Life Course Centre

WORKING PAPER SERIES

No. 2026-27

October 2026

Advancing Health and Wellbeing Across the Life Course through System-Wide Development of Data Research Infrastructure

The Case of Australia

Mark Western

Marcel van Kints

An updated version of this paper has been published as Western, M., & van Kints, M. (2026). Advancing Health and Wellbeing Across the Life Course Through System-Wide Development of Data Research Infrastructure: The Case of Australia. International Journal of Population Data Science, 11(3). https://doi.org/10.23889/ijpds.v11i3.3437

The Australian Research Council Centre of Excellence for Children and Families over the Life Course

Phone +61 7 3346 7477 **Email** lcc@uq.edu.au

lifecoursecentre.org.au

Australian Government
Australian Research Council

The University Of Queensland
AUSTRALIA

The University of Sydney

The University of Western Australia

The University of Melbourne

Research Summary

Why was the research done?

Population health and wellbeing are shaped by much more than medical care. Family circumstances and upbringing, education, employment, income, housing, access to services and community resources, all influence how people fare over their lives. However, the information to support research, policy and practice to enhance health and wellbeing is often collected and maintained in separate systems. Bringing relevant population data together carefully and securely could give researchers, policymakers and practitioners a clearer picture of how life circumstances affect health wellbeing over the life course and how to respond to advance health and wellbeing.

What were the key findings?

Australia has distinctive policy settings, institutional arrangements and partnerships linking universities, governments, data custodians and other organisations. These enabling conditions are an opportunity to develop research infrastructure that responds to current needs while also adapting to future questions.

A national data resource could support both longitudinal studies, which follow changes over time, and broader snapshots of the population at particular points in time. It could also help assess whether programs and services work, for whom and under what circumstances. More detailed analysis could improve the evidence available for communities that are often poorly represented in general population research, including groups that may benefit from targeted or more intensive support.

What does this mean for policy and practice?

A more coordinated national data system built on Australia's strong foundations, including partnerships between researchers, policymakers and communities, could support improved policymaking, implantation and service delivery, enhancing population health and wellbeing in Australia.

Citation

Western, M. C., & van Kints, M. (2026). 'Advancing Health and Wellbeing Across the Life Course Through System-Wide Development of Data Research Infrastructure: The Case of Australia', Life Course Centre Working Paper Series, 2026-27. Institute for Social Science Research, The University of Queensland.

The authors

Mark Western
The University of Queensland
Email: m.western@uq.edu.au
https://about.uq.edu.au/experts/909

Marcel van Kints
Australian Bureau of Statistics
Email: Marcel.van.kints@abs.gov.au

Acknowledgements/Funding Sources

This research was supported by the Australian Government through the Australian Research Council's Centre for Excellence for Children and Families Over the Life Course [Project ID CE200100025].

DISCLAIMER: The content of this Working Paper does not necessarily reflect the views and opinions of the Life Course Centre. Responsibility for any information and views expressed in this Working Paper lies entirely with the author(s).

This work is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License.



We acknowledge the Traditional Custodians of the lands on which we work and live across Australia.
We pay our respects to Elders past and present and recognise their continued connections
to land, sea and community.

Advancing Health and Wellbeing Across the Life Course Through System-Wide Development of Data Research Infrastructure: The Case of Australia.

Mark Western, Emeritus Professor of Sociology, Faculty of Humanities, Arts and Social Sciences, The University of Queensland. m.western@uq.edu.au

Marcel van Kints, General Manager, Statistical Production and Digital Services Division, Australian Bureau of Statistics. Marcel.van.kints@abs.gov.au

Abstract

In Australia a distinctive set of policy arrangements, institutional conditions, and partnerships between universities and other stakeholders is enabling a coordinated system-wide approach to developing research infrastructure for large-scale, whole population, research such as required to implement the WHO life course approach to health and wellbeing. Research infrastructure includes equipment and facilities, standardised data assets, research infrastructure expertise, specialist personnel, and services such as training and capacity building. This paper describes Australia's research infrastructure policy and funding environment and its national research infrastructure planning and development process. It also describes how Australian health and social science researchers have coordinated across universities and government to collectively plan and develop major research infrastructure projects capable of supporting the WHO life course approach and wellbeing.

Introduction

The WHO life course approach to health and wellbeing focuses on optimising healthy trajectories at every life stage for all population groups. Health trajectories connect life stages, and healthy trajectories and transitions during pregnancy, childhood and adolescence positively influence health and wellbeing in adulthood and later life [1, 2]. Advancing health and wellbeing over the life course requires a whole-of-system approach based on the social determinants of health. It also requires integrated multi-sectoral responses across governments, service providers and healthcare professionals. Health interventions and services need to reflect people's life stages, variable needs, capacities and circumstances, in life domains including health, labour, education, and participation in civil society.

Implementing this approach depends on research and evaluation that builds an evidence-base about how to optimise health and wellbeing for different populations at different life stages [2]. Primary research and evidence synthesis is crucial to understand social determinants of health and to identify and evaluate interventions and services against appropriate outcome measures. These research and evaluation tasks depend on comprehensive population data across life stages and life domains, such as education and the labour market, to support longitudinal and cross-domain research and evaluation [2]. The population data also need to support research and evaluation into new and novel interventions and whole-of-population and detailed subpopulation analyses, including for marginalised and vulnerable population cohorts who may require targeted and/or intensive programs, services and investments [1].

The research infrastructure for a population life course approach to health and wellbeing is an enduring coherent, integrated, national data system, accessible to researchers, policy makers and practitioners across government, research institutions and non-government sectors. The life course approach spans life domains beyond health and research infrastructure needs to be similarly broad in covering health and wellbeing, life domains such as employment, education and civil society, social determinants, and social environments. The data system needs to incorporate longitudinal and cross-sectional information on populations and subpopulations, services, health interventions, policies and programs. It should support monitoring and evaluation as well as research, with measures and indicators that are robust (reliable and valid) and comparable. Systems and data need to promote interoperability and accessibility, while also maintaining privacy, confidentiality, data integrity, public trust and accountability.

This paper describes Australia's "enabling conditions" for developing large-scale research infrastructure of this kind. Australia has an internationally distinctive set of policy arrangements, institutional conditions and partnerships between universities and other stakeholders, that support coordinated action towards a national data system to serve the public interest. These conditions allow research and policy communities to pursue coordinated action to develop national infrastructure which meets their current and anticipated future research needs. Science disciplines such as astronomy and physics have long recognised that leading-edge research and field development is impossible without large-scale research infrastructures such as telescopes, high performance computing and particle accelerators, but advanced research infrastructures are increasingly understood to be essential for frontier research and addressing societal challenges in all fields [1-3]. Australia's "enablers" for a national data system supporting a life course approach to health and wellbeing include its research policy and research infrastructure funding environment, federal legislation and frameworks around public sector data sharing and integration, and coordinated partnering across universities, government and other organisations to develop and pursue research infrastructure initiatives. These developments lay the groundwork for new advanced whole population studies in health, social science and cognate fields. This paper describes the Australian context and enabling conditions for research infrastructure development, outlines

some of the new infrastructure that is emerging, and concludes by discussing some of the challenges and opportunities arising from this process.

National Research Infrastructure Policy in Australia

In Australia, the term National Research Infrastructure ((NRI) has a specific meaning: "nationally significant assets, facilities, services, and associated expertise to support leading-edge research and innovation." [4] Research infrastructure includes equipment and facilities, data, expert personnel, and services like user-support and training. National research infrastructure is one "layer" in Australia's research infrastructure ecosystem alongside landmark and global infrastructure, which supports global research capability ((examples include the Australian Synchrotron and Square Kilometre Array), and institutional infrastructure managed within single organisations, such as universities.

More specifically for this paper, digital national research infrastructure includes: enabling platforms; tools for tasks such as data discovery, access and analysis; services such as researcher training in analytic methods or the use of infrastructure; a research infrastructure support workforce with associated expertise; curated data collections or data assets; and governance arrangements supporting accessibility, use, ethics, and community involvement. Publicly funded national research infrastructure should support open research and open user accessibility, while also safeguarding appropriate access and use, data confidentiality, privacy protection, and community and societal benefit. This case study focuses particularly on the institutional enabling conditions for Australia developing a national data system to support life course health and wellbeing for the Australian population. We do not attempt an exhaustive description of the entire data infrastructure landscape in Australia, or of all relevant data assets, rather we focus on the components relevant to the WHO life course framework. We also concentrate primarily on data infrastructure supporting the use of administrative public sector data with complete population coverage.

Linked public sector administrative data has several advantages as the foundation of a national data system for population health and wellbeing. Government administrative datasets typically have complete population coverage, enabling whole population and sub-group analyses and comparisons. Data integration across jurisdictions and areas of government supports analysis and evaluation across life course domains, life stages and health trajectories. Comprehensive, whole population data systems can also provide an underlying architecture for ongoing evaluation of health policies and interventions limiting the need for additional bespoke data collections.

However, relying exclusively on linked administrative data does not address all data requirements. Jurisdictional differences between states, territories and the national government in data governance, collection methods or measurement standards, may hinder the cross-jurisdictional linkage necessary for complete population coverage and comprehensive measurement. Administrative data is confined to information required to deliver policies or services; subjective information about circumstances and outcomes, or measures based on direct observation, may not be available. Gaps in policy and service delivery, in relation to subpopulations or aspects of health and wellbeing will be associated with data gaps, and finally variables relating to strengths-based approaches, including for priority populations, are seldom collected because they are not administratively required. Some of these issues can be resolved by legislative and regulatory frameworks, standards, and resourcing to support cross-jurisdictional data linkage. Linking administrative data selectively to bespoke data collections such as national longitudinal household and cohort surveys, and greater involvement of the

population and priority subpopulations in data governance and data collection oriented to societal and community objectives can also strengthen data systems.

Funding for National Research Infrastructure is provided primarily through the Australian Government's National Collaborative Research Infrastructure Strategy (NCRIS) funding program [5], one of the major federal funding schemes within the Australian Government's national expenditure on research and development. NRI strategy and implementation is coordinated under five yearly National Research Infrastructure Roadmaps, with the 2021 Roadmap being the most recent[1]. The 2026 Roadmap is currently being developed. NCRIS funding is guided by priorities and activities identified in NRI Roadmaps and made available through targeted national funding rounds every 1-2 years. Accessing federal NCRIS funding to establish and operate national research infrastructure requires co-funding from research infrastructure partners such as subnational state and territory governments, host organisations (often universities) for national facilities, and the private sector [6]. · Between 2018 and 2029 the Australian Government will have invested $4 billion in National Research Infrastructure, with NRI investments in NCRIS one element of overall federal spending on research and innovation [6].

Health and Wellbeing Research Infrastructure and the National Collaborative Research Infrastructure Strategy.

Although Australia has not had separate funding for stand-alone national research infrastructure to support a comprehensive life course approach to health and wellbeing, several NRI initiatives have laid relevant groundwork. The most direct NCRIS investment to support Australian health and wellbeing is in Data Connect Australia (DCA, formerly the Population Health Research Network), established in 2009 as a national collaboration hosted from the University of Western Australia to facilitate funding, coordination and the development of enabling infrastructure for the use of linked administrative data to enhance population wellbeing. DCA includes partners from Australian, state and territory governments, universities and research institutes. The network partners operate national and regional data linkage infrastructure comprising state, territory and national data linkage units, secure research environments and e-research tools and services to facilitate researcher access to linked datasets [7]. DCA partners hold administrative data collected during health services delivery (e.g. birth records, hospital admissions, prescriptions and medical benefits, aged care, child development and child protection data), clinical data including patient records, and clinical registry data collected for healthcare quality assessment [8]. DCA partner data is not primarily collected for research but for government planning, administration, monitoring, funding and health service delivery. DCA partners link administrative data to build datasets for research and make them available through a distributed trusted research environment. The network also manages governance and access to linked data and supports researchers through services including training, metadata development and data documentation, and support for ethical review.

DCA is an important element in a life course health and wellbeing data system, but incomplete in relation to social determinants of health, some health and wellbeing outcomes, and personal and other information which is not mandated for collection or necessary for healthcare administration or delivery of services. Under Australia's federal system, most health services are delivered and monitored by state and territory governments so data linkage can often

[1] This section draws upon the 2021 National Research Infrastructure Roadmap. The 2026 NRI Roadmap may modify this definition.

involve cross-jurisdictional linkage, which introduces complications relating to governance, common standards and data harmonisation.

A second stand-alone NCRIS facility relevant to a national data system for health and wellbeing is the Australian Urban Research Infrastructure network (AURIN) [9]. AURIN is national data infrastructure for e-research into urban, regional and spatial issues. AURIN provides researcher access and analytical tools access to spatially-structured datasets, that is datasets built around spatial units of analysis and observation (a statistical geography), rather than person-centred datasets. AURIN data spans health, transport, housing, the economy, land use, and demography and datasets are built primarily from integrated public sector data from administrative sources and national collections such as Australia's five yearly population census. AURIN and DCA are currently collaborating on the Australian Urban Health Indicators project which is creating new data assets for topics including Cancer Determinants, Heat Health Vulnerability and Urban Liveability and Health [10]. Introducing a spatial dimension rounds out a life course approach by incorporating factors associated with the natural, physical and built environments, and with spatially- bounded social processes such as community determinants of health and wellbeing.

Other personal and contextual information about life domains such as individuals' education, employment and social participation, in relevant institutional settings such as schools, families and workplaces, is not included in administrative public sector health services data unless required for policy or service delivery. Some of these data are collected and held by other government agencies (for instance Departments of Education, Employment, Social Services, the Australian Taxation Office and the Australian Bureau of Statistics) and through vehicles such as the Australian Census of Population and Housing. These data are also more usually the province of the social sciences. Unlike public health and urban and regional research, the social sciences have not received direct research infrastructure funding through the NCRIS program for dedicated social science research infrastructure. However, the 2016 and 2021 NRI Roadmaps acknowledge the research infrastructure needs of humanities, arts and social sciences [4].

The 2021 Roadmap contains a priority to support Humanities and Social Science (HASS) research infrastructure as national research infrastructure. Under the 2021 Roadmap, this has been realised by funding an existing NCRIS facility, the Australian Research Data Commons (ARDC), to develop research infrastructure for humanities, arts and social sciences. ARDC is a national research data infrastructure organisation supporting e-research. Its core projects include building thematic research data commons which integrate computing, storage and services with data assets, analysis platforms and tools. ARDC's HASS and Indigenous Research Data Commons (HASS-I) [11] includes people, data, platforms and tools. It uses shared technical standards, includes governance frameworks and processes, open-source tools to help researchers find and analyse data, and training facilities. Specific activities to develop the HASS-I data commons have included projects to enhance metadata for national datasets that support inequality research, harmonisation and data integration of "person-structured" and "space/place" structured datasets (an ARDC-AURIN collaboration) and development of an expanded national social science research infrastructure network linking different social science research infrastructures [11]. ARDC collaborates with university researchers, research infrastructure partners, and government data custodians to deliver this work.

AURIN, ARDC and DCA are platform services providing access and supporting researchers to work with predominantly public sector-linked datasets. Separately from these initiatives, the Australian Government is also building national whole population data assets and linkage systems from public sector administrative and census data to support research, policy

development, planning and evaluation. Several of these would form core elements of a national data system for health and wellbeing.

Australia's national Person Level Integrated Data Asset (PLIDA) is an individual level longitudinal integrated data asset based on linked information on healthcare, education, government payments, personal income tax and population information [12]. The Business Longitudinal Analytics Data Environment (BLADE) is a business centred longitudinal data asset supporting economic research [13]. It provides a comprehensive view of Australian businesses. There is also a Linked Employer and Employee dataset, which when linked with PLIDA and/or BLADE could support analyses of workplace and employment related determinants of health and wellbeing.

The National Health Data Hub (NHDH) is a linked data system of the Australian Institute of Health and Welfare. It brings together federal, state and territory government administrative datasets on health, welfare, disability and aged care which are curated and harmonised for research and analysis. It is longitudinal and person-centred [14].

Two other relevant national data assets are the national disability data asset (NDDA) and the Life Course Data Initiative (LCDI). The NDDA draws population information from every part of government [13]. While intended to understand and support the needs of people with disability the NDDA will contain de-identified information on all Australians. The NDDA is currently being delivered through the existing data integration systems of the Australian Institute of Health and Welfare and the Australian Bureau of Statistics as an interim measure and has already been used to examine outcomes for people with disability against Australia's Disability Strategy 2021-2031 [15, 16]. Finally in 2023-24, the Australian Government introduced a $200 million package in the federal Budget to address entrenched community disadvantage and improve child and family wellbeing. A key component is the Life Course Data Initiative pilot, undertaken by the Australian Bureau of Statistics, with the subnational Australian Capital Territory and South Australian governments, plus a university partner. The pilot data asset links the federal Person Level Integrated Data asset , PLIDA, with state and territory datasets from the Australian Capital Territory and South Australia [17]. Data linkages across jurisdictions and levels of the Australian federal system are necessary to capture internal migration, and because national policy responsibilities sit with the federal government, but service and program delivery in areas such as health and education, are state and territory responsibilities. The Life Course Data Initiative is in early stages with access confined to the initial participants in the data integration project. The pilot data asset is not sufficient to support a whole population approach to life course health and wellbeing and the Life Course Data Initiative will not serve a national data system without concrete timelines and governance pathways to national data integration beyond the initial jurisdictions, and researcher access beyond original participants.

These assets combine multiple single linked population datasets in a research-ready format with de-identified standard and customised longitudinal and cross-sectional data extracts available to researchers. Building data assets through population data linkage brings distributed information together, supports fine-grained subpopulation analyses, including of small or outlying groups, and is a cost-effective way to update and maintain data at different population and geographic scales. Whole population data assets linked to information on policy, program and service utilisation also create a rich evaluation data infrastructure that supports ongoing evaluation using experimental and quasi-experimental methods and overall health system performance monitoring.

In addition to data and platforms, a national data system for research and evaluation requires a strong governance framework that protects privacy, confidentiality and sensitive information, regulates access appropriately and supports public trust and public accountability in relation to

information holdings and their use [18]. Data linkage and data sharing of Australian Government data for research is legislated through the federal Data Availability and Transparency Act. An accompanying "DATA Scheme" supports data custodians, data service providers (typically authorised data linkage agencies) and users. The scheme has safeguards to protect privacy and manage user access [19]. The enabling research infrastructure includes secure trusted research environments, and privacy protecting access protocols, that allow qualified researchers to work on whole population data containing sensitive information, for approved research projects. The Data Act provides the legislative basis for federal public sector data sharing and release, and the Data Scheme operationalises this by accrediting data service providers (such as linkage authorities and trusted research environments) and organisational data users (such as universities). Researcher access is managed at the organisational level. Disclosure risk and privacy protection is managed through the Five Safes Framework [20]. This infrastructure is available alongside open data tools and basic microdata downloads.

In 2025 the Data Availability and Transparency Act underwent a statutory review [21]. The review recommended continuing the Act with modifications to facilitate greater public sector data sharing and access. The recommendations involve changing regulatory settings and the functions and powers of the National Data Commissioner (the regulator), simplifying accreditation under the Act for data sharing and use, and expanding access to users outside university and government, for broader purposes. It also recommended further investment in Australia's national data ecosystem to promote better outcomes for data custodians, providers and data users [21]. The Australian Government is considering its response to the review.

Australian social science researchers have responded to opportunities to develop national research infrastructure by coming together to develop the Decadal Plan for Social Science Research Infrastructure 2024-2033 [22]. The Decadal Plan was developed by the Academy of Social Sciences in Australia, Australia's learned academy for social science, overseen by a Steering Group of senior researchers. An Expert Working Group of 90 researchers and data and infrastructure organisations undertook development work and consultations across university, government, civil society and industry, producing a national blueprint for research infrastructure over the next 10 years. The central goals of the Decadal Plan are:

- A coordinated Australian Research Infrastructure Ecosystem for the Social Sciences (ARIESS)that caters for diverse disciplines, methodologies and communities and supports goals and aspirations of Aboriginal and Torres Strait Islander Peoples;
- Social science data and analytics are easily found and reused by researchers across Australia;
- Robust investment in research capabilities and infrastructure so that Australia has world-class assets, instruments and a research and research infrastructure workforce.

The Decadal Plan also emphasises that social science research infrastructure needs to support the range of disciplines, methodologies, data types, technical and skill levels across Australian social science, must support Indigenous self-determination in research planning, and address Indigenous Data Sovereignty and Indigenous Data Governance [23, 24]. Indigenous Data Sovereignty refers to the rights of Indigenous people to exercise ownership over data about Indigenous peoples while Indigenous Data Governance covers Indigenous people's rights over Indigenous data collection, access and use.

Australian social science researchers come from disciplines employing diverse methodologies and types of data. Technical and skill levels also vary considerably with researchers capable of analysing unit-record longitudinal integrated data inside secure research environments being only a subset of the researchers who may want to use some of the assets and capabilities in a national system. Australia has several initiatives underway to enhance access by researchers

with diverse objectives and skills. The Australian Bureau of Statistics has an experimental "snapshot of the Australian population" built from integrated PLIDA data available through its open access TableBuilder web product [25]. This data infrastructure allows researchers to build customised tables from underlying microdata through a simple web interface. TableBuilder also provides access to data from various national surveys on topics in health and wellbeing including for Aboriginal and Torres Strait Islanders, and on topics such as nutrition and physical activity, and disability and ageing [26]. The HASS and Indigenous Research Data Commons includes a project to establish the Social Science Research Infrastructure Network [27] linked to the goals of the Decadal Plan. The Social Science Research Infrastructure Network is NCRIS funded, via the Australian Research Data Commons, to support the use of linked government data by Australian researchers. It includes work programs to improve the research useability of the existing PLIDA data asset, support data integration between health and social data assets, develop a national data skills framework for social sciences and an associated training program, and develop guidelines for users of government administrative data relating to Indigenous people [28].

The 2021 National Research Infrastructure Roadmap also highlights the importance of implementing Indigenous Data Governance and the CARE Principles [collective benefit, authority to control, responsibility, ethics] [29] for data relating to Aboriginal and Torres Strait Islander peoples (Figure 1). The CARE Principles are a set of global principles for data governance that support Indigenous self-determination by enabling Indigenous people to use Indigenous data for their own purposes. The CARE Principles also recognise and are consistent with FAIR (findable, accessible, interoperable and reusable) principles of scientific data management and governance [30] which are intended to support data discoverability, accessibility, integration and reuse. While the CARE Principles are specific to Indigenous Data Governance, the high level CARE framework supporting collective benefits, the rights and interests of affected parties in relation to data, responsibility around relationships between all parties associated with data (data subjects, data stewards, data users), and a commitment to ethical action, is a model for data stewardship and data governance for all data pertaining to humans.

Together, FAIR and CARE Principles provide guidance for a governance framework for a national data system that supports data production, access and reuse, for collective benefit in the interests of those the data relate to. The implementation of the CARE Principles across Australian public sector administrative data is still in very early stages. The HASS and Indigenous Research Data Commons includes a project to build Indigenous research capability framed by Indigenous Data Governance and FAIR and CARE Principles [31]. The Social Science Research Infrastructure Decadal Plan also foregrounds secure systems and ethical and responsible governance, infrastructure and research practice. It also commits to open research and open science principles, equitable access, and training and skill development for researchers and for the research infrastructure workforce [22].

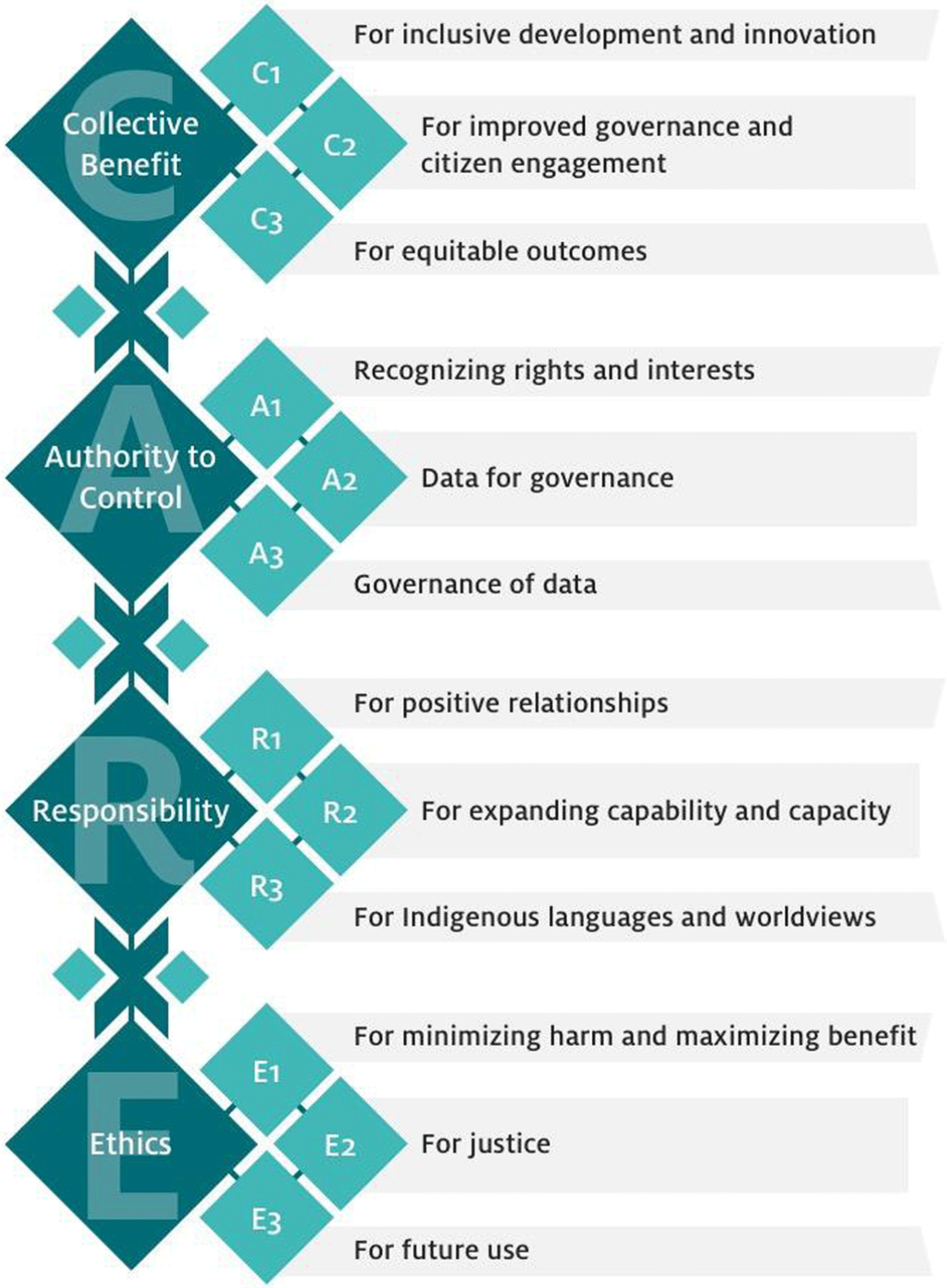


Figure 1 The CARE Principles for Indigenous Data Governance [23]

Towards a National Data System Advancing a Life Course Approach to Health and Wellbeing

Australia's emerging system of mature social science research infrastructure includes high quality population data assets, platforms, tools and services for data integration, data discovery and reuse. There are emerging governance frameworks and coordinated planning that involves data custodians and providers, research infrastructure organisations, researchers and societal stakeholders. Australia also has publicly available high quality national longitudinal and cross-sectional survey studies, not mentioned here. The system is underpinned by open and responsible research principles, equitable access, and commitments to co-design with researchers and civic communities. These principles are not yet uniformly implemented across the emerging data ecosystem but are embedded in relevant policy and funding frameworks, including NCRIS.

These are strong foundations for a cohesive national data system but to this point many elements across government and the research sector have developed independently. There is also fragmentation across data assets and systems, such as PLIDA, the National Disability Data Asset and the NHDH. None of these singular linkage systems can independently support comprehensive research and evaluation towards advancing life course health and wellbeing. There is much to gain from more coordinated and concerted action to bring components together and build systematically on them, promoting greater coverage, interoperability, harmonisation, standardisation, access and useability. Increasingly there is also greater collaboration between researchers and those who create and provide research infrastructure at scale. The Academy of Social Science's Research Infrastructure Decadal Plan illustrates this collaboration, but other examples include the Australian Research Data Commons development of the HASS and Indigenous Research Data Commons, and national projects to enhance data assets and undertake collaborative research across university and government researchers, including from the Australian Bureau of Statistics and the Australian Institute of Health and Welfare.

Examples of such current projects using the NHDH include studies of burden of vaccine preventable diseases and vaccine uptake, the epidemiology and health economics of Parkinson's disease, an analysis of relationship between support funding and life outcomes under Australia's National Disability Insurance Scheme, health trajectories of children with juvenile arthritis and high risk medication use in older adults [32]. Current projects using PLIDA include studies of health outcomes of people living with mental illness, of how Australia's Medicare Benefits Schedule and Pharmaceutical Benefits Scheme affects health equity, the development of a forecasting engine for health and aged care utilisation and financing, the impact of private health insurance on health expenditure and mortality, and the role of cancer treatments in cardiovascular health [33].

There is also growing recognition among researchers and data stewards of the need to include affected populations and communities in data governance, and data use. The Australian Government's Framework for the Governance of Indigenous Data [34] illustrates one approach. The building blocks for a more joined up national system are in place. Bringing planning, co-design, data assets, governance, platforms and services together in a more coordinated way would enable a comprehensive picture of Australians' engagement and touchpoints with government services and programs across all different environments and domains of social life over the life course. This is a fundamental requirement for a national data system to advance life course health and wellbeing. Expanding existing data linkages with rich, high quality survey data such as Australia's national household panel study and birth and education cohorts would further enrich data assets.

A National Data System would also enable advances in population data science. Elsewhere in the Australian system, researchers are developing infrastructure to capture and analyse social media and other "born digital" data such as sensor and tracker data. Infrastructure is also being developed to support computational linguistics through machine learning, generative AI and large-language models and to provide tools, datasets and documentation for historical, cultural and language data in libraries, museums and archives [11]. NCRIS funded research infrastructures in environment and climate increasingly look to include social, economic, demographic and community information, to inform environmental planning, policy development and monitoring. Because many of these infrastructures are supported through NCRIS and related programs, they incorporate the same principles around collaboration and partnering, open research and open standards, equity, access and interoperability, that informs the other research infrastructure described here.

These developments have several implications. First more large-scale and finely grained empirical research on populations and subpopulations will be possible around health and wellbeing and in subfields from disciplines across health and the social sciences. There will be better resources to study complex inequalities across class, socioeconomic status, gender, ethnicity, migration status and Indigeneity. Cross-domain topics such as the health and wellbeing implications of climate change, natural disasters and extreme weather events will become more researchable. Methods such as machine learning and AI will become more important for managing data, analysing data, modelling, forecasting and prediction. Researchers will increasingly be able to incorporate "space" and "place" into empirical studies in a more detailed and advanced way.

Beyond this, the expanding capacity to characterise populations and subpopulations more richly linked to data on interactions and transactions, may lead to a greater focus on simulation and modelling as well as empirical analysis and policy and program evaluation and intervention. However, research infrastructure should not develop in an unfettered way, even if it is technically possible to do so. Societal engagement and involvement is critical for public acceptance and social licence, and governance frameworks and safeguards need to be designed so that research with publicly funded research infrastructure is directed towards public and social benefits, public accountability, and harm minimisation, especially for vulnerable subpopulations.

References

1. World Health Organisation. Creating healthy life trajectories: universal health coverage and a life course approach. Geneva; 2023.
2. World Health Organisation. Framework to implement a life course approach in practice. Geneva; 2025.
3. Czech Presidency of the Council of the European Union. Brno Declaration on Fostering a Global System of Research Infrastructures. 2022.
4. Australian Government. 2021 National Research Infrastucture Roadmap Commonwealth of Australia; 2021.
5. Australian Government Department of Education. National Collaborative Research Infrastructure Strategy (NCRIS) 2024 [Available from: https://www.education.gov.au/ncris].
6. Australian Government Department of Education. National Research Infrastructure 2024 [Available from: https://www.education.gov.au/national-research-infrastructure]. 2024.
7. Population Health Research Network. Strategic Plan 2021-2030. Population Health Research Network; 2021.

8. Data Connect Australia. Types of Data: Data Connect Australia; 2026 [Available from: https://www.dataconnectaustralia.org.au/data/types-of-data/].
9. Australian Urban Research Infrastructure Network. AURIN 2026 [Available from: https://aurin.org.au/].
10. Australian Urban Research Infrastructure Network. Australian Urban Health Indicators 2026 [Available from: https://aurin.org.au/about-aurin/projects/ausurb-hi/].
11. Australian Research Data Commons. HASS and Indigenous Research Data Commons: Australian Research Data Commons; 2022 [Available from: https://ardc.edu.au/hass-and-indigenous-research-data-commons/].
12. Australian Bureau of Statistics. Person Level Integrated Data Asset (PLIDA) [Available from: https://www.abs.gov.au/about/data-services/data-integration/integrated-data/person-level-integrated-data-asset-plida].
13. Hansell D, Rafi B. Firm-Level Analysis Using the ABS' Business Longitudinal Analysis Data Environment (BLADE). Australian Economic Review. 2018;51(1):132-8. DOI: https://doi.org/10.1111/1467-8462.12253
14. Australian Institute of Health and Welfare. National Health Data Hub 2025 [Available from: https://www.aihw.gov.au/reports-data/nhdh].
15. Australian Institute of Health and Welfare. Australia's Disability Strategy 2021–2031 Outcomes Framework: 3rd Annual Report. Australian Institute of Health and Welfare; 2025.
16. Digital Transformation Agency. Australian National Data Integration Infrastructure, : Digital Transformation Agency, ; n.d. [Available from: https://architecture.digital.gov.au/design/australian-national-data-integration-infrastructure-andii].
17. Australian Burea of Statistics. Life Course Data Initiative: Australian Bureau of Statistics,; n.d. [Available from: https://www.abs.gov.au/about/key-priorities/life-course-data-initiative].
18. Lane J. Democratizing Our Data: A Manifesto: Massachussets Institute of Technology; 2020. DOI: https://doi.org/10.7551/mitpress/11990.001.0001
19. Office of the National Data Commissioner. Introducing the Data Scheme 2022 [Available from: https://www.datacommissioner.gov.au/the-data-scheme].
20. Australian Bureau of Statistics. "Five Safes Framework" 2021 [Available from: https://www.abs.gov.au/about/data-services/data-confidentiality-guide/five-safes-framework].
21. Department of Finance, Commonwealth of Australia. Statutory Review of the Data Availability and Transparency Act 2022 - Final Report. 2025.
22. Academy of the Social Sciences in Australia. Decadal Plan for Social Science Research Infrastructure 2024-33 2024 [Available from: https://socialsciences.org.au/projects/decadal-plan/].
23. Kukutai T, Taylor J, editors. Indigenous Data Sovereignty: Towards and Agenda: ANU Press; 2016. DOI: https://doi.org/10.22459/CAEPR38.11.2016
24. Maiam nayri Wingara. Indigenous Data Sovereignty Communique. 2018.
25. Australian Bureau of Statistics. TableBuilder: Administrative data snapshot of population and housing [Available from: https://www.abs.gov.au/statistics/microdata-tablebuilder/available-microdata-tablebuilder/administrative-data-snapshot-population-and-housing].
26. Australian Bureau of Statistics. Available Microdata and TableBuilder [Available from: https://www.abs.gov.au/statistics/microdata-tablebuilder/available-microdata-tablebuilder].
27. Australian Research Data Commons. Social Science Research Infrastructure Network 2026 [Available from: https://ardc.edu.au/project/social-science-research-infrastructure-network/].
28. Tomaszewski W, Zajac T, Curry M, Burton N. Social Science Research Infrastructure Network. Project Plan 2024 [Available from: https://ardc.edu.au/wp-

content/uploads/2025/07/public-ardc-project-plan-social-science-research-infrastructure-network-.pdf].
29. Carroll SR, Garba I, Figueroa-Rodríguez OL, Holbrook J, Lovett R, Materechera S, et al. The CARE Principles for Indigenous Data Governance. Data Science Journal. 2020;19. DOI: https://doi.org/10.5334/dsj-2020-043
30. Wilkinson MD, Dumontier M, Aalbersberg IJ, Appleton G, Axton M, Baak A, et al. The FAIR Guiding Principles for scientific data management and stewardship. Scientific Data. 2016;3(1):160018. DOI: https://doi.org/10.1038/sdata.2016.18
31. Australian Research Data Commons. Improving Indigenous Research Capability 2026 [Available from: https://ardc.edu.au/project/improving-indigenous-research-capabilities/].
32. Australian Institute of Health and Welfare. National Health Data Hub Approved Projects and Status 2025 [Available from: https://www.aihw.gov.au/reports-data/nhdh/current-approved-projects].
33. Australian Bureau of Statistics. Data Integration Project Register 2026 [Available from: https://www.abs.gov.au/statistics/data-integration/data-integration-project-register#all-projects-detailed-download].
34. Commonwealth of Australia. Framework for Governance of Indigenous Data. Commonwealth of Australia; 2024.

Data Availability Statement

No research data were used in this paper.

Acknowledgements

This research was supported by the Australian Government through the Australian Research Council's Centre for Excellence for Children and Families Over the Life Course [Project ID CE200100025]. We thank the reviewers and editor for their exceedingly helpful comments.

Statement of Conflict of Interest

Mark Western is a member of the National Research Infrastructure Advisory Group to the Australian Government and a Director of the Board of the Academy of Social Sciences in Australia. The views in this paper are his individual views and do not represent the views of the National Research Infrastructure Advisory Group or the Academy of Social Sciences in Australia.

Ethics Statement

This is a case study drawing on published materials only. There is no research with human participants. Ethical approval has not been sought.

AI Disclosure Statement

The authors declare that no generative AI tools were used in the preparation of this manuscript.