

# Humanities, Arts and Social Sciences Research Data Commons and Indigenous Research Capability Program

BACKGROUND

15 June 22

PRESENTED BY

Jenny Fewster  
HASS RDC Program Manager

# 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

The Australian Research Data Commons (ARDC) enables the Australian research community and industry access to nationally significant, data intensive digital research infrastructure, platforms, skills and collections of high quality data.

As a national research infrastructure provider, the ARDC facilitates partnerships to develop a coherent research environment that enables researchers to find, access, contribute to and effectively use services to maximise research quality and impact.

# ARDC STRATEGY



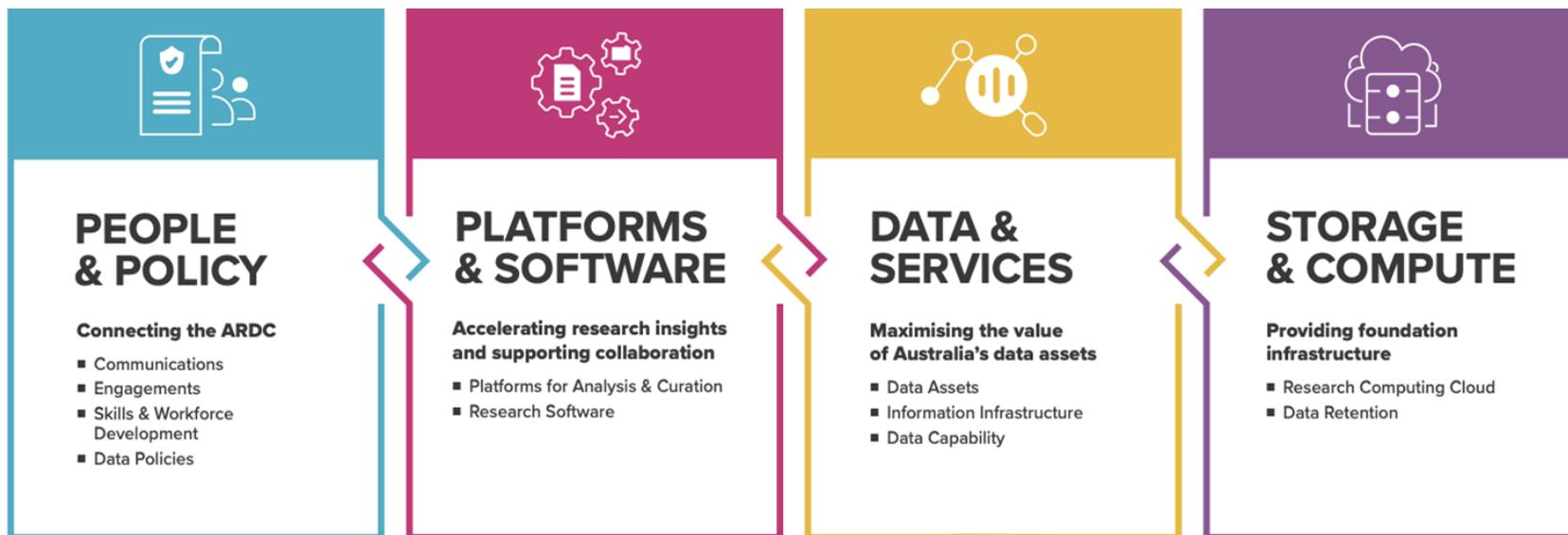
## PURPOSE

To provide Australian researchers with competitive advantage through data.

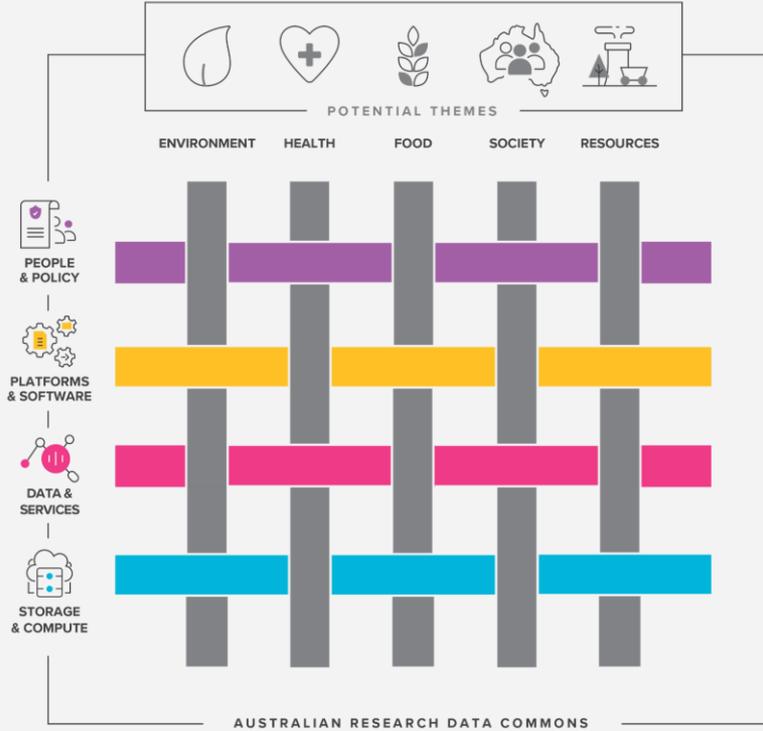


## MISSION

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



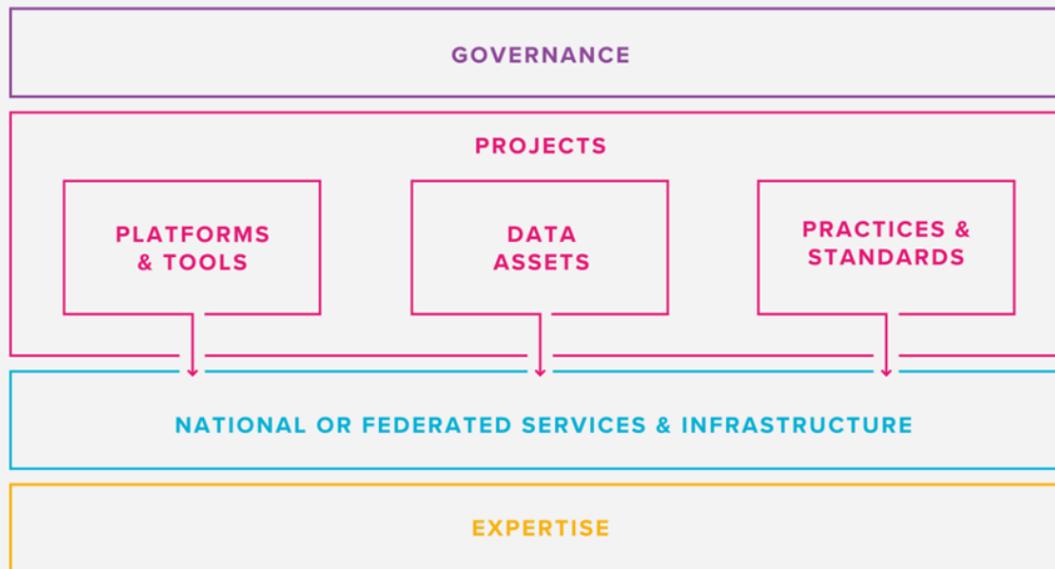
Add your project name



# THEMATIC RESEARCH DATA COMMONS

- HASS
- People
- Planet

# DELIVERY MODEL OF A THEMATIC RDC



# BACKGROUND

## Humanities, Arts and Social Sciences Research Data Commons and Indigenous Research Capability Program investment

- The need for investment was detailed in the 2016 National Research Infrastructure Roadmap
- DESE subsequently commissioned three studies which identified a number of investment ready programs
- Activities earmarked for funding displayed an advanced state of readiness to participate in and benefit from a HASS Research Data Commons
- Funding for the four identified activities was guided by the recommended investment ratios in the DESE studies

# PROGRAM OBJECTIVES

The work undertaken will:

- *“provide easier access to data and analysis methods enabling effective data mining and re-use, maximising the return on previous research investments”* and will,
- *“help institutions share HASS data more freely and cooperatively following FAIR data principles”* [Research Infrastructure Investment Plan, 2020].

It will be guided by the NCRIS Principles and will be...

*“developed on a collaborative, national, non-exclusive basis”* and *“serve the research and innovation system broadly, not just the host/funded institutions”* [National Collaborative Research Infrastructure Strategy 2021 Guidelines, 2021].



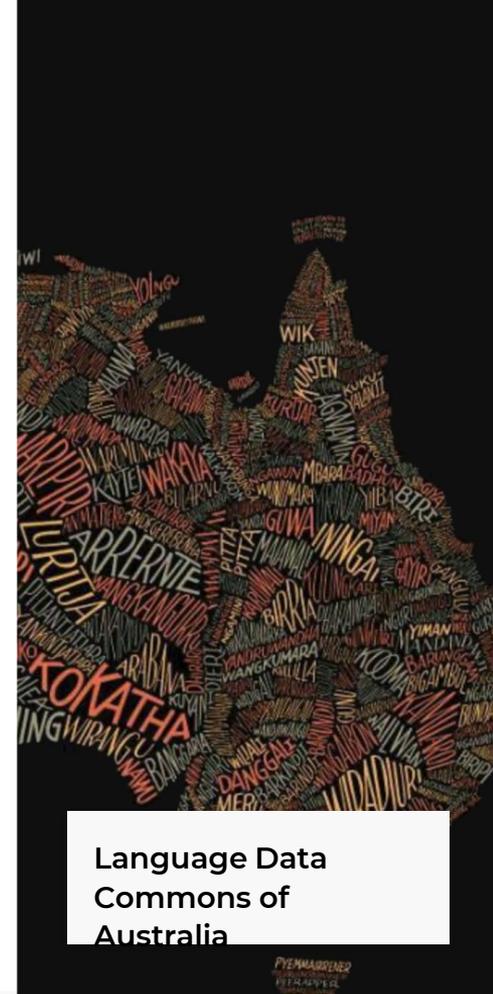
# Language Data Commons of Australia

Project Lead: Prof Michael Haugh (UQ)

ARDC co-investment: \$1,932,996

Project co-investment: \$1,942,750

- Four Streams of Work Packages:
  - Securing Language Data Collections
  - Aggregating Language Data Collections
  - Improving Text Analysis Environments
  - Strategic Partnerships, Engagement and Training



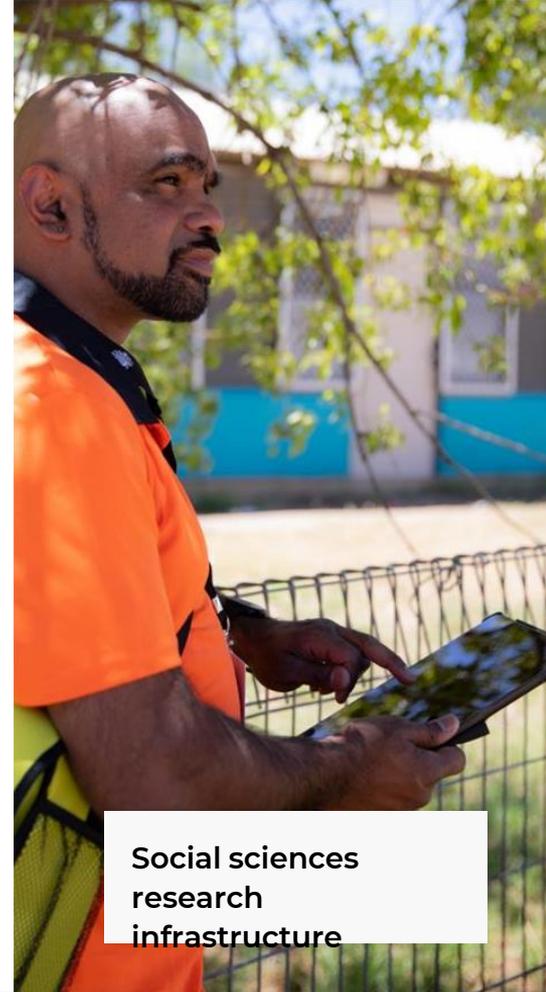
# Integrated Research Infrastructure for Social Sciences

Project Lead: Dr Steven McEachern

ARDC co-investment: \$1,380,711

Project co-investment: \$1,452,000

- Six Work Packages:
  - IRISS Project Coordination
  - VASSSAL – Vocabulary Access Service for Social Science in Australia
  - GeoSocial data integration service
  - IRISS Demonstrator projects
  - SPIRE - Survey Project Integrated Research Environment
  - CARDSS – Curation of Australian Research Data in the Social Sciences



**Social sciences  
research  
infrastructure**

# Improving Indigenous Research Capabilities

Project Lead: Prof Marcia Langton

ARDC co-investment: \$1,610,830

Project co-investment: \$1,829,831

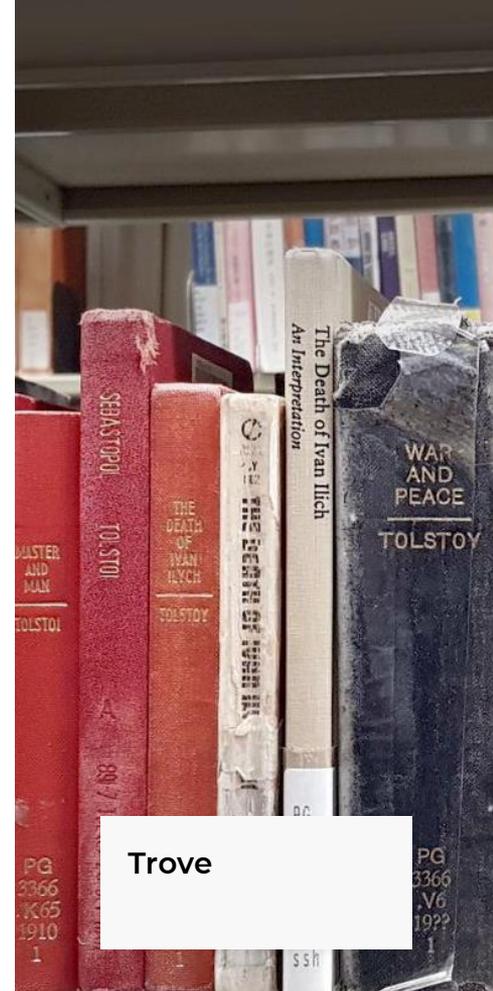
- Three Streams:
  - Social Architecture: Indigenous Data Governance and Sovereignty
  - Technical Architecture: Building the foundations for an Aboriginal and Torres Strait Islander Research Data Commons
  - Core National Indigenous Data Assets: Building an Aboriginal and Torres Strait Islander Spatio-temporal Framework



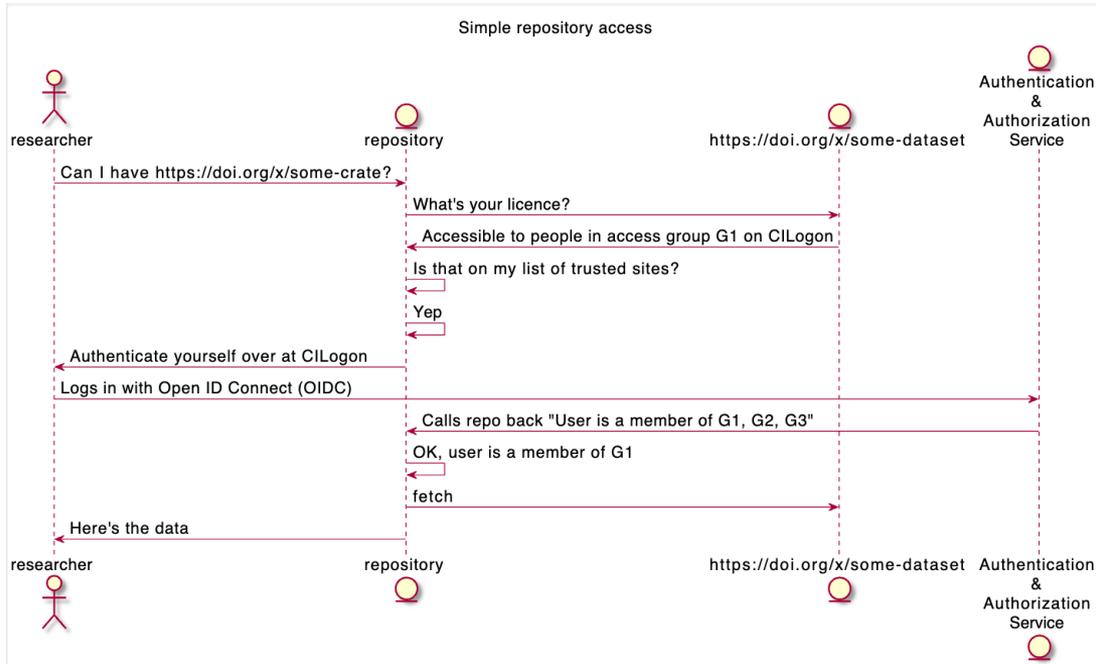
Indigenous Data  
Network

# Trove

- An independent consultation phase is being carried out to scope researcher requirements for a Trove Research Platform.
- A researcher panel with broad representation has been appointed to advise on the consultation phase.
- The consultation will inform the technical Researcher Portal functionality for the project.



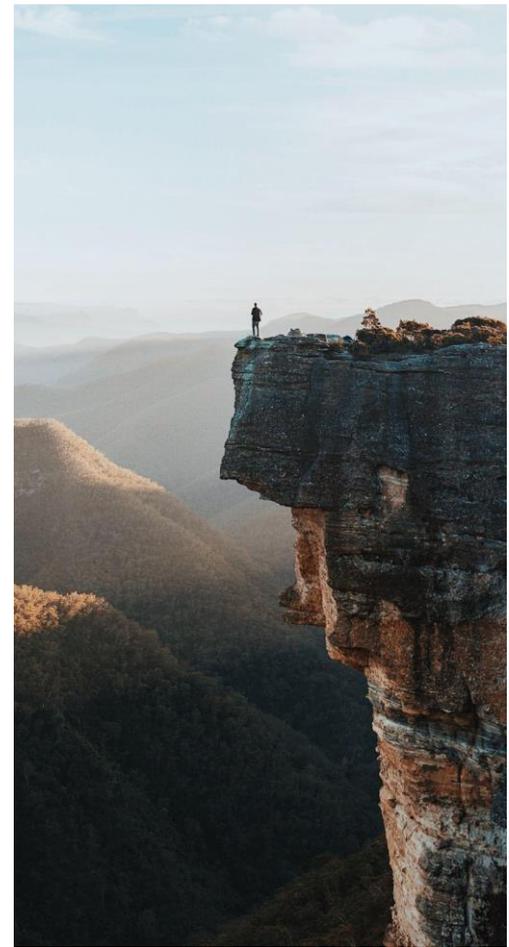
# CILogon





# Capability Gaps

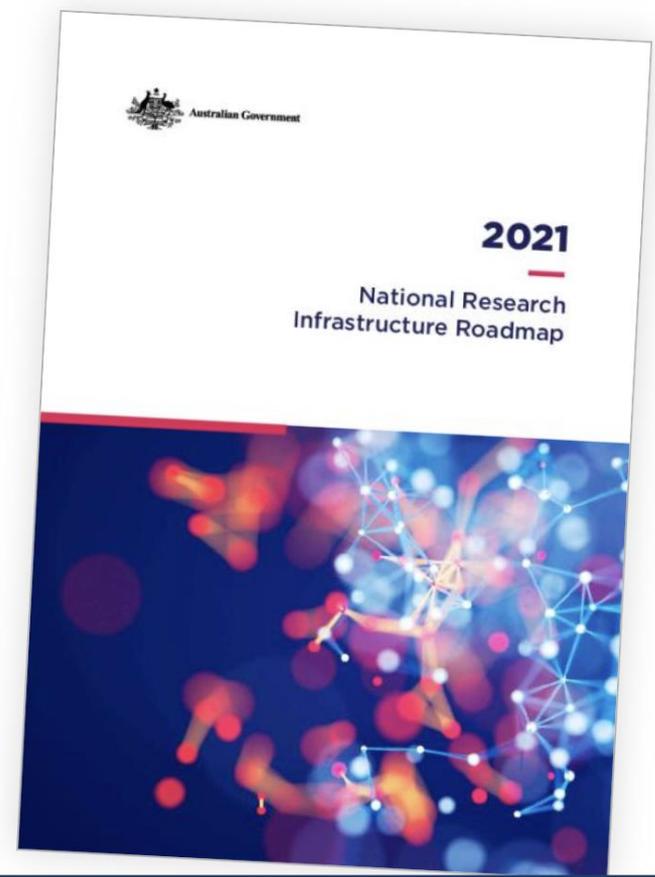
- Broadening sources available and accessible collections
- Social Media data
- Data Retention, Interoperability and Longevity
- Trove API improvements and additions
- Digitisation



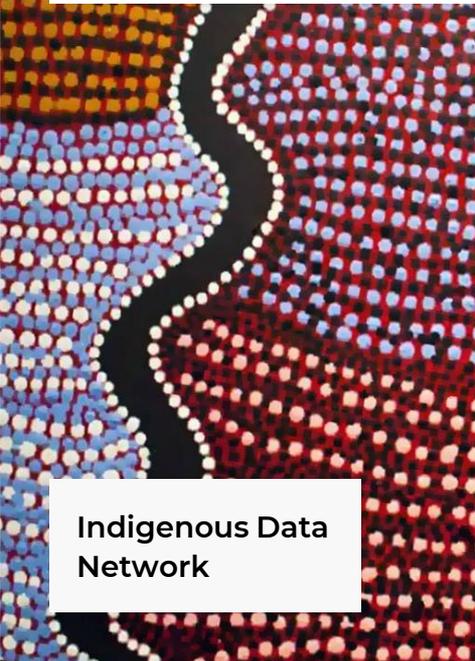
# 2021 Roadmap

“The HASS RDC is nascent and critical capabilities still need to be developed”

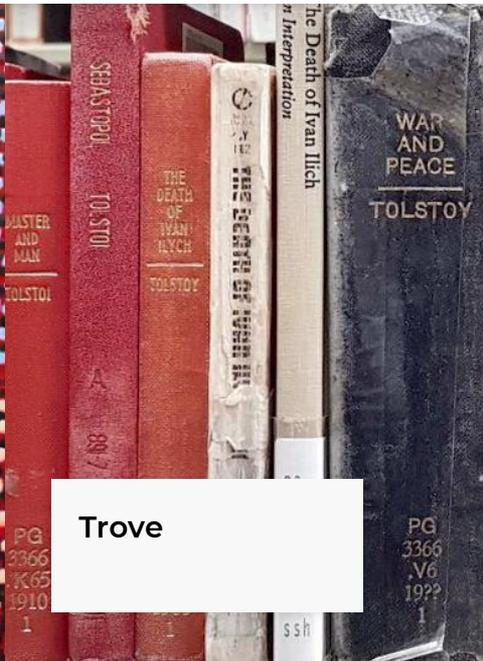
- use of big data
- interoperability of data
- secure data labs
- support for the GLAM sector
- collection digitisation
- improved availability of non-textual data
- tools and guidelines for citizen science approaches
- machine learning tools



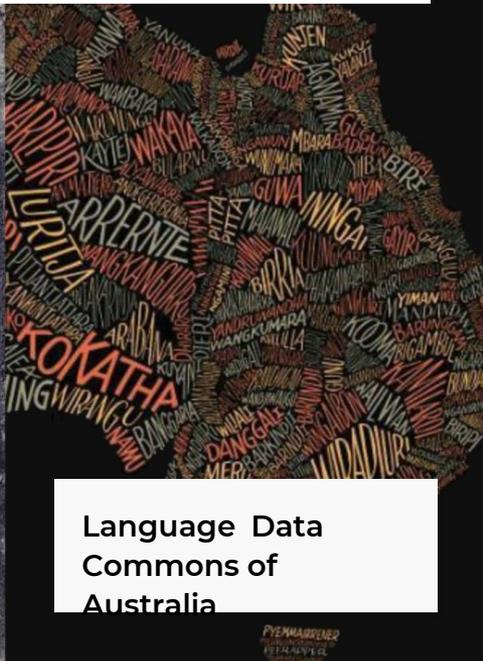
# Humanities Arts and Social Sciences Research Data Commons and Indigenous Research Capability Program



Indigenous Data  
Network



Trove



Language Data  
Commons of  
Australia



Social sciences  
research  
infrastructure



Subscribe to the  
**ARDC CONNECT**  
newsletter

## THANK YOU



[ardc.edu.au](http://ardc.edu.au)



[contact@ardc.edu.au](mailto:contact@ardc.edu.au)



+61 3 9902 0585



[@ARDC\\_AU](https://twitter.com/ARDC_AU)



[Australian-Research-Data-Commons](https://www.linkedin.com/company/Australian-Research-Data-Commons)

# Australian Text Analytics Platform (ATAP)

The Australian Text Analytics Platform (ATAP) is an open source environment that provides researchers with tools and training for analysing, processing, and exploring text. It will encourage researchers to adopt new methods, leading to greater flexibility and transparency in research workflows.

Text analytics is a suite of methods which enable data-driven research by extracting and analysing machine-readable information from within unstructured text. Due to the increasing availability of large amounts of unstructured text, such techniques are becoming more and more important across diverse research disciplines.



# The Digital Observatory

Social media is a rich source of data that researchers from many disciplines increasingly incorporate into their research. For a number of years, many groups across Australia have independently collected significant social media holdings to serve their local research community. The Australian Digital Observatory project seeks to sustainably align and connect selected established platforms into a national infrastructure that supports an extended and diverse array of research communities.

