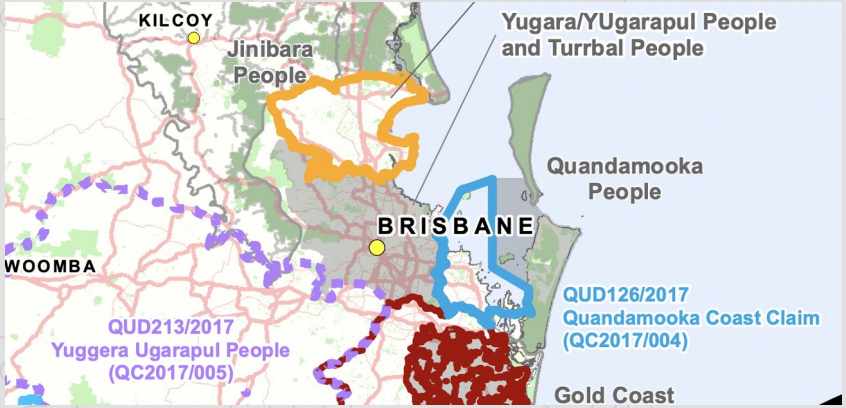
2025 eResearch Australasia Conference

# Towards a Coherent & Coordinated Spectrum of Data Stewardship & Sharing Services

**Session Organizers**
Reyna Broadhurst, World Data System
Helen Glaves, British Geological Survey

# Acknowledgement of Country

We acknowledge the traditional owners of the land on which we meet and pay our respects to their Elders past, present and emerging.




Credit: "Map reproduced with the kind permission of the National Native Title Tribunal".

Turrbal Tribe - https://www.turrbal.com.au/

Yuggera Ugarapul Peoples - https://the-yup.com/

CARE Principles for Indigenous Data Governance

# Agenda

01

## Introduction

Introduce concepts and a brief overview of the World Data System

02

## Lightning Talks

Relevant initiatives and networks: commonalities and coordination

03

## Common Challenges

Mentimeter for demographics and input on challenges

04

## Working Together

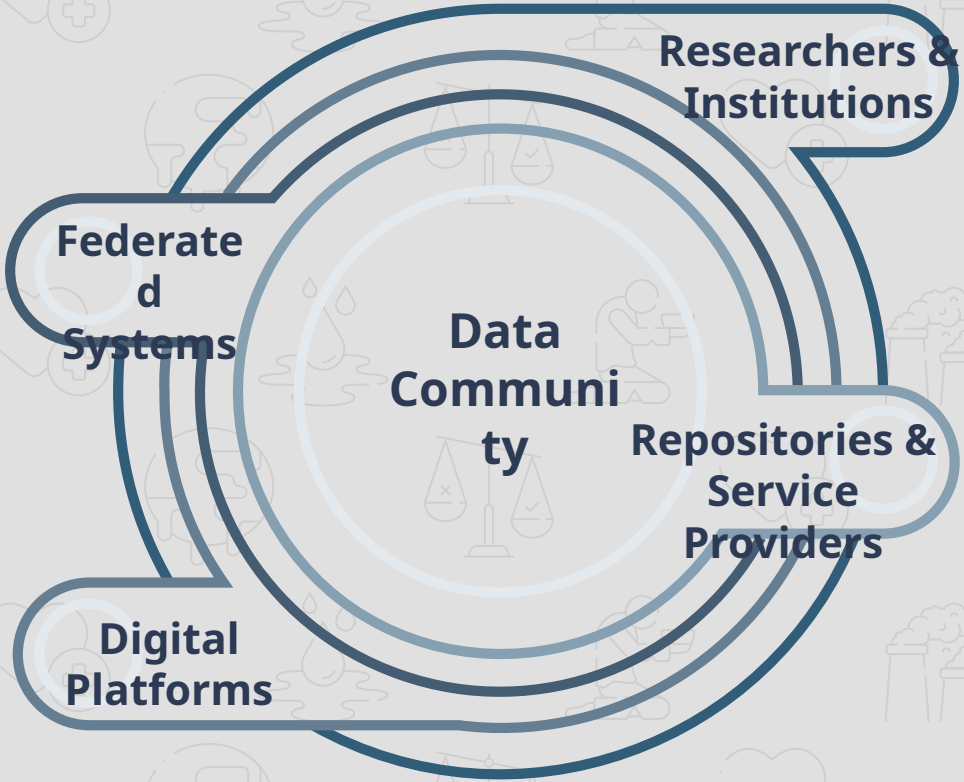
How do we work together to achieve tangible outcomes?

# Coordinated Data Stewardship & Sharing

Many initiatives and networks share some common goals with respect to data curation, sharing, archiving and preservation.

Initiatives take many forms with numerous points of intersection and overlap.

Coordinating organizations and working groups provide cohesiveness and move the entire community forward,




**Coordinating Organizations & Working Groups**

# Fostering community & collaborati

In an increasingly challenging global landscape for scientific research data stewardship, there is a growing need to **work collaboratively to address common goals & maximize potential benefits** of pooling our shared global resources.

UN Decades - https://www.un.org/en/observances/international-decades

UN Years - https://www.un.org/en/observances/international-years

# World Data System

**Mission:** to enhance the capabilities, impact, and sustainability of our member data repositories and data services by:

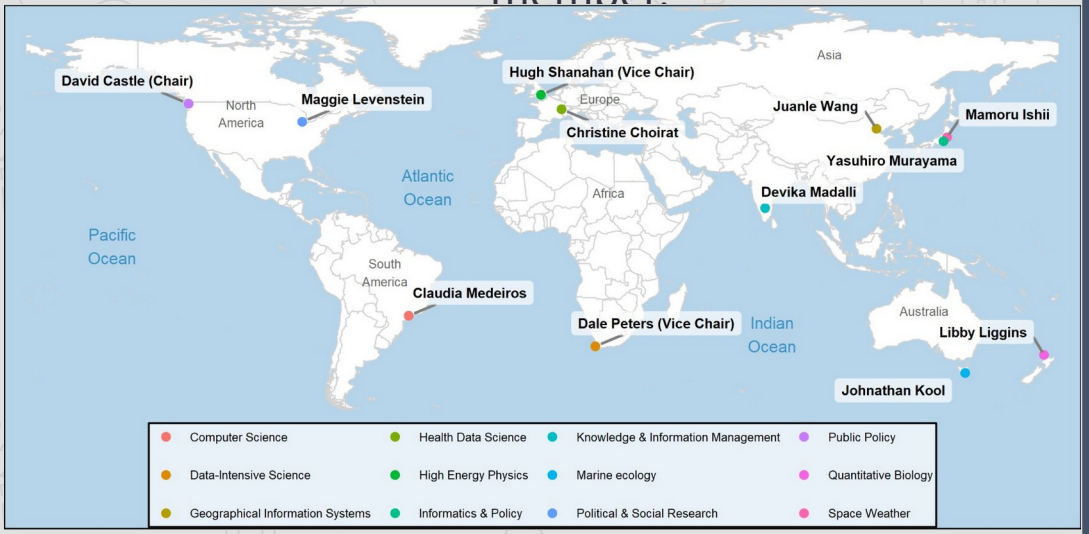
**Creating** trusted communities of scientific data repositories

**Strengthening** the scientific enterprise throughout the entire lifecycle of data and all related components creating first-class data that feeds first-class research output

**Advocating** for accessible data and transparent and reproducible science

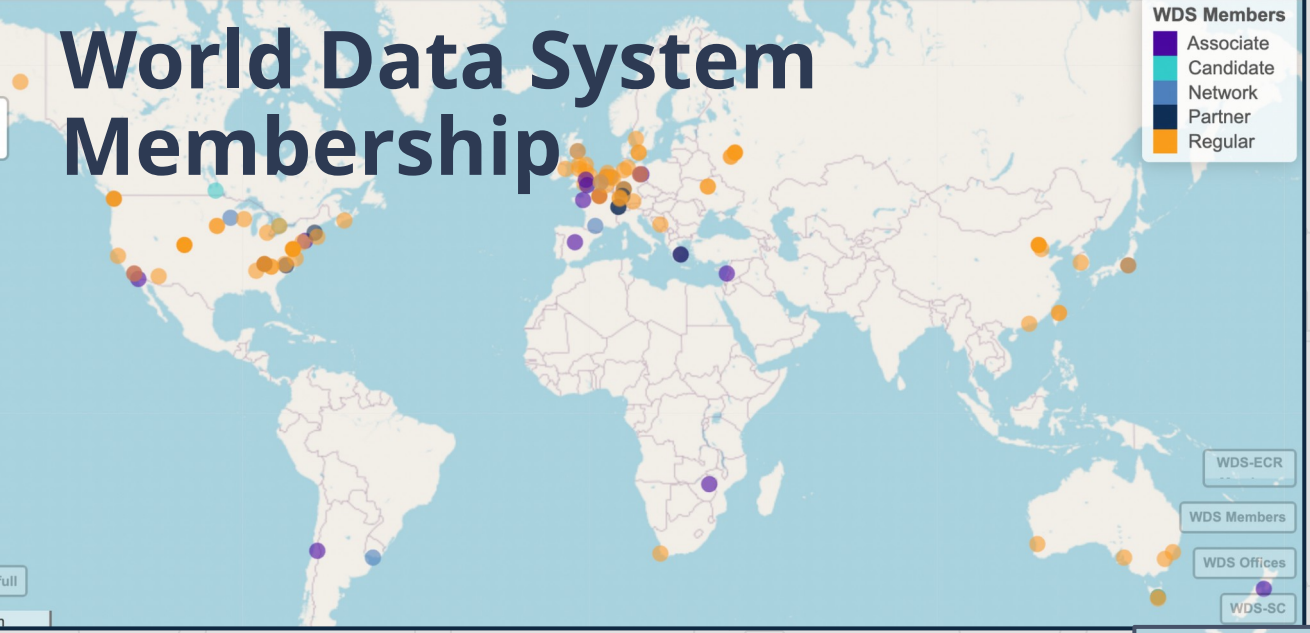
International Science Council
The global voice for science

WDS was initiated by the the predecessor of ISC (ICSU) in 2008, and continues that relationship as an ISC member.

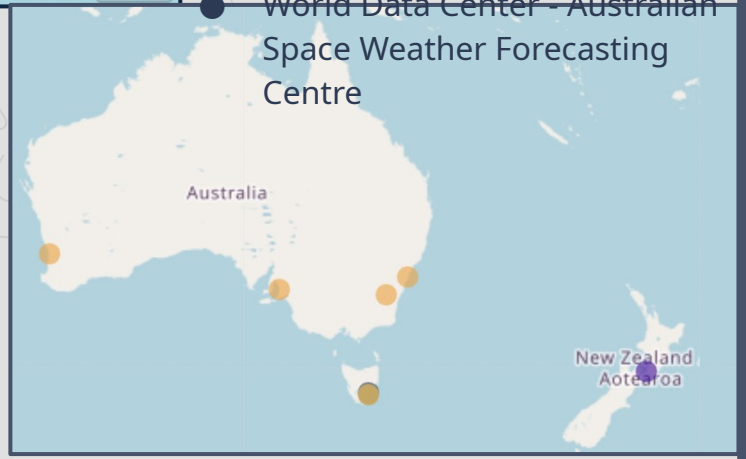



**Scientific Committee**

# World Data System Membership




- AuScope EarthBank Platform
- Australian Antarctic Data Centre
- Australian Data Archive
- Australian Ocean Data Network
- NCMI Information and Data Centre
- Pacific and Regional Archive for Digital Sources in Endangered Cultures
- Royal Society Te Apārangi
- Southern Ocean Observing System
- ~~World Data Center - Australian~~ Space Weather Forecasting Centre




Become a member:
https://worlddatasystem.org/members/application_membership/

# Lighting Talks

01

## Mark Rehbein

Australian Ocean Data Network

02

## Angus Nixon

AuScope Geochemistry Network

03

## Kelsey Druken

ACCESS-NRI

04

## Helen Glaves

British Geological Survey

01

# Australian Ocean Data Network

Mark Rehbein
Director, Australian Ocean Data Network
Integrated Marine Observing System

# Australian Ocean Data Network (AODN)

National hub for Australian marine and coastal data

A certified repository for IMOS data - a national scale marine observing program (55 facilities from 10 Institutions)

Also a federated network of Partners, underpinned by

- Metadata standards
- Spatial data standards

Delivers open and freely accessible data via:

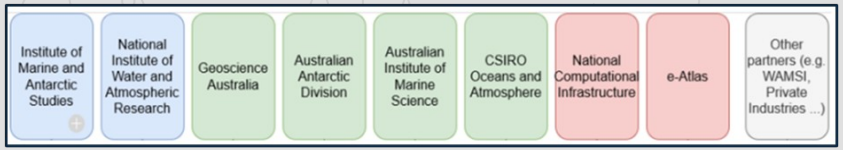
- National marine data catalogue
- Interactive portal (human friendly)
- Data web services (modelling friendly)
- Data web services (modelling friendly)
- Cloud optimised formats (data science friendly)

Supporting various technical and non-technical use-cases & end-user needs

Connected into other national and international data systems




AODN Federated Partners



# Successes

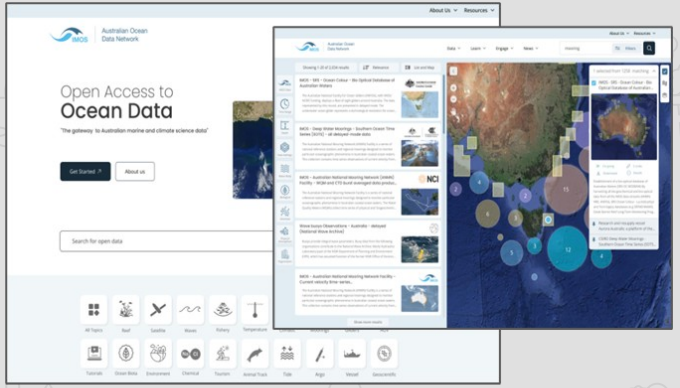
Collaboration – Strong partnerships across IMOS facilities, and federated Partners (government, and research organisations)

A comprehensive national catalogue of marine data

- Federated network

Standards and best practices (interoperability)

- OceanBestPractices
- NetCDF CF conventions
- OGC spatial data standards
- Metadata (ISO-19115)




New IMOS AODN Portal

High quality, well curated IMOS data, supported by Standard Operating Procedures (SOP's)

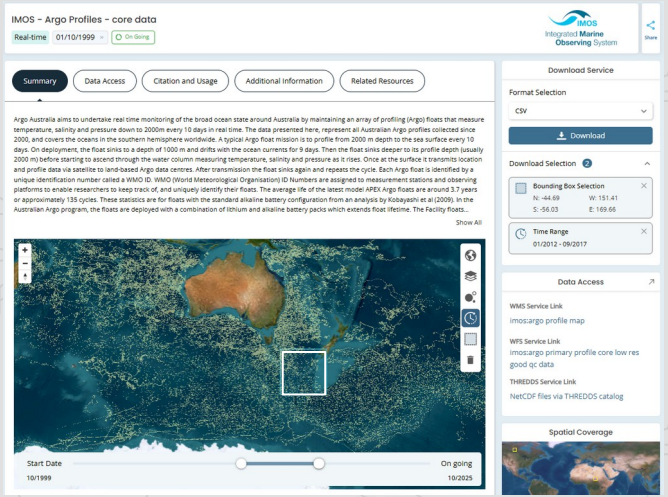
Free and open access to IMOS data

Repository certifications

- CoreTrustSeal
- WorldDataSystem
- IODE Accredited National Oceanographic Data Centre (NODC)

# Challenges & Gaps

- Dealing with restricted data
  - Indigenous
  - Industry - driving open data in industry
- CARE in ISO-19115 (metadata)
- DOI's (or not) for dynamic subsetting and repeatability
- Data versioning policy – when, when not, how




- Tracking impact beyond website metrics
- Emerging cloud optimised standards/best practices (Zarr/Parquet)
- Federated network
  - Cyber posture
  - Barriers to entry
- Talent retention / recruitment

02

# AuScope Geochemistry Network

Angus Nixon
Postdoctoral Researcher
AuScope Geochemistry Network

# AuScope Geochemistry Network

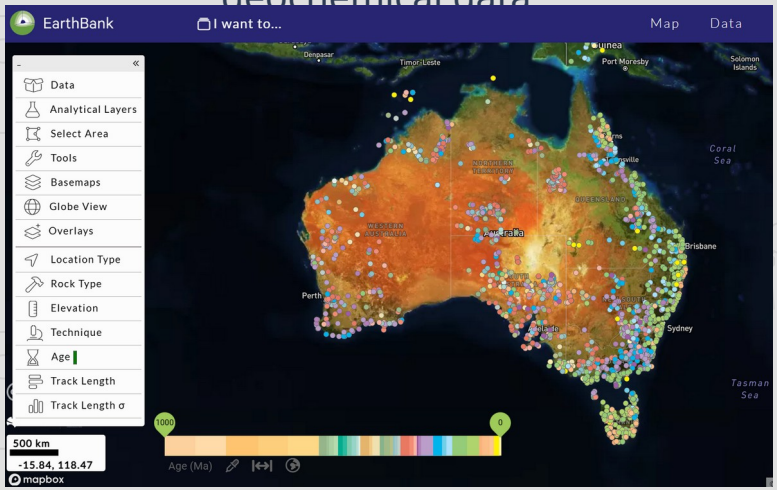
Collaboration of geoscientists across Australian universities

Enabled by funding from NCRIS and AuScope

Connecting discipline experts for standard building and enabling collaborative research

Particular focus on linking with laboratories, aiming to cover sample/data progress from generation to publication and reuse

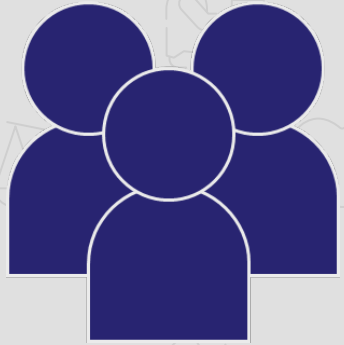
EarthBank is a domain platform designed to store, publish and visualise geochemical data



# Successes

Working with community to establish **expert working groups** for geochemical data and applications

- Low-temperature thermochronology (fission track + (U-Th)/He)
- Inorganic geochemistry
- U/Pb geochronology
- Ar/Ar geochronology
- Beta-decay geochronology
- Radiogenic isotopes
- Instrumental metadata (ICP-MS, EPMA)



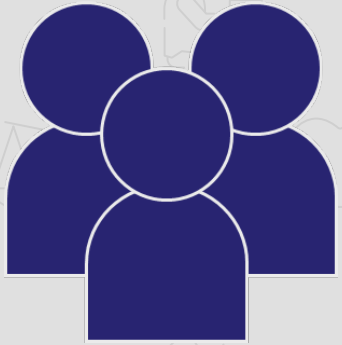
Working groups led by AGN member who supplies data science experience and time

Data model outputs coded into EarthBank, validated with key dataset(s)

# Successes

Working with community to establish **expert working groups** for geochemical data and applications

- Low-temperature thermochronology (fission track + (U-Th)/He)



Vol. 46 — N° 2 06 22 p. 245–259

GEOSTANDARDS and GEOANALYTICAL RESEARCH

**AusGeochem: An Open Platform for Geochemical Data Preservation, Dissemination and Synthesis**

Samuel C. **Boone** (1)*, Hayden **Dalton** (1), Alexander **Prent** (2), Fabian **Kohlmann** (3), Moritz **Theile** (3), Yoann **Gréau** (4), Guillaume **Florin** (4), Wayne **Noble** (3), Sally-Ann **Hodgekiss** (4), Bryant **Ware** (2), David **Phillips** (1), Barry **Kohn** (1), Suzanne **O'Reilly** (4), Andrew **Gleadow** (1), Brent **McInnes** (2) and Tim **Rawling** (5)

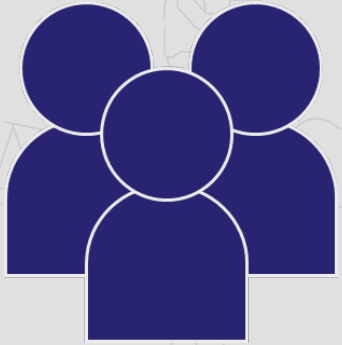
ng groups led by AGN
er who supplies data
e experience and time

nodel outputs coded
rthBank, validated with
key dataset(s)

# Successes

Working with community to establish **expert working groups** for geochemical data and applications

- Low-temperature thermochronology (fission track + (U-Th)/He)



Vol. 46 — N° 2 06 22 p. 245–259

GEOSTANDARDS and GEOANALYTICAL

**AusGeochem: An Open Platform for G Preservation, Dissemination and Synth**

Samuel C. **Boone** (1)*, Hayden **Dalton** (1), Alexander Moritz **Theile** (3), Yoann **Gréau** (4), Guillaume **Florin** (4), Sally-Ann **Hodgekiss** (4), Bryant **Ware** (2), David **Phillips** Suzanne **O'Reilly** (4), Andrew **Gleadow** (1), Brent **McInn**

scientific reports

Check for updates

OPEN

**A geospatial platform for the tectonic interpretation of low-temperature thermochronology Big Data**

Samuel C. Boone[1,2], Fabian Kohlmann[3], Wayne Noble[3], Moritz Theile[3], Romain Beucher[4], Barry Kohn[1], Stijn Glorie[2], Martin Danišík[5], Renjie Zhou[6], Malcolm McMillan[1], Angus Nixon[2], Andrew Gleadow[1], Xiaodong Qin[7], Dietmar Müller[7] & Brent McInnes[5]

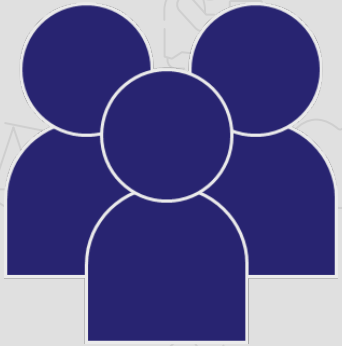
ELSEVIER

Chemical Geology

Available online 13 October 2025, 123092

In Press, Journal Pre-proof ? What's this?

# Volcanoes to vugs: Demonstrating a FAIR geochemistry framework with a diverse application of major and trace element data through the AuScope EarthBank platform



Vol. 46 — N° 2 06 22 p. 245–259

GEOSTANDARDS and GEOANALYTICAL

# AusGeochem: An Open Platform for G Preservation, Dissemination and Synth

Samuel C. **Boone** (1)*, Hayden **Dalton** (1), Alexander
Moritz **Theile** (3), Yoann **Gréau** (4), Guillaume **Florin** (4),
Sally-Ann **Hodgekiss** (4), Bryant **Ware** (2), David **Phillips**
Suzanne **O'Reilly** (4), Andrew **Gleadow** (1), Brent **McInn**

**scientific** reports

Check for updates

OPEN

# A geospatial platform for the tectonic interpretation of low-temperature thermochronology Big Data

Samuel C. Boone[1,2], Fabian Kohlmann[3], Wayne Noble[3], Moritz Theile[3], Romain Beucher[4], Barry Kohn[1], Stijn Glorie[2], Martin Danišík[5], Renjie Zhou[6], Malcolm McMillan[1], Angus Nixon[2], Andrew Gleadow[1], Xiaodong Qin[7], Dietmar Müller[7] & Brent McInnes[5]

# Challenges & Gaps

The vast majority of geochemical data still goes to generalist repositories, appendices or data dumps - if it is available at all

**Community** engagement, participation and willingness is key!

- Reducing time/effort required - automation
- Carrots - citation metrics, research outcomes community and individual benefit
- Sticks - working with journals, funding agencies, international communities

Collaboration between data infrastructures

Building shared foundational resources or crosswalks

Medium- long-term sustainability

- IT support
- Personnel change
- Funding cycles
- etc.

# 03 a’s Climate Simulator (ACCESS-NRI)

Kelsey Druken
Associate Director, Release Management
ACCESS-NRI

# Australia’s Climate Simulator (ACCESS-NRI)

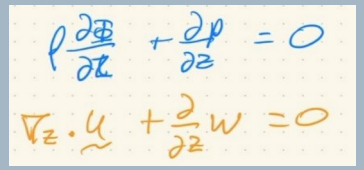
- National research infrastructure funded by the Australian government (NCRIS) to support climate modelling (software and data)
- Provides computer simulations for climate, weather and Earth systems, specifically designed for Australia and the Southern Hemisphere.

# Australia's Climate Simulator (ACCESS-NRI)

- National research infrastructure funded by the Australian government (NCRIS) to support climate modelling (software and data)
- Provides computer simulations for climate, weather and Earth systems, specifically designed for Australia and the Southern Hemisphere.

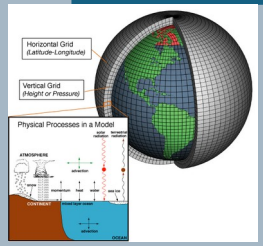
Software and data workflow

Equations based on fundamental physics

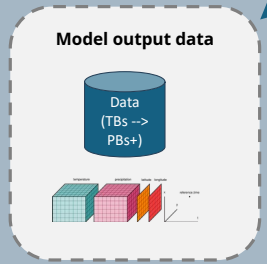


Supercomputer

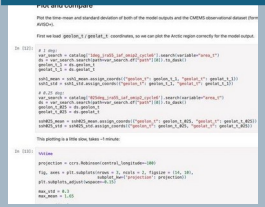

Numerical simulations



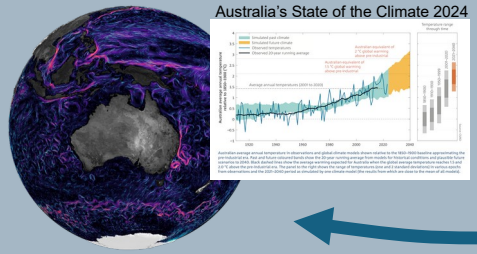

Model output data




Analysis code



Model visualisation & downstream climate/environmental data products

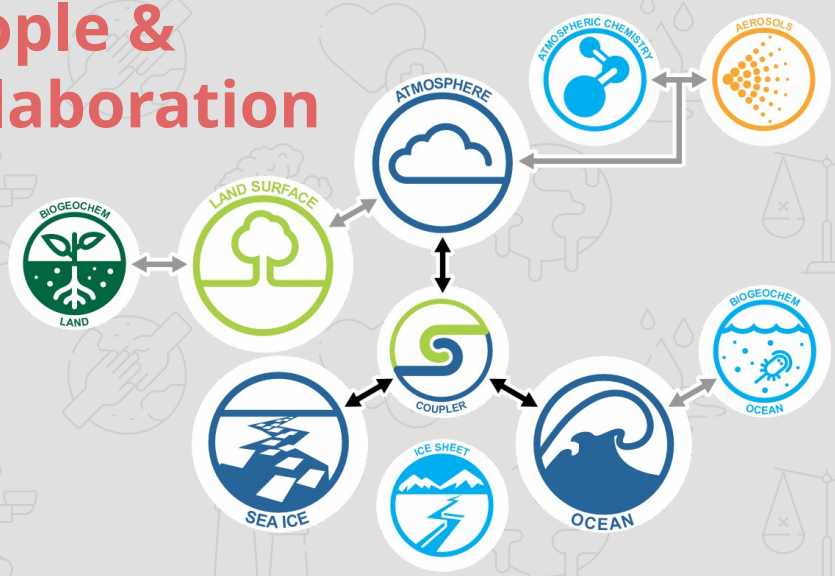


# Australia's Climate Simulator (ACCESS-NRI)

- National research infrastructure funded by the Australian government (NCRIS) to support climate modelling (software and data)
- Provides computer simulations for climate, weather and Earth systems, specifically designed for Australia and the Southern Hemisphere.

People & collaboration

Softwar e




Data

Australian Community Climate and Earth System Simulator (ACCESS)

# Successes

| | |
|---|---|
| **Open science** | |
| **Common platforms** | |
| **Scale** | |
| **International collaboration** | |

# Successes

| | |
|---|---|
| **Open science** | Open licensing for data and software<br>Open development |
| **Common platforms** | |
| **Scale** | |
| **International collaboration** | |

# Successes

| | |
|---|---|
| **Open science** | Open licensing for data and software<br>Open development |
| **Common platforms** | HPC requirements creates common platforms for model developers, data creation, analysis, visualisation, etc. |
| **Scale** | |
| **International collaboration** | |

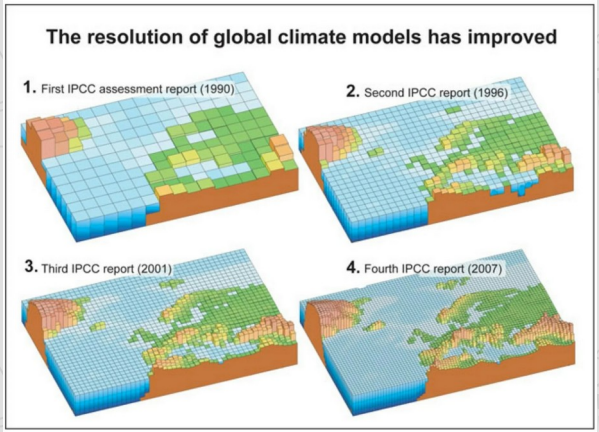
GitHub

ACCESS
National Research Infrastructure

# Successes

| | |
|---|---|
| **Open science** | Open licensing for data and software<br>Open development |
| **Common platforms** | HPC requirements creates common platforms for model developers, data creation, analysis, visualisation, etc. |
| **Scale** | Increase in computation power → higher resolutions, larger data outputs → greater coordination |
| **International collaboration** | |

GitHub

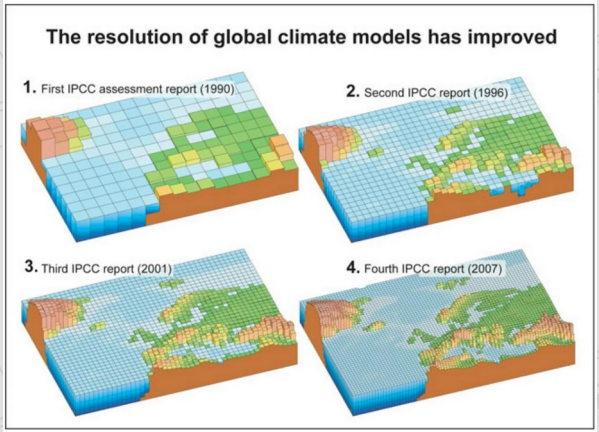



# Successes

| | |
|---|---|
| **Open science** | Open licensing for data and software<br>Open development |
| **Common platforms** | HPC requirements creates common platforms for model developers, data creation, analysis, visualisation, etc. |
| **Scale** | Increase in computation power → higher resolutions, larger data outputs → greater coordination |
| **International collaboration** | Data sharing and global intercomparison projects (i.e., CMIP/ESGF) has driven robust international standards and practices and federated data sharing. |

ACCESS National Research Infrastructure

GitHub

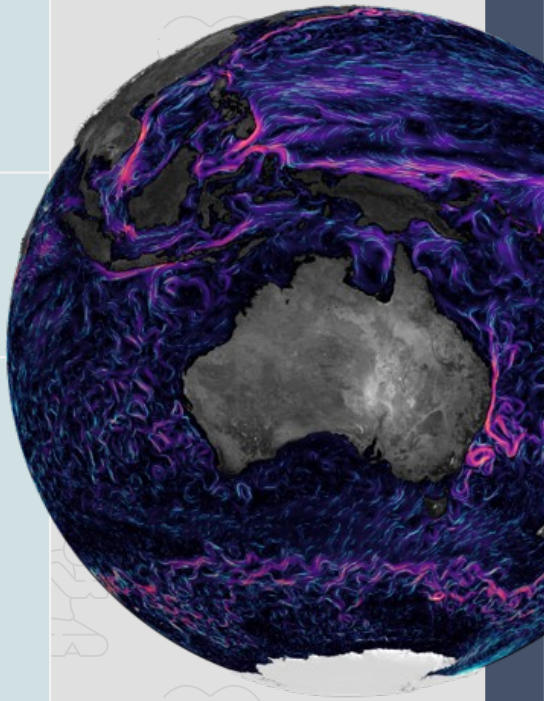




CMIP Coupled Model Intercomparison Project

ESGF Earth System Grid Federation

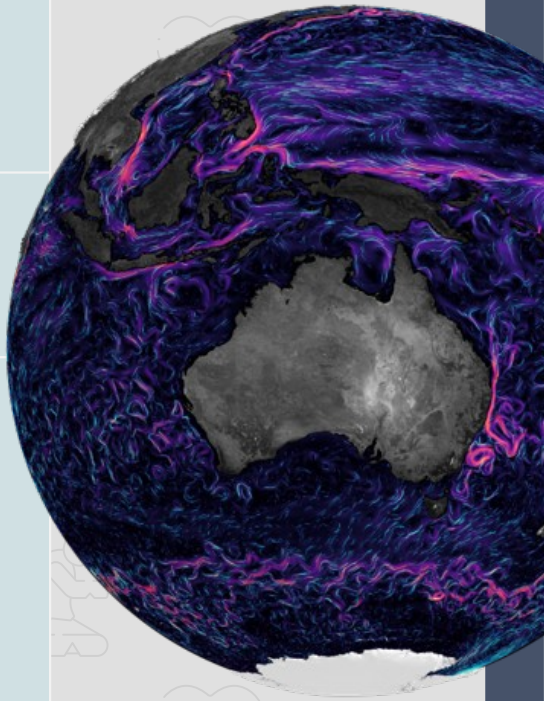
# Challenges & Gaps

| | |
|---|---|
| **Open science** | Not everything is open - data and code bases that are also used for operational weather forecasting are not open, have legal barriers and hurdles for research use.<br><br>Open development has social/people aspects - geospatial, governance, etc. |
| **Common platforms** | |
| **Scale** | |
| **International collaboration** | |



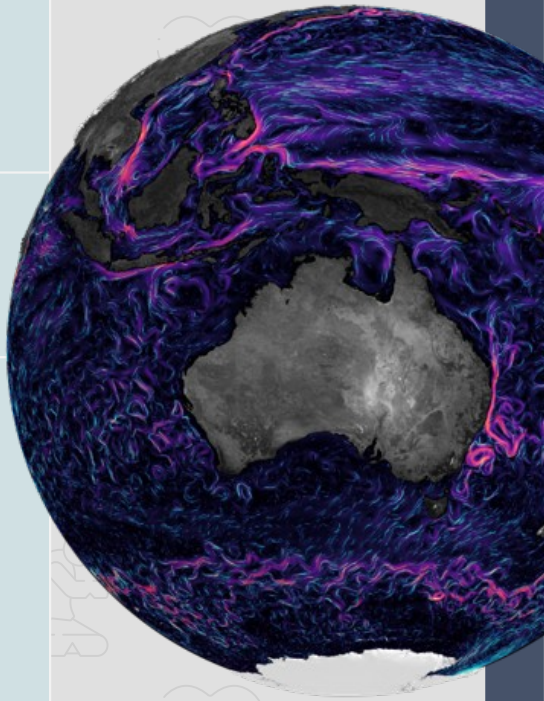
# Challenges & Gaps

| | |
|---|---|
| **Open science** | Not everything is open - data and code bases that are also used for operational weather forecasting are not open, have legal barriers and hurdles for research use.<br><br>Open development has social/people aspects - geospatial, governance, etc. |
| **Common platforms** | Working across HPC and compute/data intensive platforms challenging:<br>● Porting code, moving data not trivial<br>● Platforms themselves have restricted access and require HPC resources |
| **Scale** | |
| **International collaboration** | |



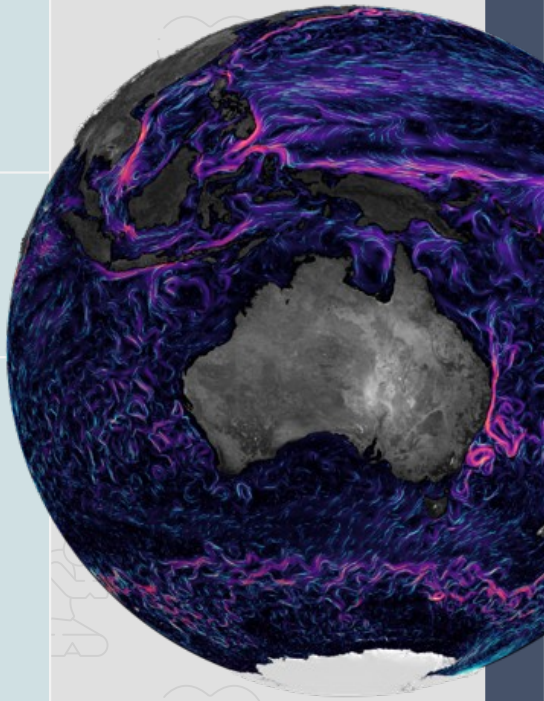
# Challenges & Gaps

| | |
|---|---|
| **Open science** | Not everything is open - data and code bases that are also used for operational weather forecasting are not open, have legal barriers and hurdles for research use.<br><br>Open development has social/people aspects - geospatial, governance, etc. |
| **Common platforms** | Working across HPC and compute/data intensive platforms challenging:<br>● Porting code, moving data not trivial<br>● Platforms themselves have restricted access and require HPC resources |
| **Scale** | Increase in computation power has led to higher resolutions, larger data outputs:<br>● Expensive → high compute/storage costs<br>● Data too large to easily move AND desire to increasingly integrate with data across other domains<br>● **Little to no support for data publication and archival**<br><br>Interoperability with other datasets challenging. |
| **International collaboration** | |



# Challenges & Gaps

| | |
|---|---|
| **Open science** | Not everything is open - data and code bases that are also used for operational weather forecasting are not open, have legal barriers and hurdles for research use.<br><br>Open development has social/people aspects - geospatial, governance, etc. |
| **Common platforms** | Working across HPC and compute/data intensive platforms challenging:<br>• Porting code, moving data not trivial<br>• Platforms themselves have restricted access and require HPC resources |
| **Scale** | Increase in computation power has led to higher resolutions, larger data outputs:<br>• Expensive → high compute/storage costs<br>• Data too large to easily move AND desire to increasingly integrate with data across other domains<br>• **Little to no support for data publication and archival**<br><br>Interoperability with other datasets challenging. |
| **International collaboration** | Data sharing, reuse, and interoperability works very well for the well coordinated large activities, much harder for everything else! |



04

# GEO: Earth Intelligence For A

Helen Glaves
In-Situ Data subgroup (co-chair)

# GEO: International Collaboration on Earth Observations

- What is the Group on Earth Observations (GEO)?
  - International network of member government institutions, partner organizations, and associate commercial entities
  - Working collaboratively to leverage Earth observations, and related data products and services, to improve the sustainable development and management of the environment.

- How does it operate?
  - Nominated representatives serve on the GEO Executive Committee, Programme Board, and within Working Groups.
  - Work Programme activities and Working Groups address specific topics utilizing EO data

# GEO: Data & Knowledge Working Group

Working with community to address data policy, ethics, and governance issues to improve the use of Earth observation data and knowledge by:

- promoting the adoption and implementation of good data management practices including widely adopted practices and principles including FAIR, CARE, TRUST and GEO Data Sharing and Data Management Principles;
- recommending ways to advance interoperability of Earth observation data and associated products and services
- advancing discussions of critical data-related topics with stakeholder communities

# GEO: Tools and Practices Supporting Open Data Sharing

## Best practices

- Common Standards: OGC, ISO, community
- Data Licensing (Law & Policy subgroup guidance document)
- Principles:
  - FAIR, CARE, TRUST, Others
  - GEO Data Sharing & Data Management Principles (DS&DMP)

# GEO: Data Sharing and Data Management Principles

## Data Sharing Principles

- Data is open by default: metadata, data and other research assets are openly shared without charge or restrictions
- Where policies or legislation preclude open sharing, data should be made available with minimal restrictions and at no more cost than that of reproduction and distribution
- All data should be made available with minimal time delay

# GEO: Data Sharing and Data Management Principles

## Data Management Principles:

**Sub-divided into 5 categories**:

1. Discoverability
2. Accessibility
3. Usability
4. Preservation
5. Curation

**Each principle includes:**

- an explanation
- guidance for implementation
- metrics for measuring adherence
- potential resource implications

# GEO: Tools and Practices Supporting Open Data Sharing

**Technologies**

- **GEO Knowledge Hub** : open-source digital repository for authoritative and reproducible open knowledge and data created by the Group on Earth Observations (GEO)

- **G-REQS**: **G**eospatial in-situ **REQ**uirement**S** database tool and standard methodology to collect user requirements for in-situ datasets in the GEO context

- **Libinsitu**: open-source python library to
- transform heterogeneous in-situ time series measurements into binary CF convention compliant NetCDF files

# GEO fosters open sharing of EO data by:

01

**Providing policies and practices** to drive wider adoption of good data stewardship within GEO and beyond

02

**Delivering tools, services and resources** to support open sharing of EO data, identify gaps, and support sustainability of existing observing systems

03

GEO has a unique convening power as an intergovernmental organisation to foster open and equitable access to EO data, products and services to deliver **Earth Intelligence For All!**

# Thanks!

## Any questions?

hmg@bgs.ac.uk

*More information about GEO available here*


Content provided by:

- Joan Maso
- Lionel Menard
- Marie-Francois Voidroit
- Bente Lilja Bye
- Bob Downs
- Paola da Salvo

# Common Challenges

## Go to www.menti.com
## Enter the code: 3904 3877

# Working Together to Achieve Tangible Outcomes

## Practicalities

How do we work effectively across different geographical locations

## Collaboration tools

What tools are required to support collaboration?

## Defining common challenges

What are those common issues that we can solve

# Wrap Up & Next Steps

# Thanks!

## Do you have any questions?

**CREDITS:** This presentation template was created by **Slides go**, and includes icons by **Flaticon**, and infographics & images by **Freepik**