Melbourne School of Population and Global Health

Indigenous Studies Unit

Road Map v1

Indigenous Data Network

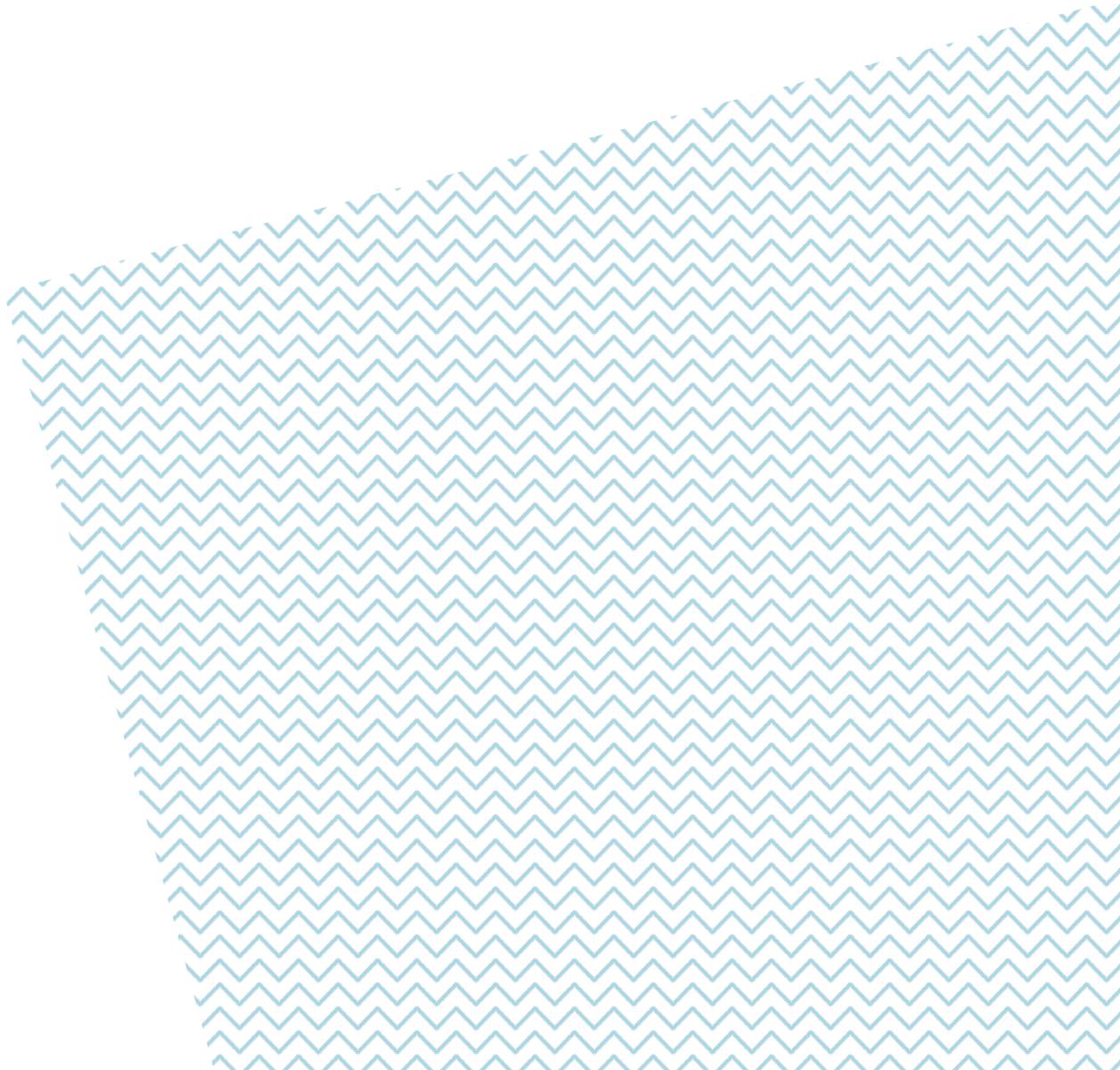1 March 2019

Table of Contents

An Australian Indigenous Data Network

Addressing the complex issues of disadvantage for Aboriginal and Torres Strait Islander people requires relevant high-quality data. Such data provides the evidence-base for the development, implementation and evaluation of effective policy solutions at local, state/territory and national levels. As the Prime Minister's 2018 *Closing the Gap Report* stressed, building an evidence base is key to success (Commonwealth of Australia, p. 12). However, the lack of reliable and consistent data for Indigenous Australians results in a paucity of evidence-based Indigenous policy-making.

It is of equal importance that local Indigenous communities are empowered 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. This sentiment is echoed in the *Empowered Communities Empowered People Design Report* (2015, p. 11) which emphasises that better outcomes for Indigenous communities are dependent on the level of autonomy and control that Indigenous people have in the decision-making process.

The Indigenous Data Network will assist communities to make informed decisions by equipping them with the knowledge, skills and resources to best manage and utilise their data.

Several Indigenous communities and organisations have already made significant progress in marshalling data about themselves within a framework of Indigenous empowerment (Wunan Foundation 2015). Examples of community approaches to data collection and interpretation were presented at the Indigenous Data Sovereignty Symposium, held in October 2017 (Indigenous Studies Unit 2018). The role of the Data Network will be to build on these strengths and to provide technical assistance and resources to help other groups to follow that path

The Data Network will also take a lead role in negotiating agreements with government agencies to make relevant data available to communities, organisations and researchers.

As the Productivity Commission's Report into Data Availability and Use (2017, p. 24) says: *"Governments across Australia hold enormous amounts of data, but mostly lag behind other comparable economies in beneficially using data beyond the purposes for which it was initially collected, or allowing others access to do so."* Similarly, there is a wide variety of data about Indigenous people and communities held by research institutions and individual researchers which should be available to the communities and other researchers. The Network will increase visibility and accessibility of this data, and encourage transparency and sharing.

Vision

The Indigenous Data Network will assist Indigenous communities in developing the technical capability and resources to enable them to manage their data for community advancement. By strengthening communities' agency in their data, the network will empower them to make informed decisions about their own development.

The Network will bring together community, university, government, non-government and private sector partners to create an ongoing community of practice to address a range of shared concerns, such as:

- Identifying best practice in community data collection, management and access

- Assisting Indigenous communities to apply best practice in data management by providing technical and educational resources
- Developing specific strategies and approaches to make better use of data over which Aboriginal and Torres Strait Islander people have ownership
- Creating a directory of databases to increase awareness of existing data sets and how to access them
- Integrating and archiving Indigenous datasets and preventing the orphaning of important datasets which would be detrimental to communities
- Negotiating with government and non-government organisations to ensure data activities are aligned with Indigenous priorities, and that data collected is available for sharing under appropriate conditions
- Working with the Indigenous Research Exchange to develop guidelines and best-practice case studies for research and data analysis in evaluation to improve Indigenous outcomes
- Coordinating educational programs to ensure the development of a critical mass of Aboriginal and Torres Strait Islander people with expertise in the data sciences
- Developing panels of experts who can provide advice and assistance on data issues.

The University of Melbourne will provide an administrative hub for the Indigenous Data Network. It is well positioned to do so, given its breadth of experience in Indigenous scholarship and in digital social sciences and humanities and health informatics, its existing capabilities in projects such as AURIN (the Australian Urban Research Infrastructure Network), PARADISEC (the Pacific And Regional Archive for Digital Sources in Endangered Cultures) and the eScholarship Research Centre, and its longstanding working relationships with AIATSIS (the Australian Institute of Aboriginal and Torres Strait Islander Studies and NCRIS (the National Collaborative Research Infrastructure Strategy).

Background: Indigenous Data Sovereignty

In October 2017, the Indigenous Studies Unit in the University of Melbourne's School of Population and Global Health convened an Indigenous Data Sovereignty Symposium, bringing together 115 people from Aboriginal communities, research institutes, Aboriginal corporations, influential scholars, and state and federal government representatives (Indigenous Studies Unit 2018, p. 4). Delegates discussed how Indigenous data is used and reported, and the right of Indigenous people to control data about themselves, their culture and their land in terms not only of collection, analysis and reporting but also of governance, ownership, access and consent.

Following the resolution of the Symposium to establish an Indigenous Data Network, a Steering Committee was established composed of prominent and experienced researchers in Indigenous data projects and leaders from universities, research institutes and government agencies (see Appendix 1). The Network is committed to ensuring better outcomes for Indigenous communities by assisting them to develop their capacity in all aspects of the data management life cycle, including identifying best practice, leveraging available data, building new data resources, and ensuring the development of Indigenous data scientists.

Background: Community Data Initiatives

Six case studies were presented at the Symposium of projects that engage with Indigenous data collection, community priorities and local analysis:

- Maranguka Community Hub, Just Reinvest, Bourke, New South Wales;
- Yawuru Knowledge and Wellbeing Project, Broome, Western Australia;
- Ngemba Data Research Hub, Brewarrina, New South Wales;
- Algabonyah Data Unit, Kaiela Institute, Shepparton, Victoria;
- Wadeye Museum and Thamarrurr Development Corporation, Wadeye, Northern Territory;
- National Empowerment Project: Development and Implementation of the Queensland NEP Pilot Sites in Kuranda and Cherbourg, Queensland.

These case studies provided important examples of projects that use community consultation, participation and ownership in their data management to achieve better results for Indigenous communities. Below we describe one of them, the Algabonya Data Unit.

Case study: Algabonyah Data Unit, Kaiela Institute

The Algabonyah Data Unit (ADU) is an example of Indigenous data self-determination and provides practical leadership in this nascent policy area. It was established by the Kaiela Institute in Shepparton to maintain and refine the continuity of data collection activities in Shepparton started by the Koori Resource Information Centre, the Rumbalara Football and Netball Club; and the Academy of Sport Health & Education and Strategic Policy and Planning Unit (University of Melbourne). The ADU is an important part of the implementation of the Empowered Communities vision in the Goulburn Valley region. It aims to enable access, analysis and interpretation of locally relevant data to address community priorities; provide an evidence base to inform the development of programs and policy; and to develop a structure to support the Algabonyah Community Cabinet processes.

The Algabonyah Data Unit aims to:

- Secure access to and analysis and reporting of relevant data from Census, administrative and statutory datasets
- Collect local data, for example through a local Community Census
- Develop indicators of wellbeing and social impact that reflect community organisations' specific role and aims
- Provide data analysis support for community-controlled organisations
- Provide support for local program evaluation
- Be the first port of call for proposed research projects that involve the local Aboriginal community

The anticipated benefits include:

- Building the capacity of local Aboriginal people to run, own, control and manage their own data.

- Identifying strengths and weaknesses of service models, including program evaluation.

- Over time it would allow for building governance models that would support the generation of bespoke data sets that can respond to community data needs, and have the capacity to access raw data from partnering organisations.

- Providing the infrastructure and workforce capacity to support organisations in their own data management and analytical needs.

(Indigenous Studies Unit 2018, p. 10-11)

Background: Orphan datasets

Many valuable Indigenous datasets, collected in the course of specific projects, are in danger of being lost because of the lack of ongoing funds for preservation and curation. The Network refers to these endangered datasets as 'orphan datasets'. The Steering Committee has already identified several datasets as endangered, including the following high-priority examples:

- the Koori Health Research Database (The University of Melbourne),
- Sandra Smith Archive (Museum of Victoria and the University of Melbourne),
- Aboriginal and Torres Strait Islander Data Archive (Australian National University),
- Victorian Aboriginal Child Mortality Study dataset (Victorian Department of Health),
- WA Linked Health Dataset (Western Australian Department of Health)
- Barwick Archive (State Library of Victoria)
- Native Title database
- National Trachoma and Eye Health Program records.

There are many examples of orphan datasets with similar significance for Indigenous data governance and Indigenous public health held at the Universities, libraries, archives and elsewhere. These will be investigated and prioritised and a management plan produced to ensure they are preserved and made available for re-use. Below we describe one example, the Koori Health Research Database.

Case study: Koori Health Research Database and the Sandra Smith Collection

The Koori Health Research Database (KHRD) began in 1999 as a partnership between the Bunjilaka Aboriginal Cultural Centre at Museum Victoria and the Onemda Koori Health and Community Development Unit at the University of Melbourne. Its purpose was to reconstitute the Aboriginal population of Victoria, using family histories, genealogies, civil registrations, and other historical records of the colonising state. The KHRD is unique internationally as a cradle-to-grave dataset documenting from the mid-nineteenth century the demography and health of an Indigenous people under colonisation. It is held in a customised system called Yggdrasil, developed by Sandra Silcot and Dr Len Smith with funding from the Australian National Data Service (ANDS), stored on the University of Melbourne's Nectar server.

The KHRD was built from a family tree database constructed for the Victorian Koori communities by Sandra Smith, a research officer at the Bunjilaka Centre. Sandra's family histories identified individuals and their family connections, and the University of Melbourne and the Australian Research Council funded the supporting documentation and research. The Sandra Smith Collection consists of 53 unaccessioned boxes of family files stored in the offices of Bunjilaka at Museum Victoria, containing all the paper files of research notes and birth, death and marriage certificates.

As is often the case, the KHRD database and the family history records on which it is based form a single collection, and need to be curated as such. The paper records provide the context without which the database cannot be properly understood or interpreted. Moreover, it is a living collection, providing a framework for the incorporation of new information as it becomes available.

Because the software is outdated and the physical records are inadequately curated and insecurely managed, the dataset is at high risk of being compromised and ceasing to be of relevance to Indigenous and research needs. Placing the KHRD on a sustainable basis will involve redeveloping and re-factoring the out-of-date software on a sustainable platform, and indexing and curating the physical archive to ensure its safe keeping and managed access by other researchers.

Roadmap Framework

The Indigenous Data Network (IDN) will operate three forms of support oriented to two general types of project partner, depending on organisational need assessed on a case-by-case basis. Forms of support will include technical facilitation, training, and governance, which will be differentially oriented to local organisations and government organisation respectively. This operational model gives rise to a reiterative partner-identification and service-delivery timeline, illustrated at Figure 1Figure 1.
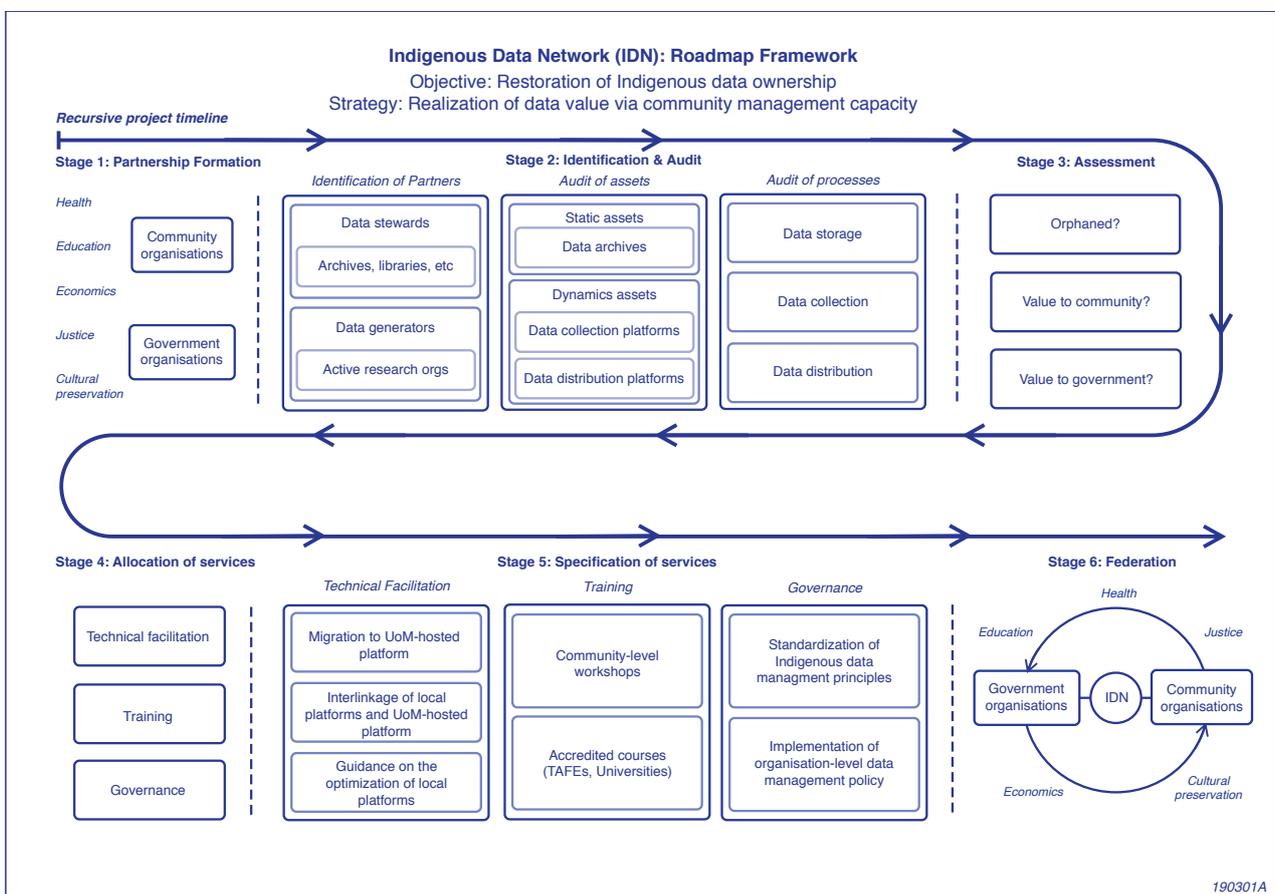

Figure 1: IDN Roadmap (see Appendix 2 for full-scale image).

Roadmap Framework: Identification of Partners, Audit, and Assessment

The first three stages in this reiterative framework involve the formation of partnerships, identification of partners according to an audit of the data services in which they are engaged, and assessment of partnerships by value status.

Partnerships will be formed across the health, education, economic, justice, and cultural preservation sectors, with both community and government organisations. As a first step to assessing the value status of the data services provided by partners, an audit of partners' data assets and processes will be conducted. Partner organisations may be classified into two broad groups according to whether they perform data stewardship or data generation activities, including data storage, data collection, and/or data distribution.

The assessment process will involve allocation of a status value to the services provided by each partner organisation, according to whether corresponding data assets are either or both orphaned, of value to Indigenous communities, and/or of value to government. Priority will be given to partner organisations storing orphaned data sets, and to partner organisation that are providing all three classes of service.

Roadmap Framework: Allocation and Specification of Services, Federation

The final three stages of the reiterative roadmap framework involve the allocation of specified services to partner organisation according the identification, audit, and assessment stages described above, followed by the aggregation of partnerships into a national federation.

Services provided by the IDN to partner organisations may be grouped in to three broad classes, including technical facilitation, training, and governance support. Each class of services will then be further specified according to the audit and assessment stages conducted beforehand, factoring in the requirements articulated by partner organisations themselves.

For example, whereas a government organisation may have a well-established technical capacity, it may be lacking a governance framework for the data services they provide. Conversely, and community organisation may have highly developed governance structures in place for managing the storage and distribution of data, yet lack the technical infrastructure to properly facilitate it.

Aggregation of each new partner organisation into a national federation will occur concurrently with each iteration of the roadmap process. Ultimately, the quality, efficacy, and efficiency of service delivery are expected to improve for each service organisation across all sectors relevant to Indigenous Australians.

Key priorities

The first priority for the Network will be to establish a working relationship with AIATSIS and other partners, and develop an overview of the domain, in order to answer basic questions such as:

- What are the key organisations that need to be involved? (e.**g.** land councils, community councils, medical and legal services)
- What data do they currently have access to? What data is needed?
- Where are the models for best practice of the data management life cycle?
- Where are the gaps in available data? How can they be filled?
- What expertise and technical resources are required?
- What relationships need to be established with other organisations active in the field? (e.g. NCRIS, ARDC, Indigenous Research Exchange, Indigenous STEM Network)

- What research activities need to be undertaken to support the mission of the Network?

Further actions to be undertaken are as follows:

- A series of consultations will be undertaken, leading to a detailed plan of action including resource requirements

- Expert panels will be established to provide advice and assistance on specific issues

- A best practice governance framework will be established for the management of the Indigenous data life cycle, including the development of ethical and access guidelines

- A digital media profile will be established, including a web site, newsletter, blog and social media presence

- Cloud resources and communication facilities will be mobilised to provide off-site backup, data management and software resources, and other crucial digital infrastructure as required

- In collaboration with data.gov.au and the Australian Research Data Commons, a public directory of datasets will be created, so communities know what datasets exist and how to access them

- Advice will be sought from data experts on the technical and administrative issues involved in establishing federated databases, drawing together information from federal, state, and non-government sources

- Priority will be given to identifying sources of data on Indigenous economic participation, which has been identified as a deficiency in the Closing the Gap framework. Key experts will be identified and consulted, and a specific strategy developed.

- In conjunction with national and international collaborators, a research program will be established aimed at furthering understanding the basis of Indigenous identification in databases, especially the census which is basic to all other sources.

Partners

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

The Australian Institute of Aboriginal and Torres Strait Islander Studies (AIATSIS) is a research, collections and publishing organisation within the portfolio of the Department of Prime Minister and Cabinet. It is the major repository for documents and audio, visual and digital materials collected in relation to Aboriginal and Torres Strait Islander cultures and societies. The CEO of AIATSIS, Mr Craig Ritchie, is a member of the Indigenous Data Network's Steering Committee.

With its high level of capability in archiving and its centrality in the nation as the statutory body with the responsibility to provide research services, maintain a library and archives of materials, and increasingly digitise its collections and maintain them on a variety of platforms, AIATSIS is a key partner in the Indigenous Data Network. The AIATSIS digitisation program preserves important Indigenous cultural items in digital form so they can be used as a research resource to be shared with all Australians and the rest of the world. AIATSIS has recently established an Indigenous Research Exchange, which focuses on research translation, and may provide some resources to the Network.

As a Commonwealth Statutory Authority, AIATSIS will play a lead role in data integration and in encouraging the collaboration of Indigenous, governmental and research entities in sharing data.

Kaiela Institute, Shepparton

The Kaiela Institute is a funded backbone organisation under the Commonwealth Government's Empowered Communities initiative and is Secretariat to the Algabonyah Community Cabinet. The Kaiela Institute is located on Yorta Yorta country, in the Goulburn Murray region of Victoria. The Kaiela Institute is already a major Indigenous organisation partner of the University of Melbourne and has recently approved a formal partnership with the Indigenous Data Network, and they will be a crucial advisor on matters of Indigenous data, and future host of capability building workshops. The Algabonyah Data Unit based at the Kaiela Institute (see Case studies), will be important to this partnership. Mr Paul Briggs is the Executive Chair of the Kaiela Institute

Australian National University (ANU)

ANU will assist in the non-accredited capability building courses for Indigenous scholars and community workers. They have connections with the Yawuru and Cowra communities. These connections will provide potential presenters, attendees and host organisations for the short courses. Collaborators include Professor Richard Baker, Pro-Vice Chancellor (University Experience); Roxanne Missingham, University Librarian; Associate Professor Edith Gray is Head and Dr Len Smith is an Academic Visitor at the School of Demography; Dr Nicholas Biddle, Dr Mandy Yap and Professor John Taylor are from the Centre for Aboriginal Economic Policy Research.

Australian Data Archive (ADA)

The Australian Data Archive (ADA), located at ANU, provides a national service for the collection and preservation of digital research data. Dr Steve McEachern is Director.

Australian Institute of Health and Welfare (AIHW)

AIHW is the leading health and welfare statistics agency in Australia. Dr Fadwa Al-Yaman is responsible for the Indigenous & Maternal Health Group. Dr Al-Yaman has expressed her willingness to assist in the capability building courses and provide in-kind support from the Institute.

National Collaborative Research Infrastructure Strategy (NCRIS) and Australian Research Data Commons (ARDC)

We will work with the National Collaborative Research Infrastructure Strategy (NCRIS) in the Commonwealth Department of education, which is developing the major research infrastructure to facilitate collaboration between the research sector, industry and government in Australia and internationally. AIATSIS has a central role in the development of NCRIS engagement with Indigenous Australians. By collaborating with the institutions in our Indigenous Data Network and Indigenous corporations and communities, together we will have the capacity and expertise to design and develop data systems to enable relevant, useful and innovative research to improve outcomes for Indigenous Australians.

The Australian Research Data Commons (ARDC) is funded by (NCRIS), and is a potential source of funding for the Network. Dr Adrian Burton is Director, Data, Policy & Publication Services. Dr Burton has expressed ARDC's willingness to provide training and in-kind technical support through their existing services.

South Australian Health and Medical Research Institute (SAHMRI)

The South Australian Health and Medical Research Institute (SAHMRI) is South Australia's first independent health and medical research institute. Dr Ward is the Head of the Infectious Diseases Research Program – Aboriginal Health, at SAHMRI is a member of the Network's Steering Committee.

Department of Prime Minister and Cabinet (PM&C)

PM&C is responsible for the Indigenous Advancement Strategy and with COAG for the Closing the Gap program. Professor Ian Anderson is Deputy Secretary, Indigenous Affairs and Dr Shane Johnson and Dr Joanna Abhayaratna are from the Evaluation Branch. We have asked for financial support from the Department.

Empowered Communities

We will create a formal partnership with the Empowered Communities group, working with Sean Gordon and Dianne Hawgood.

Maiam nayri Wingara Indigenous Data Sovereignty Collective

The Collective is the Australian arm of the international movement for Indigenous Data Sovereignty. We have invited the Collective to participate in the Network, and will continue to inform them of our intentions and progress.

Technical Services

The University of Melbourne will provide the base technical platforms for the Indigenous Data Network. The University has invested heavily in state-of-the-art research computing platforms and services that are capable of supporting complex data-driven projects such as the Network. At no cost, researchers can leverage a large-scale self-service research cloud, Nectar, for which the University is the lead agent within the Australia Research Data Commons. The research cloud is complemented by a petabyte data storage facility, complete with tools to manage research metadata, longitudinal data changes, secure sharing and collaboration, and fine-grained access control. These facilities are operated within the University's data centres under strict security and privacy protocols.

The University's Research Platform Services group will provide the enabling link between the Network and the research computing platforms, providing expert advice and consultation to solve complex data problems. Together, these systems have been used to support the Koori Health Research Database, Founders & Survivors Project, and the First 1000 Days Project. The University's commitment to research data management, processing, and analysis is demonstrated by its five-year Petascale Campus and Digital Onramp projects, currently underway. These substantial initiatives support research at the forefront of new technologies, accelerating the rate of scientific discovery and research translation.

Further technical support will be sought from NCRIS and ARDC as required.

References

1. Indigenous Studies Unit 2018, *Report on the Indigenous Data Sovereignty Symposium: The importance of data sovereignty for communities*, University of Melbourne, Melbourne.

2. Commonwealth of Australia 2018. *Closing the Gap Prime Minister's Report 2018*, Department of the Prime Minister and Cabinet, Canberra.

3. Productivity Commission 2017, *Data Availability and Use: Overview & Recommendations*, Report No. 82, Canberra.

4. Wunan Foundation 2015, *Empowered Communities: Empowered Peoples, Design Report*, Canberra.

Appendices

Appendix 1 – Indigenous Data Network Steering Committee

Professor Sandra Eades (Committee Chair)
Associate Dean (Indigenous), Faculty of Medicine, Dentistry and Health Sciences (MDHS), University of Melbourne

Professor Marcia Langton
Associate Provost, Redmond Barry Distinguished Professor, University of Melbourne

Professor Shaun Ewen
Pro Vice Chancellor (Indigenous), Foundation Director Melbourne Poche Centre, University of Melbourne

Dr Len Smith
Visiting Scholar, Australian National University

Mr Craig Ritchie
CEO, Australian Institute of Aboriginal and Torres Strait Islander Studies (AIATSIS)

Associate Professor James Ward
Head of Infectious Diseases Research Program - Aboriginal Health, South Australian Health and Medical Research Institute (SAHMRI)

Professor Janet McCalman
Redmond Barry Distinguished Professor, Melbourne School of Population and Global Health, University of Melbourne

Dr Kristen Smith
Research Fellow, Indigenous Studies Unit, Melbourne School of Population and Global Health, University of Melbourne

Dr Lyndon Ormond-Parker
Research Fellow, Indigenous Studies Unit, Melbourne School of Population and Global Health, University of Melbourne

Dr Kalinda Griffiths
Scientia Fellow, Centre for Big Data Research in Health, University of New South Wales

Dr James Rose
Senior Research Fellow, Indigenous Studies Unit, Melbourne School of Population and Global Health, University of Melbourne

Appendix 2 – Roadmap Framework

Indigenous Data Network (IDN): Roadmap Framework
Objective: Restoration of Indigenous data ownership
Strategy: Realization of data value via community management capacity

190301A

Recursive project timeline

Stage 1: Partnership Formation

Identification of Partners

- Data stewards
  - Archives, libraries, etc
- Data generators
  - Active research orgs

- Health
- Education
- Economics
- Justice
- Cultural preservation

Community organisations

Government organisations

Stage 2: Identification & Audit

Audit of assets

- Static assets
  - Data archives
- Dynamics assets
  - Data collection platforms
  - Data distribution platforms

Audit of processes

- Data storage
- Data collection
- Data distribution

Stage 3: Assessment

- Orphaned?
- Value to community?
- Value to government?

Stage 4: Allocation of services

- Technical facilitation
- Training
- Governance

Stage 5: Specification of services

Technical Facilitation

- Migration to UoM-hosted platform
- Interlinkage of local platforms and UoM-hosted platform
- Guidance on the optimization of local platforms

Training

- Community-level workshops
- Accredited courses (TAFEs, Universities)

Governance

- Standardization of Indigenous data managment principles
- Implementation of organisation-level data management policy

Stage 6: Federation

- Justice
- Health
- Education
- Economics
- Cultural preservation

Community organisations

IDN

Government organisations