# Planet Research Data Commons: Vision and Recent Developments

**18 October 2023**

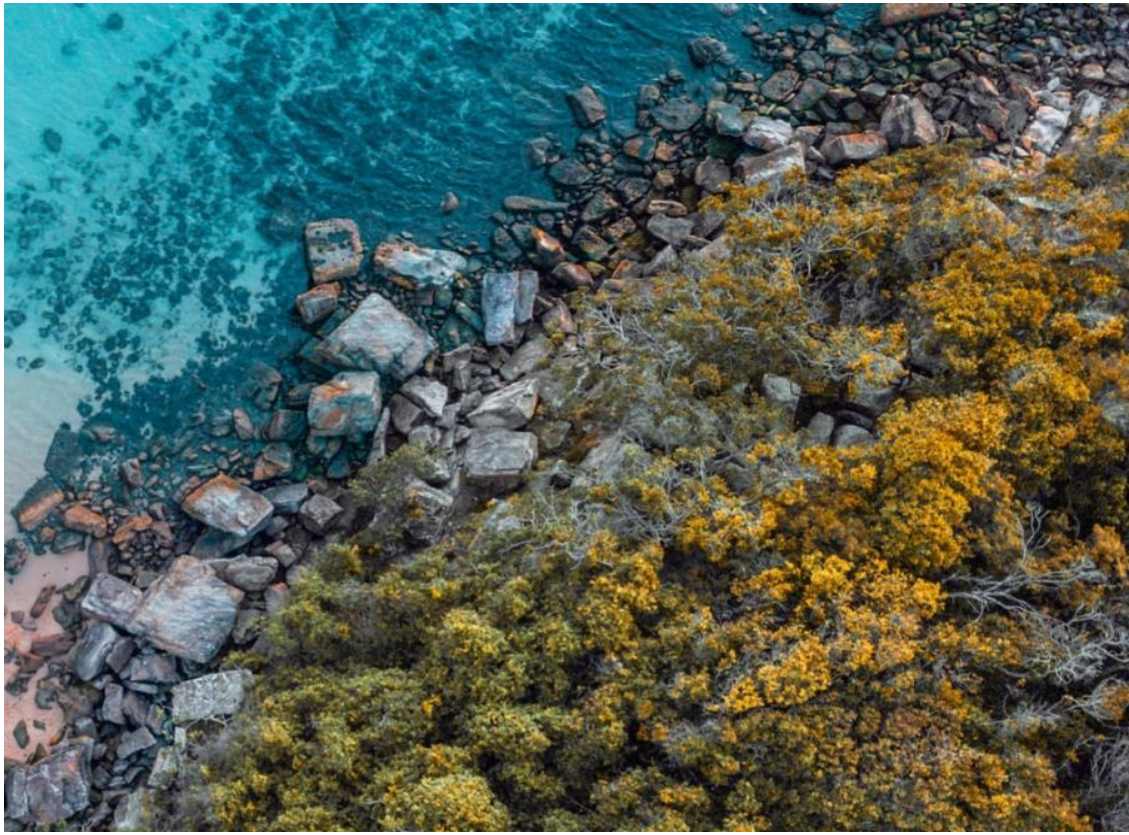
**PRESENTED BY**

Muhammad Ali
Kerry Levett
Hamish Holewa

*Image — themorninglory - 391783296 / STOCK.ADOBE.COM*

# 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.



# **Content and key takeaways from this presentation** (15 mins)

1. Overview of Planet RDC
    a. Data challenges and outcomes from national consultations
    b. Four focus areas
    c. What we are going to do? Activities
2. Most recent developments and success stories
    a. Use cases
3. 5-year vision
    a. Impact and outcome


# 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.

# Outcomes of National Consultations

## Data challenges

### Reliable and Trusted Data Supply Chains

Data **quality** is important for **trust** in data sources

**Government(s)** and **industry** are significant generators of data although difficult to access

Data infrastructure is needed for **continuous** environmental monitoring

**Seamless** access across NCRIS, government, industry

### Curated, Integrated and FAIR Datasets and Services

Data integrated **within and between** Planet research domains

**Geospatial/temporal** data are fundamental but disparate

Data need to be ready for **new methods** (AI/ML)

Need **strategic** approach to identify integration needs and mature processes

### Modelling and Decision Support Infrastructure

Common, **reusable tools and platforms** (repeatable reference architectures)

Allow researchers to utilise and build upon **previous models**

Enable research to **decision making**

Data needs to be available to **all cloud platforms**

FAIR, CARE and Indigenous Knowledge Principles

Trusted Data and Information Supply Chains

Integrated FAIR Datasets and Services

Modelling, Analytics and Decision Support Infrastructure

Indigenous Knowledge Management, FAIR and Skills

# 4 Focus Areas and Integrated Activities

- Address national priorities and challenges
- Intersection across earth and environmental disciplines to enable multi-transdisciplinary research
- Increase interoperability and integration between domains
- Enhancing collaboration and access between research, government and industry
- Enduring data infrastructure with a codesign methodology

ARDC Australian Research Data Commons

NCRIS National Research Infrastructure for Australia An Australian Government Initiative

ARDC is enabled by NCRIS

# Recent developments and success stories

1. Program description released June 2023

2. Trusted data supply chains
   a) SAFE guide published
   b) Partnership with WABSI & DCCEEW / Pilbara
   c) Implementation: Diverse uses cases and regional exemplars- e.g. Pilbara region

3. Planet Infrastructure → MADSI
   a) Developing a reference architecture for Planet RDC digital infrastructure
   b) Planet RDC Ecological and Biodiversity Workshop Report
   c) Opportunity canvas—> to document research community needs

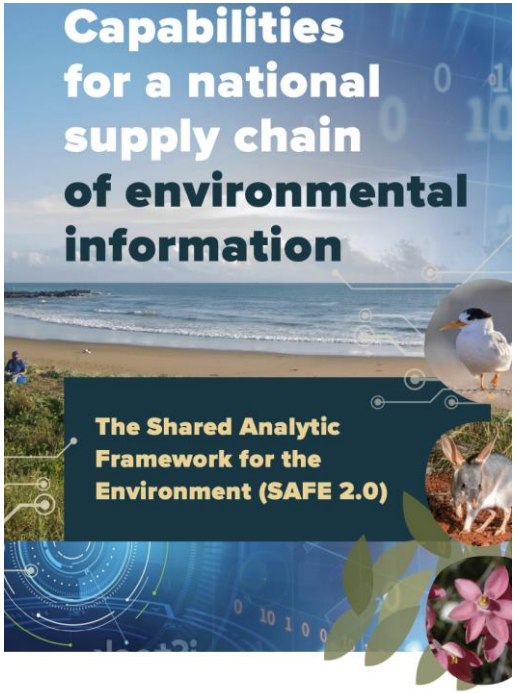
4. Integrated Data → Federated Data Collections with NESP
   a) On the horizon - data spaces/ data lake

5. Indigenous Knowledge - Consultation with NEESFF
   a) Drafted a position paper and recommendations


## Planet RDC Program Description released June 2023

This program description is important because it outlines:

- the drivers behind establishing the Planet RDC and the national consultation process.
- It describes the vision, the high-level program strategy, and activities that the Planet RDC will undertake during its establishment.
- The program description will be used to inform collaborations, programs and projects undertaken by the Planet RDC.




https://ardc.edu.au/resource/shared-analytic-framework-for-the-environment-safe-2-0/

## Trusted data supply chains
## SAFE guide published

"SAFE depicts the capabilities – the building blocks – which work together across the environmental information and analytic supply chain to provide input into … **environmental assessments** … to improve confidence in decisions made."

"SAFE provides a common framework and vocabulary" to show how the different components of the environmental data work together

# MADSI
# Digital Infrastructure for ecology and biodiversity modelling workshop

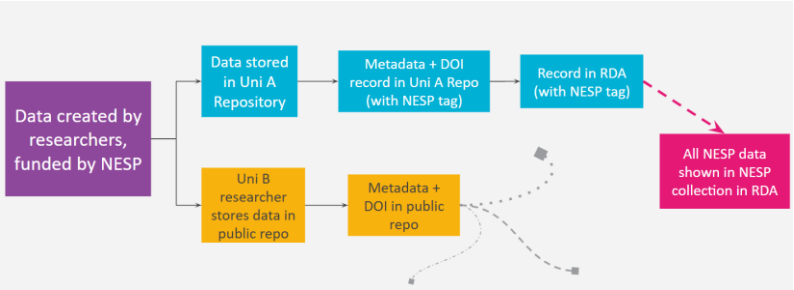

*Planet Research Data Commons workshop on research infrastructure for ecological and biodiversity modelling.*

The ARDC is investigating what digital research infrastructure is needed to support researchers in model development and use.

Objectives of the workshop:

1. Identify high-priority/national priority trusted ecological models and associated input data that could be made available to a broader community of users.

2. Contribute to the development of Planet RDC's understanding of modelling, analytics and decision support infrastructure




*NESP DATA PUBLISHING AND LINKING SCENARIOS*

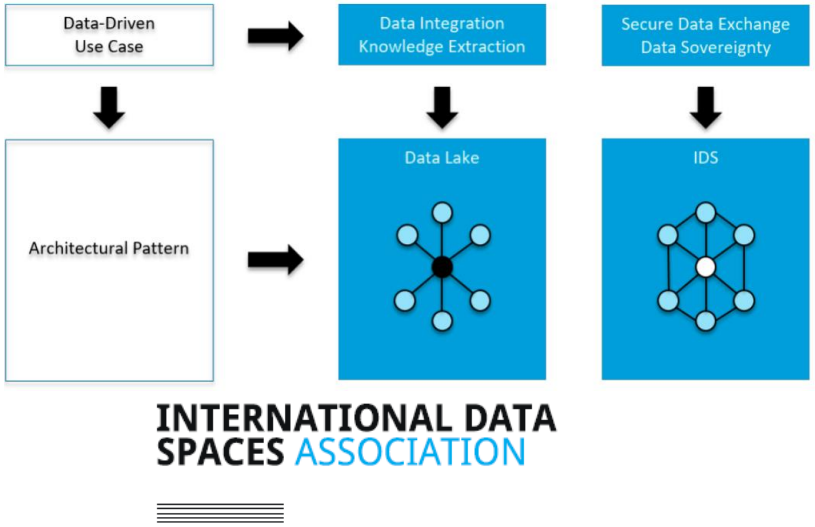
## Engagement with National Environmental Science Program NESP (Australia's largest long term environmental research program)

Objective of this program:

1. Develop the capability to systematically create National Federated Collections for partner organisations or specific disciplines, using Research Data Australia.

In terms of recent developments:

- Planet RDC drafted joint proposal with National Environmental Science Program to develop a NESP Lens.




*Architectural patterns for data exchange and data sharing using data lakehouse and Data space (IDSA)*

## On the horizon - Data Spaces

Large amounts of environmental data remain hidden from view due to security concerns and legal obligations.

1. Data spaces facilitate a decentralized and secure data exchange across organizations and industries that come together to pool, access, process, and share data.
2. Data spaces will change the behaviour, culture, and etiquette when it comes to storing, sharing and exchanging data.

In terms of recent developments:

- Planet RDC has drafted the first version report on Environmental data spaces. Although the Data space concept has gained popularity, its use in the earth and environmental area is still in its early stage.

# Planet RDC vision for 5 years

Australian Research Data Commons

NCRIS National Research Infrastructure for Australia
An Australian Government Initiative

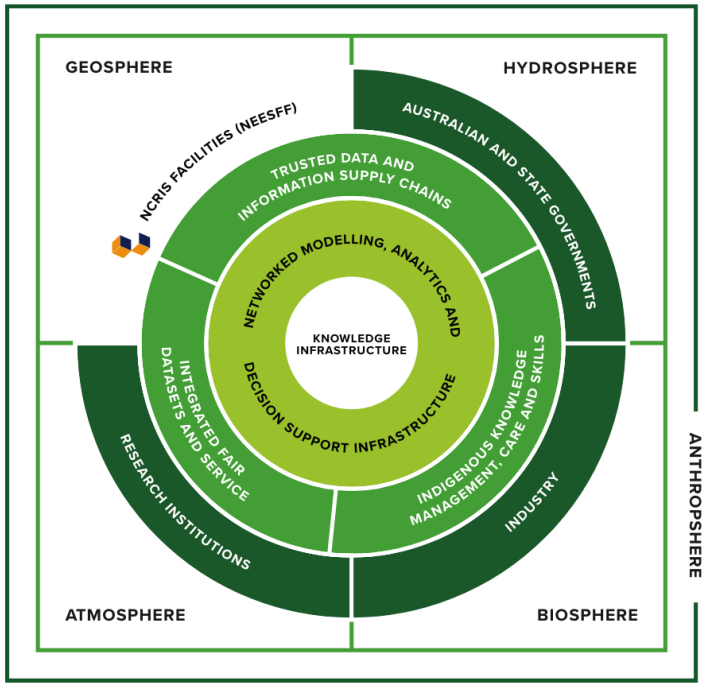
ARDC is enabled by NCRIS

# Planet RDC - Working Collaboratively Towards a Networked Earth & Environmental Infrastructure

The Planet RDC envisages a future where research, government and industry can seamlessly access, visualise, analyse and model trusted data across earth and environmental domains.

**A joined up approach**

a. Network all environmental research facilities, government and research institutions;
b. Allow users to analyse and predict environmental change across all Planet related spheres.




*Planet RDC is working collaboratively towards a networked earth & environmental infrastructure*

# Vision and Plan

The Planet RDC will continue to work on:

1. the four data challenge areas;
2. interconnected and enduring digital infrastructure;
3. strong research translation pathways between research, government and industry; and
4. use cases based on the data infrastructure needs in selected regions of national interest and priority.

# Impact of Planet RDC

The expected impacts of the Planet RDC include:

- faster research through access to datasets, models, workflows and analytic tools
- improved evidence-based environmental and earth science decision making by policy makers and managers
- enhanced opportunities for cross-sector and multidisciplinary research nationally and internationally.

**PLANET**
RESEARCH DATA COMMONS

# 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

ARDC is enabled by NCRIS

# Outcomes

The Planet RDC will deliver national data assets and digital capabilities to support environmental and earth science research and evidence-based decision making. This will include:

- collaborative analytics and decision support platforms
- infrastructure to enable trusted models to be shared, reused and networked to solve complex multidisciplinary challenges
- exemplar trusted data and information supply chains for selected priority regions
- integrated FAIR data assets to enable multidisciplinary research
- skills, governance, policy, technology, collaboration and participation frameworks to enable success.