualifications, to jobs etc. So don't add extra burden of training where it can be embedded in existing frameworks - e.g. land and sea Country management activities.

Anonymous Participant, Improving Indigenous Research Capability. *Workshop Two*.

Australian Research Data Commons
NCRIS National Research Infrastructure for Australia An Australian Government Initiative
ARDC is enabled by NCRIS

WITH

Skills are improved in an interconnected system of community, relationships and ethics.

Australian Research Data Commons
NCRIS National Research Infrastructure for Australia An Australian Government Initiative
ARDC is enabled by NCRIS

WITH

What looks like a skills deficit may need other things to change, not more training. Like:

- Simplifying tools
- Clarifying need
- Changing how skills are valued

WHAT

WHAT

- Indigenous Data Governance
- Knowledge of each other's techniques and tools to collaborate
- Data collection
- Data ethics
- Consent, permissions, licencing training
- Digital literacy
- Collecting, using, sharing data
- Text analytics
- Application of metadata
- FAIR and CARE

Australian Research Data Commons
NCRIS National Research Infrastructure for Australia An Australian Government Initiative
ARDC is enabled by NCRIS

4. Five Challenges

Consultation is essential, but resource-intensive.

How to support so many suggested delivery types?

Does each RDC focus area need its own skills framework due to unique topics and users?

HASS methods are often qualitative.

Much digital research skills material relates to quantitative research.

What is the role of universities when researchers do not have basic digital skills?

5. Strategies

Indigenous internship

Australian Research Data Commons

NCRIS National Research Infrastructure for Australia
An Australian Government Initiative

ARDC is enabled by NCRIS

HASS & Indigenous Computational Skills Summer School 2024

Australian Research Data Commons

NCRIS National Research Infrastructure for Australia
An Australian Government Initiative

ARDC is enabled by NCRIS

Australian Research Data Commons

NCRIS National Research Infrastructure for Australia
An Australian Government Initiative

ARDC is enabled by NCRIS

Australian Research Data Commons

NCRIS National Research Infrastructure for Australia
An Australian Government Initiative

ARDC is enabled by NCRIS

Subscribe to the **ARDC CONNECT** newsletter

THANK YOU

ardc.edu.au

contact@ardc.edu.au

@ARDC_AU

Australian-Research-Data-Commons

ARDC_AU

Australian Research Data Commons

NCRIS National Research Infrastructure for Australia
An Australian Government Initiative

ARDC is enabled by NCRIS

Identifying key challenges - Q&A

Australian Research Data Commons
NCRIS National Research Infrastructure for Australia An Australian Government Initiative
ARDC is enabled by NCRIS

Troubleshooting user training and engagement challenges

Lightning talks

Australian Research Data Commons
NCRIS National Research Infrastructure for Australia An Australian Government Initiative
ARDC is enabled by NCRIS

VICTORIA 2024
ResBaz
RESEARCH BAZAAR
MONASH University
National Sponsor
ARDC
Australian Research Data Commons
SAVE THE DATE - ResBaz is back to Victoria in 2024
We are delighted to announce that this year ResBaz will be hosted by the Monash eResearch Centre from 26-28 November, 2024 at the Monash Clayton Campus.
We are currently finalising details, so please watch out for coming updates!
ResBazVIC-l@monash.edu

Troubleshooting user training and engagement challenges

Lightning talk #1

Ann Backhaus

Education and Training Manager, Pawsey Supercomputing Research Centre

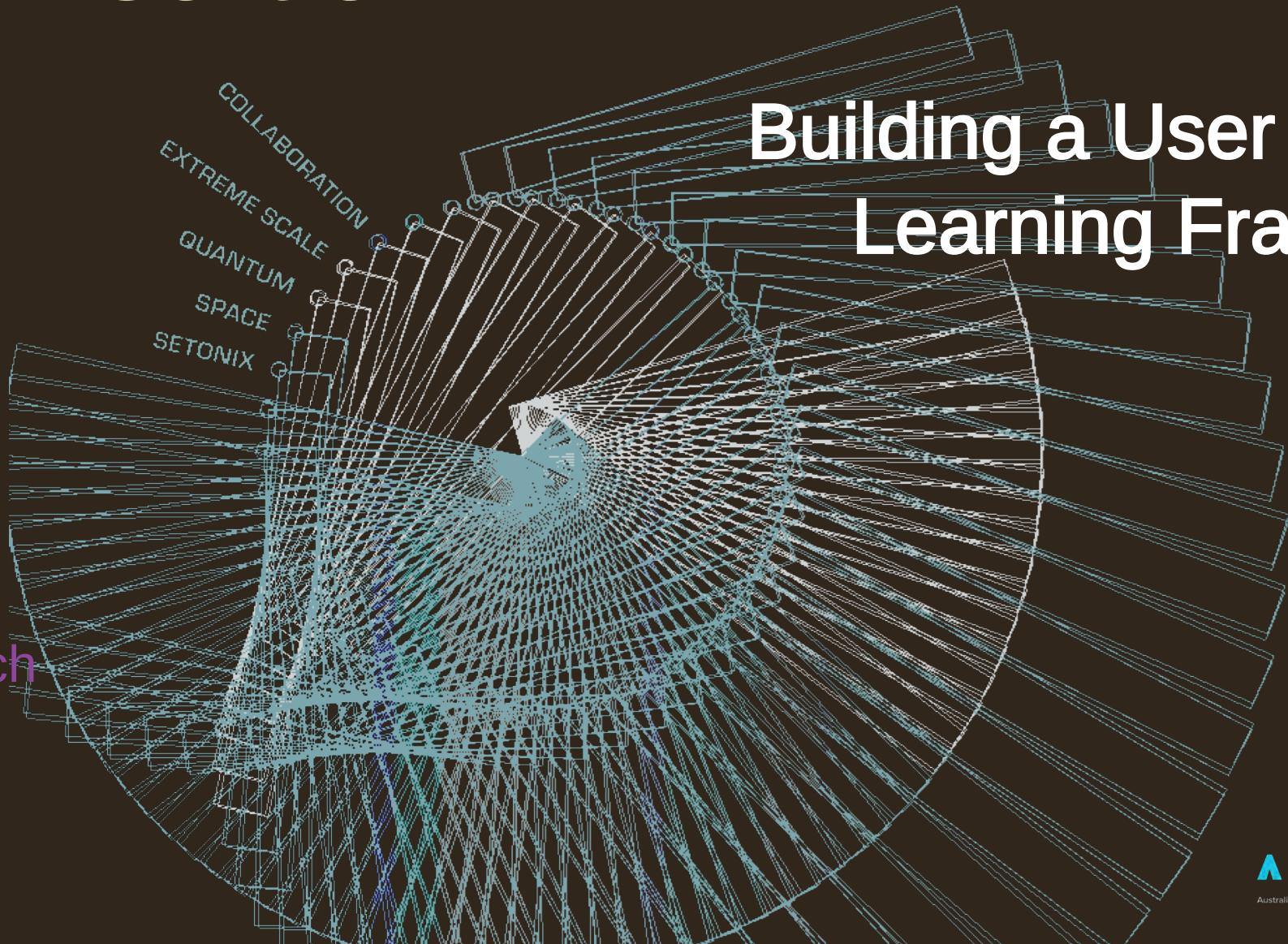
Ann Backhaus is the Education and Training Manager at Pawsey. She seeks to accelerate researcher's discoveries through the Pawsey User Training Program. In addition, Ann helps build Australia's next generation of scientists through the Internship Program, summer schools, and STEM outreach – underpinned with an extensive network of Partners. Ann also keeps Pawsey staff upskilled in current and future-looking technologies as well as essential skills, which contribute to Pawsey's "friendly vibe".

Australian Research Data Commons

NCRIS National Research Infrastructure for Australia An Australian Government Initiative

ARDC is enabled by NCRIS

Pawsey Supercomputing Research Centre



Building a User -Centric Learning Framework

Ann Backhaus

Education & Training Manager

ARDC
Digital Research Skills Summit
May 2024

Pawsey Infrastructure

200,000+ Cores, 700+ GPUs, 100Gbps

Setonix Supercomputer

Nimbus (Cloud)

ASKAP Ingest

MWA - Garrawarla

100 Gbps

100 Gbps

10 Gbps

10 Gbps

100 Gbps

S3 Interface

Fibre Channel

100Gbps

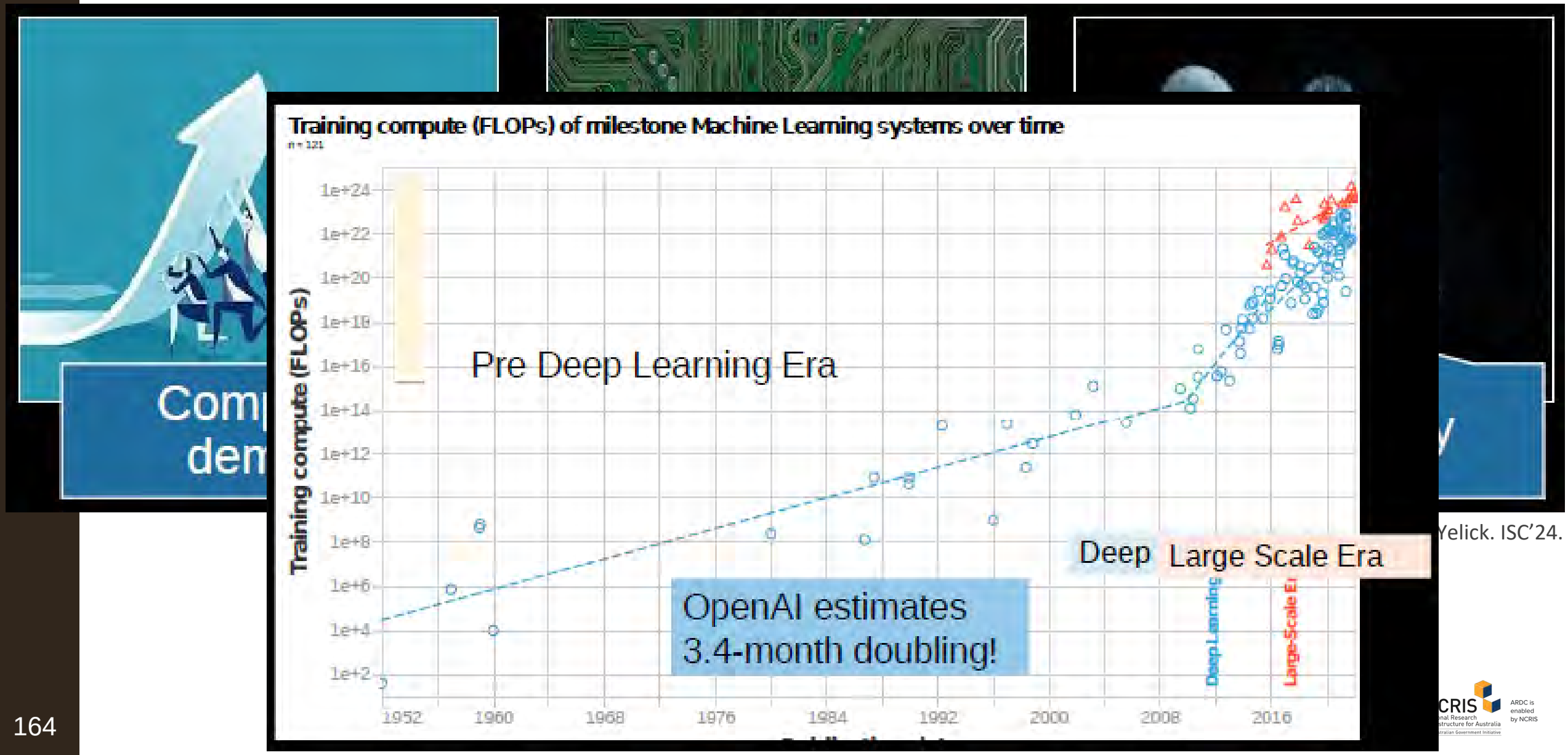
100 Gbps External Firewall

Banksia

Acacia

+ Ella – Grace Hopper nodes
Quantum computing

Business as Usual? No. The Challenge.



Yelick. ISC'24.

CRIS ARDC is enabled by NCRIS

Learners are confused. They want guidance.

Australian Research Data Commons
NCRIS National Research Infrastructure for Australia An Australian Government Initiative
ARDC is enabled by NCRIS

The Pawsey Academy.

In DRAFT

Upskilling & Building Community

Staff

upskilling

Learn @ Pawsey

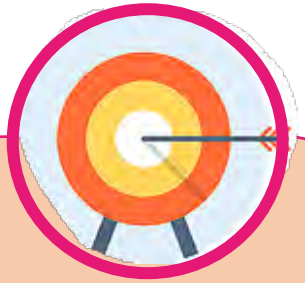
Inspire school students & teachers through STEM-driven enrichment & masterclasses.

Grow @ Pawsey

Develop work-ready students through courses, internships, & introductory HPC & digital skills.

Start @ Pawsey

Engage new & non-traditional researchers & domains with HPC, visualisation & Pawsey.



Scale @ Pawsey

Build, deepen & broaden skills in HPC & emerging technologies via Pawsey expertise & innovation.

Partners and collaborators

Australian Research Data Commons
Infrastructure for Australia
An Australian Government Initiative
ARDC is enabled by NCRIS

How does The Pawsey Academy reach its goals?

IN DRAFT

Learn@Pawsey

A STEM-driven enrichment program that:

- Inspires students and fosters curiosity; WA students; include regional & low ICSEA schools

Goals:

- Inspire through engaging STEM, career, and pathway activities
- Upskill teachers with ready-to-use materials, training
- Partner to extend awareness and education outreach

Target: National. Secondary students, teachers (Primary if time)

Reaching goals through (e.g.):

- Learn@Pawsey datasets & (associated) activities
- Hands-on science, careers
- Teacher Masterclasses & conference participation
- Leverage Intern/s & Emerging Scientists

grow@Pawsey

A Work Ready program that:

- Intersects with Australian (HPC) tertiary curricula to prepare learners for work in science, research, industry, etc.

Goals:

- Prepare through intro and intermediate training & experiences
- Collaborative upskilling
- Scalable upskilling

Target: Australian university students, HDRs, PhDs, ECRs

Reaching goals through (e.g.):

- Pawsey Internships, exchanges, summer schools
- Essential training in supercomputing, cloud, visualisation, HPC, data
- Quantum (education) hub
- Tertiary HPC-related courses, guest talks

start@Pawsey

A Pawsey Launch program that:

- Initiates use of Pawsey resource/s for new users and actively seeks users from 'non-traditional' domains

Goals:

- Provide an "on-ramp" to Pawsey resources (HPC, data, visualisation),
- Build competency in core and foundational "specialised" skills, e.g., exa-scale computing, ML, quantum, etc.

Target: Pawsey Alum (interns), new users, new partners, ECRs

Reaching goals through (e.g.):

- Allocations, uptake projs
- Pawsey domain 'guides'
- Invites to specialised (domain/targeted) trainings in HPC, data, visualisation, etc.

scale@Pawsey

A Pawsey Builder program that:

- Accelerates research outcomes by enabling users to scale their codes on Pawsey's resources
- Builds Pawsey expertise

Goals:

- Enable scaling on Pawsey resources through services and training
- Build expertise in compute, data and/or visualisation skills
- Apply best practices

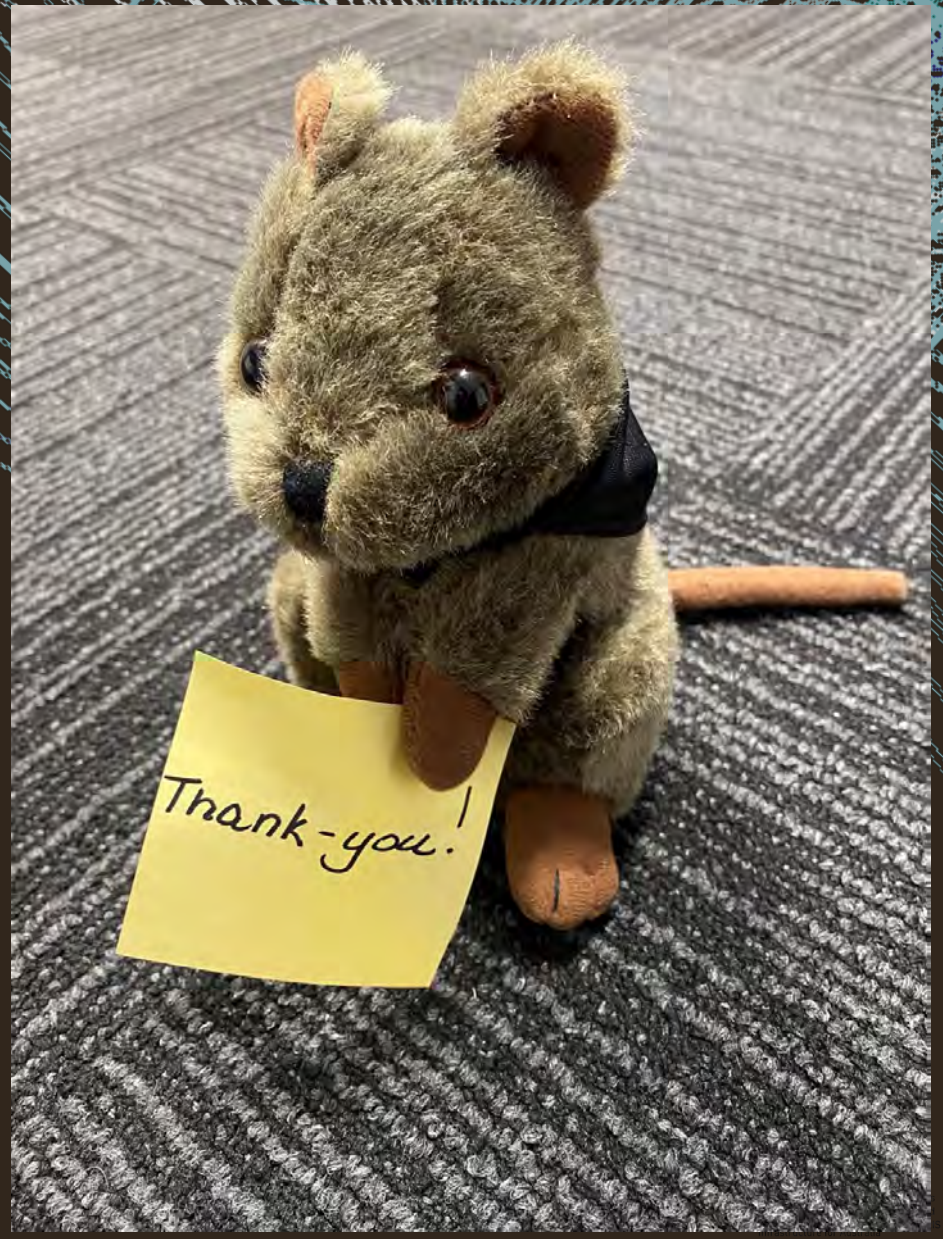
Target: Pawsey users, partners

Reaching goals through (e.g.):

- Allocations and awards
- Targeted programs (e.g., PaCER, other schemes)
- Strategic partnerships, fellowships exchanges
- Specialised training and experiences
- Conferences

pawsey

Thank you



Australian Research Data Commons

Troubleshooting user training and engagement challenges

Lightning talk #2

Dr Giorgia Mori

Data Science Trainer, SIH,
University of Sydney

Giorgia is a data scientist with a background in microbiology and bioinformatics. She is a Software, Data and Library Carpentry instructor, and she has been supervising, teaching and training students in the development of computational skills for conducting efficient and reproducible research. She volunteers in diversity initiatives (RLadies, PyLadies) to support the representation of gender minorities in the programming community

Australian Research Data Commons

NCRIS National Research Infrastructure for Australia An Australian Government Initiative

ARDC is enabled by NCRIS

Troubleshooting user training and engagement challenges

Lightning talk #3

Patrick Taylor
AWS Cloud Application Specialist, RMIT

Patrick is a cloud application specialist at the RMIT AWS Cloud Supercomputing (RACE) Hub, working with researchers to help them achieve research excellence in the cloud.

Australian Research Data Commons
NCRIS National Research Infrastructure for Australia An Australian Government Initiative
ARDC is enabled by NCRIS

Troubleshooting user training and engagement challenges

Lightning talk #3

Aishwarya Krishnan

Solutions Architect,
Amazon Web Services
(AWS)

Aishwarya is a Solutions Architect at Amazon Web Services. She supports Healthcare Organisations and works closely with researchers to understand their unique requirements and provides guidance on leveraging AWS services to implement secure, scalable, and cost-effective cloud solutions. Before joining AWS, she was a Microbiologist, and made a career transition to tech driven by a passion for leveraging technology to improve healthcare outcomes.

Australian Research Data Commons

NCRIS National Research Infrastructure for Australia An Australian Government Initiative

ARDC is enabled by NCRIS

Empowering Researchers: RACE and our Collaboration with AWS

Patrick Taylor
Cloud Application Specialist
RACE

contact: race@rmit.edu.au
Aishwarya Krishnan
Solutions Architect
AWS

What's next...

aws

aarnet
Australia's Academic and Research Network

RMIT UNIVERSITY

RACE (RMIT AWS Cloud Supercomputing)

Vision

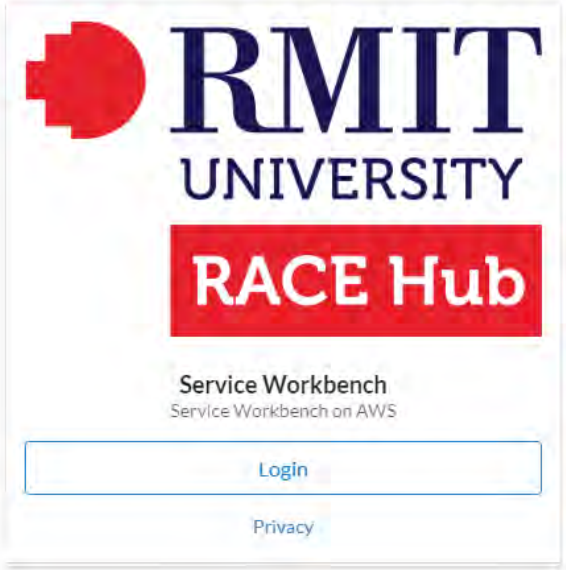
To **empower research and teaching excellence** through cutting-edge cloud infrastructure.

Mission

To provide services, **training** and expert support to enable RMIT staff, students and industry partners **to maximise the scientific return** from the existing data and computing infrastructure.

RACE Landscape

SWB



AWS ParallelCluster

Amazon Bedrock

Amazon SageMaker

aws DEEPRACER LEAGUE

Amazon Braket

Native AWS Access

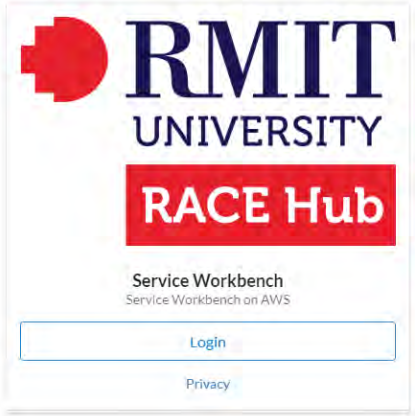
AWS Management Console

Everything you need to access and manage the AWS Cloud — in one web interface

Research Complexity

RACE Landscape

SW B



Native AWS Access

RACE Team

Training

Consulting

Expert Service

Research Complexity

Training to Maximise the Scientific Return

Engagement Success:

- **Embedded onboarding training**
- **Tailored workshops**
- **Hands-on workshops**
- **Building a network of experts**
- **Feedback**

RACE Platform Training

- Onboarding workshops
- Tutorial videos
- Written guides

Research Domain Deep Dives

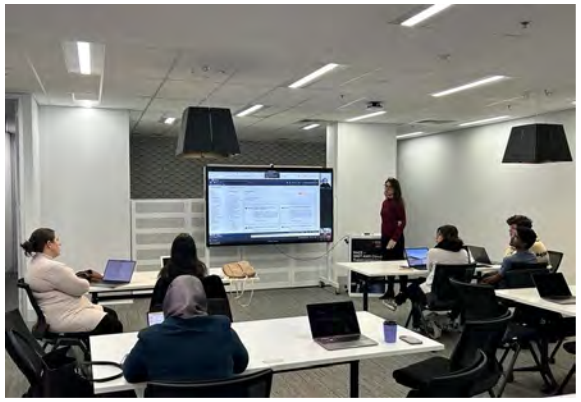
- RACE x AWS training, workshops and immersion days
- Research team support sessions

Coursework

- AWS guest speakers
- Workshops

Training to Maximise the Scientific Return

DeepRacer



HDR Course Workshop

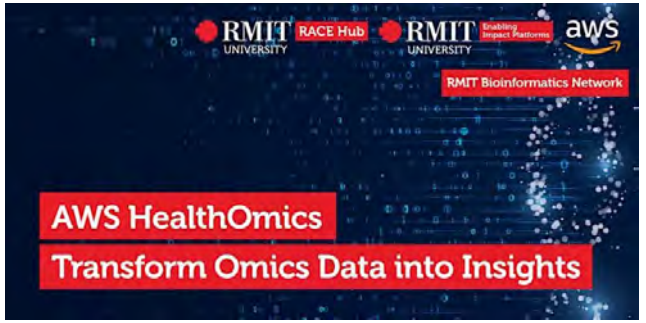
NCI Training

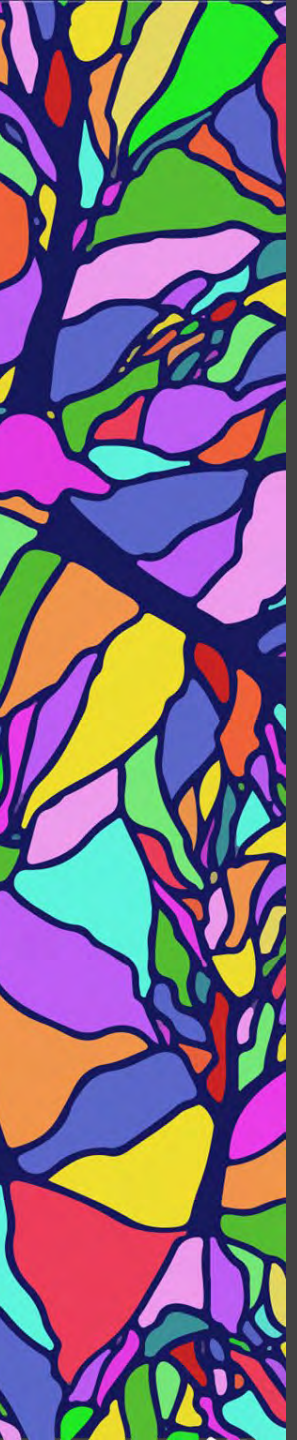
Build a GenAI Chatbot

A No-Code Machine Learning Workshop

SCT x AWS Serverless

Geonomics Workshop





Thank you

RACE & AWS Team

Patrick Taylor
race@rmit.edu.au

RMIT UNIVERSITY RACE Hub

Troubleshooting user training and engagement challenges

Table and breakout groups

Miro Board - Questions & Discussion in groups (40 minutes)

1. What does successful skills training and user engagement look like? How do we get there? Are there barriers?
2. Besides training programs, what other approaches could you suggest that both build community and upskill users of infrastructure?
3. What creative approaches can we use to tailor infrastructure training to different user experience levels?
4. For other comments or ideas - Use section 7 of the Miro board https://miro.com/app/board/uXjVKLbul04=/

Australian Research Data Commons
NCRIS National Research Infrastructure for Australia An Australian Government Initiative
ARDC is enabled by NCRIS

Troubleshooting user training and engagement challenges

Breakout highlights

1. Are there any ideas or solutions you think are innovative and/or would like to put into practice at your organisation?
2. Do you know of any current initiatives or programs that use best or innovative practices that you'd like to highlight?

Australian Research Data Commons
NCRIS National Research Infrastructure for Australia An Australian Government Initiative
ARDC is enabled by NCRIS

Big lottery win!
Blue sky thinking - Menti activity

You now have an unlimited budget.

What engagement and training activities would you propose for:

- Your digital research infrastructure staff?
- Your digital research infrastructure users (researchers)?

Australian Research Data Commons
NCRIS National Research Infrastructure for Australia An Australian Government Initiative
ARDC is enabled by NCRIS

Wrap up

- Summary / Wrap-up
- Feedback Forms
- *QR code for welcome kit here*

Use the QR code to access your Welcome Pack

Australian Research Data Commons
NCRIS National Research Infrastructure for Australia An Australian Government Initiative
ARDC is enabled by NCRIS

INTERNATIONAL DATA WEEK 2025

Trusted research data driving innovation and a better world

Australian Research Data Commons

NCRIS National Research Infrastructure for Australia
An Australian Government Initiative

The ARDC is enabled by NCRIS

COMMITTEE ON DATA CODATA INTERNATIONAL SCIENCE COUNCIL

RDA RESEARCH DATA ALLIANCE

WORLD DATA SYSTEM

Subscribe to the **ARDC CONNECT** newsletter

THANK YOU

ardc.edu.au

contact@ardc.edu.au

@ARDC_AU

Australian-Research-Data-Commons

ARDC_AU

Australian Research Data Commons

NCRIS National Research Infrastructure for Australia An Australian Government Initiative

ARDC is enabled by NCRIS