HASS Research Data Commons and Indigenous Research Capability project plan

1. Project title

Improving Indigenous Research Capabilities: An Aboriginal and Torres Strait Islander Research Data Commons

2. Lead contact

The organisation with whom ARDC will contract.

- First name: Marcia
- **Last name**: Langton
- **Organisation:** University of Melbourne
- **Group/Department:** Indigenous Data Network (Indigenous Studies Unit, Centre for Health Equity, Melbourne School of Population and Global Health)
- **Job titles:** Associate Provost, Chair of Australian Indigenous Studies, Director Indigenous Studies Unit, Co-Chair Indigenous Data Network
- **State:** Victoria
- **Email:** m.langton@unimelb.edu.au
- **Phone number:** +61383449159
- **ORCiD (optional): 0000-0002-1398-7524**

3. Proposal summary

Representing Aboriginal and Torres Strait Islander researchers, organisations and communities across Australia, this project will review and refine national and international frameworks of Indigenous Data Governance (IDG) and Indigenous Data Sovereignty (IDS) to collectively strengthen the foundations of Aboriginal and Torres Strait Islander data governance principles, of which will be translated and applied to Indigenous research data tools and infrastructure. The Indigenous Data Network (IDN) will deliver an Indigenous research capability program

that celebrates, supports and enhances the capabilities of Aboriginal and Torres Strait Islander peoples and researchers at the interface of research data science and Indigenous knowledge systems.

The program will achieve this by working with Aboriginal and Torres Strait Islander communities to determine their engagement with, and aspirations for, their data. The program will be delivered by the development of three parallel streams of activities:

(1) Development of social architecture: Indigenous data governance and sovereignty
(2) Development of technical architecture: Building the foundations for an Aboriginal and Torres Strait Islander Research Data Commons
(3) Core National Indigenous Data Assets: Building an Aboriginal and Torres Strait Islander Spatio-temporal Framework.

Collectively, the three activity streams will lay the foundation to build national Indigenous research capabilities, framed by a set of agreed Indigenous Data Governance principles that can leverage existing data assets, linking them with new and existing Aboriginal and Torres Strait Islander data assets. This project will generate a detailed account of research data ecosystems, including how research data is distributed and used for the benefit of Aboriginal and Torres Strait Islander people.

Indigenous Data Network

The IDN was established to empower Aboriginal and Torres Strait Islander communities to decide their own local data priorities, inclusive of how their data is utilised, and that this is supported and reinforced by our national legal and policy frameworks, as well as the FAIR and CARE Data Principles and principles of Indigenous data democracy, stewardship, analysis, dissemination and infrastructure. Data is relevant to self-determination and self-governance. The need for data ecosystems needs to be accountable to Indigenous Australians as much as it needs to benefit needs of importance to them (Indigenous Data Sovereignty Summit 2017).

The IDN connects individuals, communities, organisations, institutions and academic partners, nationally and internationally, to empower Aboriginal and Torres Strait Islander communities and organisations to achieve these goals. Further, the IDN aims to engage with and leverage internationally leading developments in the data sciences to maximise the optimal collection, access and use of data resources for community empowerment. It is a partnership between the University of Melbourne, the Australian National University and the Australian Institute of Aboriginal and Torres Strait Islander Studies (AIATSIS), coordinated by the Indigenous Studies Unit at the University of Melbourne. The IDN will work with its partners and members to engage the key leaders and stakeholders of Indigenous research across Australia to improve Indigenous research capabilities.

Project Aims

A core aim of this project is to develop Aboriginal and Torres Strait Islander research data capabilities by supporting and strengthening Indigenous social and technical research data architecture, bringing together key national stakeholders to collectively agree on a core set of Indigenous Data Governance principles and data for governance which will be applied to Aboriginal and Torres Strait Islander research data, tools and analytical methods. This project seeks to ensure the cultural and ethical safety of the data and the individuals and communities from which the data is about.

The project will leverage ARDC capabilities, supporting institutions to share data more freely and cooperatively following the FAIR (Findable, Accessible, Interoperable and Reusable) and CARE data principles (Collective benefit, Authority to control, Responsibility and Ethics).

Project Significance

Indigenous research data exists in a wide variety of formats in a wide variety of contexts. At present, these assets are scattered and inadequately catalogued: this limits data discovery and use, results in duplication and over-researching, and prevents aggregation of related data to form high-level national assets. The three streams of activities outlined in this project plan are designed to develop the foundations from which to remedy these problems, incorporating international best practice data science and data governance and Aboriginal and Torres Strait Islander knowledges and worldviews to provide a bridge between Indigenous research data owners or custodians and Indigenous research data users and service providers. We would also look at those other actors in the information supply chain that provide the information platforms or infrastructures to manage data.

There are five broad communities of Indigenous research data custodians:

(1) Aboriginal and Torres Strait Islander communities and organisations;

(2) universities, researchers and other research organisations;

(3) the GLAM sector, including libraries, archives and museums;

(4) Commonwealth, state, territory and local governments; and

(5) Private sector (including NGOs).

These custodians are largely siloed, both within and between communities, and have widely varying data science practices and capabilities. Thus, the contribution needed from this Program also varies. While this will be prioritised with each community according to the level of readiness and the demand from researchers, it is

expected that the main challenge will be to assist with data lifecycle management and the provision of metadata to promote discoverability. Other challenges include digitisation of legacy assets, translation of interoperability between data infrastructures/platforms, translations of traditional indexing methods to machine-readable metadata, and the development of schema which translate existing metadata to a consistent basis.

Expected Project Outputs

a. The development and application of Aboriginal and Torres Strait Islander Data Governance Principles that are agreed on by experts and communities.

b. The application of the Aboriginal and Torres Strait Islander Data Governance Principles.

c. Improved understandings of Indigenous Research communities across Australia.

d. Improved understandings of Indigenous Research data across Australia.

e. The development of an Aboriginal and Torres Strait Islander Research Data Catalogue.

f. Aboriginal and Torres Strait Islander data 'rematriation' prototyping.

g. The technical architecture will be based on a catalogue of data assets using DCAT-2, the international standard for interoperability between data catalogues published on the Web.

4. Project partners

University of Melbourne (UoM)

The University of Melbourne is an international leader in Indigenous research with a strong track record of both Australian competitive research grants and public sector research. The University's income for research in Indigenous topics in 2018 was $18.5 million, and its income from projects involving one or more Indigenous Chief Investigators was $3.45 million. The University is committed to respectful partnerships and collaborations with Indigenous communities and organisations across our research activities. The University prioritises Indigenous research and invests in flagship Indigenous research initiatives.

The most recent of these is the Indigenous Knowledge Institute, which launched in 2020 with unprecedented core funding support of $6 million over an initial five years to help sustain and advance research into Indigenous knowledge systems. The University's other flagship Indigenous research initiatives include the Indigenous Studies Unit (responsible for the coordination of the Indigenous Data Network), Centre for Health Equity;

Research Unit for Indigenous Languages; Research Unit for Indigenous Arts and Cultures; and Indigenous Studies Program, School of Culture and Communication.

The University's leading Indigenous researchers also work in collaboration with the Melbourne Institute of Applied Economic and Social Research, and the Evaluation and Implementation Science Program in the Centre for Health Policy.

Australian National University (ANU)

As Australia's national university, one of our defining roles is to contribute to the advancement of Australia's Indigenous peoples. We host some of the nation's best Indigenous academics who are enthusiastic about nurturing and mentoring the next generation of academic leaders from across the country and around the world. ANU is a place where First Nations peoples and cultures are respected, celebrated and reflected in our research and education.

As a branch of the University's executive, the First Nations Portfolio works across the University to ensure the ANU is a world leader in teaching and research of First Nations issues as well as ensuring that the University makes a leading contribution to national policy in the relationship between Indigenous Australians and the nation.

The University is committed to research excellence and supports meaningful collaboration with Indigenous communities through a range of Centres and initiatives including the Tjabal Indigenous Higher Education Centre, the Centre for Aboriginal Economic Policy Research, the Australian Centre for Indigenous History, The National Centre for Indigenous Genomics, Mayi Kuwayu: The National Study of Aboriginal and Torres Strait Islander Wellbeing, and the ANU Grand Challenges - Indigenous Health and Wellbeing project.

Australian Institute of Aboriginal and Torres Strait Islander Studies (AIATSIS)

AIATSIS is Australia's national Institute dedicated to Aboriginal and Torres Strait Islander peoples' knowledge, societies and cultures. We are both the custodian and repository of Australia's national collection of Aboriginal and Torres Strait Islander heritage materials and one of Australia's publicly funded research agencies. Among AIATSIS statutory responsibilities are to provide leadership in Aboriginal and Torres Strait Islander research, research ethics and use and protocols for collections related to Aboriginal and Torres Strait Islander peoples., AIATSIS published the Code of Ethics for Aboriginal and Torres Strait Islander Research in 2020 which includes guidance on ethical practice in relation to Indigenous data. AIATSIS contributes to meta data standards, in particular through AIATSIS language thesaurus/Austlang. In 2018 AIATSIS established the Indigenous Research

Exchange. The purpose of the Indigenous Research Exchange is to connect people and ideas, build the evidence base to support Aboriginal and Torres Strait Islander peoples' decision making and inform policy and practice. A key element of the Indigenous Research Exchange is the Knowledge Exchange Platform, a portal that aims to provide a central discovery point for Indigenous research and data that:

- showcase and support Indigenous knowledge translation.
- is a powerful resource for Indigenous leadership and community-led decision making
- support nation-building through specific nation-based data
- promote IDG, IDS and self-determination
- change the narrative on Indigenous data collection, use and ownership; and
- provide links and networks to other data sources and knowledge repositories, including the wide range of information available within AIATSIS.

Empowered Communities

First established in 2013, Empowered Communities (EC) is a nation-wide initiative designed and led by Aboriginal people to shift the 'top-down' approach to Indigenous affairs in Australia. The EC approach focuses on supporting Aboriginal empowerment, development, and productivity to address issues of social and economic disadvantage. EC shifts away from an Indigenous affairs agenda to a 'comprehensive Indigenous Empowerment agenda' that includes long-term reforms in partnership between Indigenous leaders, governments, corporations.

Commonwealth Scientific and Industrial Research Organisation (CSIRO)

As Australia's national science agency, CSIRO's purpose is to solve the greatest challenges through innovative science and technology. These challenges can only be met through collaborative action involving the community, industry, government and research sectors. CSIRO operates through three lines of business: Impact science: Nine national research business units with focus on the biggest challenges facing the nation; National Facilities and Collections: managing infrastructure and biological collections for the benefit of research and industry; CSIRO Services: Commercial, customer-centric products and services for industry, government and the community. CSIRO has also established a whole of organisation Indigenous Science and Engagement Program, which includes the Office of Indigenous Engagement, and sits within CSIRO Office of the Chief Scientist. CSIRO also has a strong Indigenous STEM education program, including the Young Indigenous Women's STEM Academy.

CSIRO acknowledges the extraordinary contributions Aboriginal and Torres Strait Islander people have made, and continue to make, to our culture, the economy, and science. CSIRO is working with Indigenous communities and organisations to create Indigenous-driven science solutions that support sustainable futures for Indigenous peoples, cultures and Country. See here for examples of CSIRO Indigenous Science - https://www.csiro.au/en/research/indigenous-science.

Queensland University of Technology (QUT)

The QUT Centre for Data Science run by Professor Kerrie Mengerson, encompasses about a hundred researchers from across the University, with a vision of developing methods for better use of data to benefit the world. Their core research focus includes fields of data management and governance, data analysis and visualisation, human-focused data science, and data-focused decision-making. Their researchers bring expertise in a wide range of areas of potential relevance to the IDN, including statistics and machine learning, information science, health, environment, business, industry, digital media and more. The CDS is also a lead node in the Australian Data Science Network, which connects around 25 research centres in data science across the country.

CDS also works with the Carumba Institute for Indigenous research and education. Professor Peter Anderson is the Institute's Director and the CDS and CI are currently working on a number of joint research projects as well as student centred research programs.

5. Project team roles and responsibilities

Name	Project role	Organisation	Responsibility
Professor Marcia Langton	Project Lead	University of Melbourne	**Professor Marcia Langton AO is** the University's Foundation Chair of Australian Indigenous Studies and Associate Provost. Professor Langton is also the Director of the Indigenous Studies Unit (MSPGH) and Co-Chair of the Indigenous Data Network. Professor Langton is an anthropologist and geographer who holds extensive experience of working with native title holders and PBCs,

			and in evaluating native title and Indigenous programs. She has produced a large body of knowledge in the areas of political and legal anthropology, Indigenous agreements and engagement with the minerals industry, and Indigenous culture and art, and has extensive experience contributing to the policy and program evaluation space. Her role in the Empowered Communities project under contract to PM&C and as a member of the Expert Panel on Constitutional Recognition of Indigenous Australians are evidence of her academic reputation, policy commitment and impact, alongside her role as a prominent public intellectual. She was Chair of the National Aboriginal and Torres Strait Islander Curricula Expert Panel under contract to PM&C. She has worked for many years with the National Native Title Council on PBCs and related issues, including the ongoing development of her ground-breaking Agreements, Treaties and Negotiated Settlements Database and Australian Research Council Linkage grants (LP0990125, LP0561857, LP0211472) and her expert advice on this topic is often sought. Working collaboratively with the project team, Professor Langton will provide overarching intellectual leadership and guidance framing the project across all three Streams of Project Activities. Professor Langton will also lead Stream 1 of the Project Activities.

| Dr Vanessa Russ | Senior Research Associate | University of Melbourne | **Dr Vanessa Russ** is a Ngarinyin/Gija women from the Kimberley region of Western Australia. She is the former Associate Director of the Berndt Museum of Anthropology at the University of Western Australia. Her extensive skills in managing that collection as well as working in both the not-for-profit and government sectors ensures that she brings the capabilities and community knowledge that is required for this project. Dr Russ will be responsible for leading contributions to Streams 1 and 2. |
| Dr Kristen Smith | Senior Research Associate | University of Melbourne | **Dr Kristen Smith** is a Senior Research Fellow and Research Director of the Indigenous Studies Unit of the Centre for Health Equity at the Melbourne School of Population and Global Health. Smith the Research Coordinator of the Indigenous Data Network and a founding member of the Indigenous Data Network Steering Committee. Dr Smith has an extensive history of implementing large multi-sited projects on Indigenous alcohol management, Indigenous digital health ICTs and family violence with regional, remote and very remote Indigenous communities. Her expertise of diverse mixed-methods research with Indigenous Australians in challenging research environments has led the translation of her collaboratively informed research findings into policy change and resource prioritisation. For the past eight years she has also worked with Prof Langton in the ongoing development of the ground-breaking |

			Agreements, Treaties and Negotiated Settlements Database, which is the largest database of its kind in the world. Her interdisciplinary work traverses the fields of medical anthropology, epidemiology, human geography, public health, international development and health promotion. Dr Kristen Smith will make leading contributions to Streams 1 and 2 of the Project Activities.
Professor Aaron Corn	Senior Research Associate	University of Melbourne	**Professor Aaron Corn** is the Inaugural Director of the Indigenous Knowledge Institute. He serves as Director of the National Recording Project for Indigenous Performance in Australia, and was Arts Learning Area Leader on the PM&C National Aboriginal and Torres Strait Islander Curricula Expert Panel. His work with Indigenous colleagues, heritage and applications of information technologies engages with intellectual traditions that remain fundamental to Indigenous cultural survival in Australia and foregrounds the unique perspectives of Indigenous Australians on public opinions, government policies and scholarly debates that impact upon the cultural, economic, political and technological futures of their communities. His work Indigenous peoples to apply new information technologies to discovering and accessing their heritage in collections has engendered new approaches to curatorial policies and practices among institutions worldwide. Professor Corn will make significant

			contributions to Streams 1, 2 and 3 of the Project Activities.
Dr Kalinda Griffiths	Senior Research Associate	University of New South Wales	**Dr Kalinda Griffiths** is a Yawuru woman and early career Scientia Lecturer at the Centre for Big Data Research at the UNSW. She also holds and honorary positions at Menzies School of Health Research and the University of Melbourne. Dr Griffiths is an epidemiologist who has worked in the research sector in a number of roles for over 20 years. Dr Griffiths is on a number of international and national data committees. A few examples currently include the International Group for Indigenous Health Measurement, the IDN Steering Committee and the Australian Burau of Statistics Aboriginal and Torres Strait Islander Health Surveys Advisory Group. She has also been a commissioned expert to advise on a range of initiatives, including the development of the Australian Institute of Health and Welfare's National Health Information Strategy (2020-2021), the National Indigenous Australia's Agency Housing Conditions and Early Childhood Development Research Project (2019-2020) as well as the Australian Human Rights Commission, National Children's Commissioner's Data 2018 Roundtable. Dr Griffiths will co-lead, with Dr Len Smith, Streams 2 and 3 of the Project Activities.

Dr Len Smith	Senior Research Associate	Australian National University	**Dr Len Smith** is a renowned demographer, with an extensive background of research focusing on Aboriginal Australian people. Dr Smith was able to reconstitute the Aboriginal population that was not recorded in the Protection Board records and demonstrate how the Victorian Aboriginal community recovered to reach its current strength of over 30,000 people who identified in the last Commonwealth census. Dr Smith sits on multiple Indigenous Data Expert Advisory Committees, including the AIHW Indigenous Data Advisory Group and is a key leader of Indigenous data research at the ANU. Dr Smith is also a founding member of the IDN Steering Committee. Dr Smith will co-lead, with Dr Griffiths, Streams 2 and 3 of the Project Activities.
Sam Provost	Research Associate	Australian National University	**Sam Provost** is a Yuin man from the far south coast of NSW, focused on developing collaborative, community-led research built around supporting Indigenous people and communities to leverage emerging digital technologies in the pursuit of self-determination over their lands and waters. Sam is the Project Lead for the Australian National University and will be responsible for the coordination and implementation of ANU partnership activities. Sam is a Lecturer and PhD scholar at the Fenner School of Environment and Society at the ANU, a project coordinator at the ANU First Nations Portfolio, and is an Executive Member of Maiam nayri Wingara

			Aboriginal and Torres Strait Islander Data Sovereignty Collective.
Dr Steven McEachern	Senior Research Associate	Australian National University	**Dr Steven McEachern** is Director and Manager of the Australian Data Archive at the Australian National University, where he is responsible for the daily operations and technical and strategic development of the data archive. He has high-level expertise in survey methodology and data archiving, and has been actively involved in development and application of survey research methodology and technologies over 15 years in the Australian university sector. Dr McEachern will make leading contributions to Stream 2 of the project activities.
Dr Lisa Strelein	Senior Research Associate	AIATSIS	**Dr Lisa Strelein** is the AIATSIS Executive Director of Research and Education. Dr Strelein is one of Australia's leading experts in Native Title law, publishing ground-breaking works, designing and leading the National Native Title Conference and building capacity in the nation's Native Title sector for over 20 years. Dr Strelein has also made immense contributions to education and cultural understanding, overseeing the development of the AIATSIS Core cultural competency program, the Guidelines for Ethical Research in Australian Indigenous Studies (GERAIS) and the landmark primary education series, Our Land, Our Stories. Dr Strelein will make contributions to Streams 1 and 3 of the project activities.

Tamsin Porter	Research Associate	AIATSIS	**Tamsin Porter** is the Project Officer, Ethics and Research Leadership/Indigenous Research Exchange at AIATSIS. Tamsin will make contributions to Stream 3 of the project activities.
Ash Pollock-Harris	Research Associate	AIATSIS	**Ash Pollock-Harris** is the A/g Director, Indigenous Country and Governance, Research and Education Group at AIATSIS. Ash is a public servant with over 7 years' experience in various roles. His background is from the Ngunnawal and Wiradjuri people of the ACT and central NSW, and his passion is reconnecting Aboriginal and Torres Strait Islander people with their culture, materials and history. Ash works on several research projects at AIATSIS, including Preserve, Strengthen and Renew in Community (PSR) and the AIATSIS Indigenous Language Preservation: Dictionaries Project. Ash will make contributions to Streams 1 and 3 of the project activities.
Anthony McLaughlin	Research Associate	AIATSIS	**Anthony McLaughlin** is an Assistant Director in the Collection Development and Management program responsible for Descriptive Metadata and Cataloguing at AIATSIS. Anthony was instrumental in the adoption of the AUSTLANG codes as a language source for MARC cataloguing and worked closely with AIATSIS and National Library of Australia colleagues in preparing the application to the Library of Congress. Anthony collaborated with AIATSIS Senior Linguist, Amanda Lissarrague, to present the May 2019 webinar 'In Search of Indigenous

			Australian Languages', and took an active role in a code-a-thon during NAIDOC week. Anthony is committed to strengthening practices around the Indigenisation of collections and collaborating on matters of cultural safety in resource description. Anthony will make contributions to Streams 1 and 3 of the project activities.
Stephanie von Gavel	Senior Research Associate	CSIRO	**Stephanie von Gavel** has over 25 years experience in technology transfer, business development and strategy development in a range of sectors - from biotechnology, agriculture, biodiversity, information management, international research for development and inclusive innovation in an Indigenous context. Stephanie has strong interpersonal, strategic, project management and conceptual (design) abilities which can be applied to the development and implementation of strategic plans, commercial arrangements and relationships in a collaborative and international context. She brings a broad skill set as well as invaluable experience on both sides of the commercialisation fence, having led projects on behalf of both technology providers and technology acquirers. Her background provides an important perspective and appreciation of the motivations and objectives of each partner in a successful commercial or research collaboration/ alliance. Stephanie will make contributions across all three project streams.

| Dr Chris Bourke | Senior Research Associate | CSIRO | **Dr Chris Bourke** (Gamillaroi) is Program Director - Indigenous Science and Engagement at CSIRO, leading and managing CSIRO's Indigenous Science and Engagement program including co-development of Indigenous science priorities for CSIRO. Dr Bourke served in the ACT Legislative Assembly from 2011-2016 and held many Ministerial portfolios including Education, Children & Young People, Aboriginal & Torres Strait Islander Affairs, Disability, Corrections, Industrial Relation, Small Business, Arts, Veterans Affairs and Seniors. Chris is a graduate of The University of Melbourne and the first Indigenous Australian dentist. After an extensive career in public dentistry Chris moved to Canberra in 1993 where he ran a successful private dental practice for 16 years. His strong focus on community health led him to provide his clinical skills pro bono and policy making skills in many political and professional arenas. Dr Bourke will make contributions across all three project streams. |
| Professor Kerrie Mengerson | Senior Research Associate | Queensland University of Technology | **Prof Kerrie Mengerson** is a statistician with about thirty years of experience in the development of analytic methods to improve the use of data to address challenges in health, the environment and industry. Prof Mengerson is a Distinguished Professor in Statistics at QUT and Director of the QUT Centre for Data Science. She is an elected Fellow of the Australian Academies of Science and Social Sciences, and an executive member of four international professional |

			societies. Prof Mengerson is committed to training the next generation in collaborative data science and have supervised around forty postgraduate students in statistics, public and allied health, and ecology. She is also strongly committed to increasing data literacy, agency and capability, so that everyone can engage in how their data are shared, analysed used for better understanding and decision-making. Professor Mengerson will make contributions across all three project streams.
Prof Peter Anderson	Senior Research Associate	Queensland University of Technology	**Prof Peter Anderson** is the Director of the Carumba Institute, which is also a cross-University Centre at QUT that advances Indigenous research and education. Professor Anderson is from the Walpiri and Murinpatha First Nations in the Northern Territory. His own research expertise is in organisational leadership, Indigenous peoples' education, and teacher and academic professional development. He was also the Director of the National Indigenous Research and Knowledges Network (NIRAKN). The two Centres are engaged in a number of joint research projects and we are planning a program for jointly supervised research students. Prof Anderson will make contributions across all three project streams.
Prof Matt Bellgard	Senior Research Associate	Queensland University of Technology	**Prof Matt Bellgard** is the eResearch Director at QUT. The eResearch Office provides computing infrastructure and support for researchers across the University. He brings expertise in health informatics, AI,

			biosecurity, HASS, data management and data governance. Matt is Chair of the APEC Rare Disease Network and leads the Clinical Data and Analytics Platform (CDAP) that aims to create infrastructure and a governance framework for sharing and rapid access/utilisation of health data for clinicians, to better inform patient care. The eResearch Office has close ties to the CDS and is a Domain in the Centre. Prof Bellgard will make contributions across all three project streams.
Prof Rachel Thomas	Senior Research Associate	Queensland University of Technology	**Prof Rachel Thomas** is the inaugural CDS Data Scientist in Residence. Rachel brings expertise in data ethics, AI accessibility and bias in machine learning. Rachel was the founding director of the University of San Francisco Center for Applied Data Ethics, which aims to address harms such as disinformation, surveillance, algorithmic bias, and other misuses of data. Her writing has been read by nearly a million people. She is also co-founder of fast.ai, which created the most popular deep learning course in the world. Prof Thomas will make contributions across all three project streams.
Kirsten Thorpe	Research Associate	University of Technology Sydney	**Kirsten Thorpe** (Worimi, Port Stephens NSW) has led the development of protocols, policies, and services for Aboriginal and Torres Strait Islander peoples in libraries and archives in Australia. Kirsten's research interests relate to Indigenous self-determination in libraries and archives. She has contributed to numerous projects that have involved the return of historical

			collections to Indigenous peoples and communities and advocates for a transformation of practice to centre Indigenous priorities and voices concerning the management of data, records, and collections. Kirsten has broad interests in research and engagement with Indigenous protocols and decolonising practices in the library and archive field. Kirsten will make contributions across all three project streams.

6. Project objectives

List the objectives of this project. Include a description of any existing infrastructure, data collections, platforms, services and tools that are being leveraged as input for this project. (500 words)

The core objectives of the Indigenous Research Capability program are to:

a. collectively agree on, and apply, a set of uniquely Australian, Aboriginal and Torres Strait Islander Data Governance Principles;

b. enable Aboriginal and Torres Strait Islander organisations and researchers to manage their data sustainably;

c. establish the breadth and depth of Aboriginal and Torres Strait Islander research data available across Australia that are relevant to the needs of Aboriginal and Torres Strait Islander peoples;

d. develop the foundations of an Aboriginal and Torres Strait Islander Data Commons; and

e. value adding through access to appropriate aggregation and analysis tools.

The project framework embeds Indigenous Data Governance and Sovereignty across all streams of activities.

7. Project activities

The project activities are organised into three key streams (see Figure 1):

1. Social Architecture
2. Technical Architecture

3. National Data Assets

The three streams of activities are designed to collectively lead to improved Indigenous research data practice across Australia.
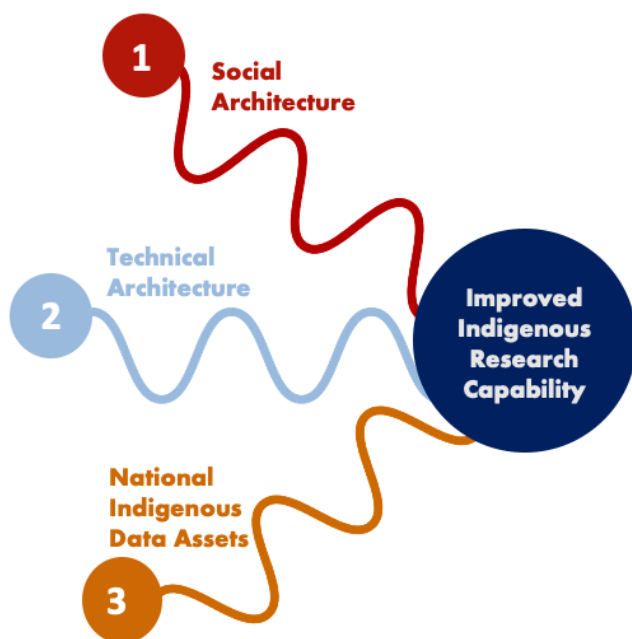

Figure 1. The Three Streams: Indigenous Research Capabilities Project Activities

Stream 1. Social Architecture: Indigenous Data Governance and Sovereignty

The activities in Stream 1. are designed to bring together Aboriginal and Torres Strait Islander national Indigenous Data Governance leaders and stakeholders to collectively review, refine and agree on a set of core, and uniquely Australian, Aboriginal and Torres Strait Islander Data Governance principles that will inform and be applied to all other streams and activities of this project. It is anticipated that the agreed Aboriginal and Torres Strait Islander Data Governance principles will have broader application beyond the life of this project, contributing to a culturally safe national research data ecosystem. This will also recognise the AIATSIS Code of ethics and other relevant scholarship available to the public.

Activity 1.1: National Scoping and Engagement of Indigenous Research Communities and Place-based Application of IDG Principles

The IDN will undertake national scoping and engagement activities to gather information on Aboriginal and Torres Strait Islander research data ecosystem, and the use of IDG Principles within these communities. This work will encompass all publicly available research and tools.

A1.2 Key Objectives

This activity will seek to:

a. Identify Aboriginal and Torres Strait Islander research communities across Australia within each of the four groups of data custodians and owners (as detailed in the Project Aims).

b. Improve understandings of Aboriginal and Torres Strait Islander community perceptions of Indigenous Research Data and Indigenous Data Governance and Sovereignty principles.

c. Inform the development of agreed Indigenous Data Governance Principles for the project.

A1.1 Key Deliverables

Information pertaining to the key groups and subgroups that make up the Aboriginal and Torres Strait Islander research data ecosystem across Australia.

Information about, and descriptions of, the various understandings and application of IDG principles in different contexts.

Activity 1.2: National Leaders of Indigenous Data Governance and Sovereignty Roundtable

Co-convened by ANU, UoM and AIATSIS, the National Leaders of Indigenous Data Governance Roundtable will bring together Aboriginal and Torres Strait Islander leaders and experts, and relevant non-Indigenous scholars, to highlight and discuss the wide variety of views, understandings, principles and protocols related to Indigenous Data Governance and Sovereignty. Roundtable participants will be selected drawing on information gathered by Activity 1.1 and IDN partners and members.

A1.1 Key Objectives

The key findings of the National Scoping and Engagement activities (1.1) will be outlined and discussed during the Roundtable, with the aim of agreeing on a core set of Aboriginal and Torres Strait Islander Data Governance Principles and Technical Specifications that will be applied to all

project activities, with the potential for broader application across Aboriginal and Torres Strait Islander research tools and platforms nationally.

A1.2 Key Deliverables

The key deliverables for this Activity are:

    a. The delivery of a National Leaders of Indigenous Data Governance and Sovereignty Roundtable.

    b. A core set of agreed Aboriginal and Torres Strait Islander Data Governance Principles that will provide the overarching framework for the Indigenous Research Capabilities project activities.

Activity 1.3: Place-based case studies of applied Aboriginal and Torres Strait Islander Data Governance Principles

The project partners will develop a series of place-based case studies that reflect local variations of Indigenous data governance principles, use and management of data. They may also want to matrix this with certain data types, cultural heritage, environment, administrative to extract the nuances related to governance that might impact Indigenous Data Governance.

A1.1 Key Objectives

The primary objective of developing a series of place-based place-based case studies is to develop improved contextual understandings of the variations in the application of Indigenous Data Governance practices within Aboriginal and Torres Strait Islander communities and organisations.

A1.2 Key Deliverables

The key deliverables for this Activity are information related to the application of IDG principles in varying contexts that will be used to dynamically inform the social and technical underpinnings of the project.

Stream 1. Milestones and Deliverables

Timeframe	Milestone	Deliverables

2021: Q4	1.1 Stream Establishment Activities	Establishment of Governance Groups Recruitment for key Stream 1 project roles Assignment of Stream 1 Activity leaders Development of Stream 1 Implementation Plan
2022: Q1/2	1.2 Engagement Activities	Undertake Project Activity 1.1 National Leaders Roundtable Development of agreed IDG Principles for project
2022: Q3/4	1.3 Case Studies	Selection of place-based case studies Implementation of case studies for dynamic refinement of ATSIDG Principles
2023: Q1/2	1.4 Extension Activities	Widely share progress and outcomes of agreed ATSIDG principles Extension of testing of agreed ATSIDG principles across broader tools, platforms and infrastructure

Stream 1. Key Partners

Australian National University

University of Melbourne

Queensland University of Technology

AIATSIS

Empowered Communities

Stream 2. Technical Architecture: Building the foundations for an Aboriginal and Torres Strait Islander Research Data Commons

Aboriginal and Torres Strait Islander research data assets are scattered and inadequately catalogued: leaving them vulnerable to loss, limiting data discovery and use, and resulting in the duplication and the over-researching of Aboriginal and Torres Strait Islander communities, preventing aggregation of related data to form important national assets. The potential benefits of

understanding Indigenous data and its uses are not only directly linked to what communities' needs might be, but also provides an economic development and commercial application for those communities.

The aim of the project activities of Stream 2 are to address these issues by working with the ARDC to establish the foundations of an Aboriginal and Torres Strait Islander Research Data Commons, exploring repository services to ensure preservation of data assets, catalogue services to make them discoverable, and metadata generation to improve asset description and facilitate data linkage and aggregation.

The first step will be to undertake engagement and scoping activities investigating the Aboriginal and Torres Strait Islander data landscape, including legacy and current data held in universities, archives and government, as well as other components of the information supply chain such as information management software/platforms/infrastructure, data aggregators, and data users and beneficiaries.

Activity 2.1 Aboriginal and Torres Strait Islander Research Data Engagement and Scoping

The IDN will conduct broader community engagement, drawing on information gathered in Activity 1.1. to facilitate the discovery of existing, legacy and proposed research data assets.

Lead by ANU, the IDN will develop an improved understanding of Indigenous research data across Australia. The IDN will develop an inventory of the Aboriginal and Torres Strait Islander data landscape across an initial set of organisations and individuals to determine the location and state of these data. This inventory will provide a picture of Indigenous research data accessibility, completeness and interoperability. The inventory will identify data asset service requirements that will facilitate ongoing discovery, identification and prioritisation of data assets for subsequent work packages.

As a part of this process, online submissions will be requested from custodians and owners of orphaned datasets and archives from researchers on Indigenous topics from any discipline. Some work has been carried out by partners already but the model of use for this research seeks to focus on Federated data ecosystems for best practice.

A2.1 Key Objectives

Using methods developed by CSIRO and Geoscience Australia, the IDN will collaborate with Institutional data owners and custodians to determine their data holdings. Finding ways of sharing through secure networks. Secure data access infrastructure (CADRE) and secure high capacity

networks (AARNet) for transfer of data between researchers and facilities, and to enable remote access to secure facilities.

A2.1 Key Deliverables

Preliminary identification of Aboriginal and Torres Strait Islander data assets held across Australia by the four categories of data custodians and owners.

Activity 2.2 National Aboriginal and Torres Strait Islander Research Data Catalogue

A2.2 Key Objectives

Based on Activity 2.1, a flexible and advanced conceptual model of Aboriginal and Torres Strait Islander data assets will be incorporated in DCAT-2, identifying the key categories, variables, classifications (i.e. the ANZRC FOR classifications) as one key mechanism that could be used by repositories and characteristics of the assets involved. This will form the basis of a best practice framework for cataloguing and describing these assets, incorporating a metadata generation and capture process, data provenance, variable and item-level linkability, and ARDC persistent identifiers.

To facilitate compatibility with other relevant national data infrastructure, the Catalogue will incorporate an export function to support synchronisation of metadata with the ARDC RDA repository. It will contribute to ARDC vocabularies with the long-term goal of harmonising them to meet the needs of Indigenous people. It will also facilitate compatibility with other emerging NCRIS initiatives, including particularly the proposed HASS Research Data Commons.

A2.2 Key Deliverables

a. an initial set of organisations and individuals involved in the Aboriginal and Torres Strait Islander Research Data Community

b. an initial set of current and desired capabilities

c. an initial set of data assets and service requirements which facilitate ongoing discovery, identification and prioritisation of data assets for subsequent work packages.

As individual assets are progressively incorporated in the Data Catalogue, the NATSI Data Catalogue will support the development and implementation of standardised custodian APIs, and provide a portal through which remote assets may accessed using links to custodial repositories and associated APIs.

Activity 2.3 Research data capability building for Aboriginal and Torres Strait Islander organisations and researchers

Activity 2.3 will build data management skills that embody Aboriginal and Torres Strait Islander Data Governance principles and enhance the capacity of Indigenous organisations and researchers to adopt best practice when managing their data and making it available for re-use.

A2.3 Key Objectives

This activity will draw on the ARDC's extensive experience in providing training in research data lifecycle management. University and project partners will be encouraged to develop short training courses, and the IDN will develop online and in-person training modules to build project relevant capability for Aboriginal and Torres Strait Islander researchers and organisations.

A2.3 Key Deliverables

- Outputs/deliverables are designed to enable members of the community to use, repeat and extend data asset construction practice which was prototyped and developed in Activity 2.2.

- Detailed training materials and well documented, repeatable workflows derived from the outcomes of Activity 2.2.

Stream 2. Milestones and deliverables

Indicative Dates	Milestone	Deliverables
2021: Q4	2.1 Stream Establishment Activities	- Establishment of Governance Groups - Recruitment for key Stream 2 project roles - Assignment of Stream 2 Activity leaders - Development of Stream 2 Implementation Plan
2022: Q1/2	2.2 Engagement Activities	- Undertake Project Activity 2.1 - Commence development and build of NATSI Data Catalogue (Activity 2.2)

		▪ Commence collaborative development of training materials and modules with partners.
2022: Q3/4	2.3 Case Studies	▪ Continue development of NATSI Data Capabilities (Activity 2.2) ▪ Create and support communities of practice ▪ Share progress and outcomes
2023: Q1/2	2.4 Building capability	▪ Refinement and delivery of Phase 1 of training packages by IDN and partners(Activity 2.3) ▪ Continue to build and support communities of practice
2023: Q3/4	2.5 Extension and refinement	▪ Refinement and further populating of NATSI Data Catalogue (Activity 2.2) ▪ Delivery of Phase 2 of training packages by IDN and Partners

Stream 2. Key Partners:

Australian National University

University of Melbourne

Queensland University of Technology

Commonwealth Scientific and Industrial Research Organisation

Empowered Communities

Australian Institute of Aboriginal and Torres Strait Islander Studies

Stream 3. Core National Indigenous Data Assets: Building an Aboriginal and Torres Strait Islander Spatio-temporal Framework

Place is central to Indigenous Australians' society, economy and culture. Information on place is crucial for Indigenous research, policy-making, program planning and evaluation. It is also often contested. Extracting data on place relevant to specific Indigenous research interests requires an appropriate geospatial framework which also recognises changes over time.

The ABS has developed an Indigenous geospatial hierarchy which is part of Loc-I, an integrated national framework for geospatial hierarchies. It is built on census mesh blocks, the smallest areas for which statistical data is available. These aggregate in turn to Indigenous localities, Indigenous areas, and Indigenous regions.

The ABS hierarchy is an excellent tool for official statistical purposes, but does not reflect Aboriginal and Torres Strait Islander geospatial priorities, which emphasise the link between particular peoples and particular places. Other Indigenous data held by government agencies is currently being made accessible to researchers, and major data integration projects such as DIPA and NIHSI will increase its value. The aim is to develop a framework which can be applied to any data which includes location information.

The aim of this Stream of activities is to develop an Indigenous Loc-I framework which parallels the ABS hierarchy and can be applied to any data which includes location information incorporating a database and formal ontology based on geocoding the AIATSIS map of Australian language boundaries. Although details of the AIATSIS map remain contested, the ontology will provide the foundation from which further map refinements can be incorporated dynamically in alignment with findings from ongoing research of the area, including that from place-based Indigenous knowledges, Native Title determinations, Indigenous Land Use Agreements and Negotiated settlements.

In collaboration with Geoscience Australia, CSIRO and AIATSIS, tools will be developed using Loc-I, the Location Index for spatial data referencing, to harmonise geospatial referencing with Indigenous place names and to facilitate the creation of datasets of national interest by aggregation of disparate data across space and time.

Activity 3.1 Vocabulary development and metadata labelling: Traditional Knowledge (TK) and Biocultural (BC) Labels

TK and BC labels have been developed by ongoing partnerships and testing internationally, facilitating Indigenous communities to define and customise localised conditions and protocols "for sharing and engaging in future research and relationships in ways that are consistent with already

existing community rules, governance and protocols for using, sharing and circulating knowledge and data"[1].

A3.1 Key Objectives

a. Collaborate with IEEE P2890 and Local Contexts NYU for labelling protocol adoption
b. Work alongside CSIRO and AIATSIS to develop consistent and interoperable metadata conventions for TK and BC labels and notices.

A3.1 Key Deliverables

This Activity will develop core vocabularies, ontologies, thesauruses and catalogues.

Activity 3.2: Geospatial Portal

A3.2 Key Objectives

A wealth of geographical information has been collected during Native Title proceedings, and the Indigenous Ranger program is producing vast data on environment and ecology. These are incorporated in Indigenous spatial frameworks which reflect people's relationship to country, including language and tribal boundaries, but are not integrated with Loc-I. This activity will remedy this by building a parallel classification to the ABS hierarchy, incorporating a database and formal ontology based on geocoding Tindale's map of Australian tribal boundaries. Some of that map's detail is contested or outdated, but the ontology will provide a starting point for the ongoing incorporation of further data and critiques from researchers, Native Title determinations, negotiated settlements and contested alternatives.

A3.1 Key Deliverables

- Loc-I as framework
- Digitise and collate Indigenous place names databases
- Incorporate AIATSIS place name thesaurus in Compound Gazetteer
- Digitisation of the AIATSIS language map
- Adding analytical layers to geographies
- Dual place names database

Activity 3.3 Aboriginal and Torres Strait Islander data 'rematriation' and archival services

---

[1] https://localcontexts.org/labels/traditional-knowledge-labels/

Drawing on the findings of Activity 3.2, one or more data assets will be selected as a case study for Aboriginal and Torres Strait Islander data rematriation practice. The term 'rematriation' was developed by Indigenous feminists as a restorative response to the patriarchal underpinnings of the term repatriation. Internationally, Indigenous Americans and Africans have employed the term when referring to the reclaiming of ancestral remains, spirituality, culture, knowledge and resources. In this context, the use of 'rematriation' embeds Indigenous worldviews and the principles of self-determination in the following activities.

A3.4 Key Objectives

This activity aims to practically embed Indigenous Knowledges, exchange and world views within the technical infrastructure of the project. The aim is to ensure that Aboriginal and Torres Strait Islander peoples are the primary beneficiaries of the outputs of the work. Over time, this practice will further embed the Aboriginal and Torres Strait Islander Data Governance principles across the national data ecosystem.

A3.4 Key Deliverables

1. Catalogue extensions to support the more detailed description of these particular data assets.
2. Thesaurus and ontology extensions to support reuse for other similar data assets.
3. Commons service extensions to support using the new National Data Assets and future similar data assets in ways useful to indigenous communities.

Stream 3. Milestones and deliverables

Indicative Dates	Milestone	Deliverables
2021: Q4	1.1 Stream Establishment Activities	Establishment of Governance Groups Recruitment for key Stream 3 project roles Assignment of Stream 3 Activity leaders Development of Stream 3 Implementation Plan
2022: Q1/2	3.2 Metadata Activities	Undertake Project Activity 3.1 - TK & BC Labels activities

2022: Q3/4	3.3 Case Studies	Establish and implement case study activities Support to build communities of practice Share progress and outcomes
2023: Q1/2	3.4 Data rematriation & archival services	Refinement and implementation of data rematriation for selected data assets
2023: Q3/4	3.5 Extension activities	▪ Widely share progress related to National core data assets ▪ Potential extension of training packages delivered in Stream 2, tailored to Stream 3.

Stream 3. Key Partners & Stakeholders

University of Melbourne

Australian National University

CSIRO

AIATSIS

Queensland University of Technology

8. Integration Component

Describe, in detail, elements of your proposal that will utilise RDC wide developments to be supported by the ARDC Integration Component.

1. **Access, Authentication and Governance**
   - Enable research groups and communities that do not currently have AAF credentials to access LDaCA.
   - Implement AAF authentication as the preferred access model.
   - This includes governance around the FAIR and CARE principles.
2. **Consultation Phase**
   a. Shared Model for engagement with communities.
   b. Develop evidenced based user for modelling services.
3. **Compute: HPC and GPUs**

   To be able to support the use of machine learning/AI in text analytics.
4. **Skills and Training**
   a. Project specific training
   b. Scope of what ARDC training exists
5. **HASS RDC Roadmap**

   The teams believe that the current work will only establish the potential need for further work. A shared work package might include the need to develop road maps for both components.

   a. **HASS RDC Roadmap**
      i. Governance framework
      ii. Legal and cultural data access framework
      iii. Technical architecture
      iv. Building communities

   b. **Indigenous Research Roadmap**
      i. The application of Indigenous governance frameworks
      ii. Legal/legislative requirements around data
      iii. Technical architecture and the ongoing development of catalogue
      iv. Indigenous Research Capabilities
6. **Legal requirements for data management and infrastructure**

   This work package would begin the conversation across government around the legal requirements of data access and use.

9. Outcome and Impact

Use cases: Policy Ready Data

The National Data Asset case study will be identified during Activities 2.1 and 2.2. The Case Study will be selected according to its potential for broad benefit to Aboriginal and Torres Strait Islander communities nationally with regard to both its significance in scope and purpose.

a. Thesaurus and/or ontology extensions to support the more detailed description of the selected data assets.

b. Catalogue content and interoperability extensions to support reuse for other data assets and other contexts.

c. Commons service extensions to support using the new National Data Assets and future similar data assets in ways useful to Indigenous communities.

d. Mobilise data and put it to use for advocacy, and for policy development and evaluation.

e. Consult on priorities, which could include:

- Public finance and the Indigenous economy
- Deaths and injuries in custody
- Child protection
- Indigenous health and well-being

Use cases: Orphaned datasets and preservation of personal data

Use cases of orphaned datasets will be developed, undertaken in a process similarly to that of PARADESIC, where ownership of the datasets are retained by the data owner/custodian, but discoverable information about the datasets are collected for the National Aboriginal and Torres Strait Islander Research Data Catalogue and/or Core National Indigenous Data Assets. Examples may include:

- National Trachoma and Eye Health Program
- Northern Territory Aboriginal Population Record
- Native title records
- Koori Health Research Database
- Native censuses

NB. ARDC together with the partners will implement a process for reporting on usage and outcomes of the infrastructure we build together. Reporting on impact allows us to demonstrate the value of the programs, activities and services to the commonwealth, states, partners, research community, and broader society. Partners are required to support the ARDC in implementing an impact reporting framework .

10. Governance

Our Project Lead, Langton, and Senior Research Associates will ensure the vision of the Indigenous Research Capabilities Project is achieved by establishing strong and dynamic governance arrangements that ensure an inclusive, transparent, equitable and collaborative approach to achieving our commitment to building national Indigenous research capability.

Project Governance will comprise the following collectives: (i) Executive Project Governance Committee (ii) Operational Project Groups within the key Activity.

1. **Executive Project Governance Committee**

The Executive Project Governance Committee will have a membership comprising of:

   a. IDN Steering Committee members

   b. Representatives from each of the Partner Organisations/Groups

   c. Aboriginal and Torres Strait Islander key stakeholders

   d. ARDC HASS RDC Program Manager

The Executive Projective Governance Committee will meet quarterly.

2. **Operational Activity Stream Groups**

Each of the Project Activity Streams will form a Operational Activity Stream Group to facilitate the development and timely implementation of activities within each of the three Streams. These groups will have a membership of:

   a. Project Team Leaders

   b. Core Activity Stream Staff

   c. Representatives from Core Stream Partners

The OASGs will report to the EPGC on all activities on a quarterly basis updating on project implementation activities and project risk management.