

Digital Transformation Of Healthcare In Australia Constrained - A Call To Action For A National Data Governance Framework

Frean, I., Belgard, M., Zeps, N., Boyd, J., Shaw, T., Cavedon, L., Gray, L.

Authors and Affiliations

Isobel Frean¹, Matt Belgard², Nikolajs Zeps², James Boyd³, Tim Shaw⁴, Lawrence Cavedon⁵, Len Gray⁶.

1. Digital Health Cooperative Research Centre, Sydney, NSW, Australia
2. eResearch Office, Faculty of Health, Queensland University of Technology, Brisbane, Australia
3. La Trobe University, Bundoora, Victoria, Australia
4. Charles Perkins Centre, University of Sydney, Sydney, New South Wales, Australia
5. School of Computing Technologies, RMIT, Melbourne, Australia
6. Centre for Health Services Research, University of Queensland, Brisbane, Queensland, Australia

Notices

This publication is provided by Digital Health CRC Limited funded under the Australian Commonwealth Government's Cooperative Research Centres Programme.

This publication is made available under the Creative Commons Attribution-NonCommercial 4.0 International License. The full text of the license is available at: <https://creativecommons.org/licenses/by-nc/4.0/deed.en>



Outline

Executive Summary	04
Introduction	06
DHCRC Data Governance Project Experiences	08
Discussion	14
A Call To Action	17





Executive Summary



Executive Summary

Australia lacks the national data and information governance tools to deliver the scale of healthcare transformation required to support effective responses to population health challenges such as pandemics, population ageing and personalised care. Our effective use of healthcare data, whether in support of the delivery of care or accessed for clinical research, is hampered by a lack of national harmonisation around information governance frameworks and resources to support consistent interpretation.

We reflect here on the challenges experienced by Australian researchers linking a variety of health datasets to inform the response to the SARS-CoV-2 pandemic and to better predict care trajectories in chronic disease management. We examine data quality and data processing challenges faced by researchers responding to calls from the Royal Commission into Aged Care Quality and Safety and for smarter use of data held in electronic health records to offer further evidence for why we must act now around a national data and information governance framework.

These reflections suggest we are still weighed down by defensive policies, legislation and mindsets related to information governance. Here we argue for a renewed call for action around information governance. We re-assert that Australia requires a harmonised set of national health, medical technology, and pharmaceutical research governance arrangements along the lines of those advocated by researchers in 2018; to address the additional compliance barriers created by the myriad laws in each state and territory; and a contemporary capability to make available tools and templates compatible with national data and information governance requirements, codes and standards for those looking to use data responsibly and innovate on behalf of Australian healthcare consumers.





Introduction



Introduction

In 2018 Australian researchers in digital health cited the many barriers hampering our world class health, medical technology, and pharmaceutical sector from delivering solutions that benefit Australia and Australian healthcare consumers.¹ They suggested that Australia lags other countries such as the USA and UK by not having a clearly documented national information governance framework. They issued a call to action for a streamlined approach to information governance that embeds privacy, security, and confidentiality by design. Their assessment and call to action are summarised in Box 1 (see Appendix).²

In this paper, we reference five Digital Health Cooperative Research Centre (DHCRC) funded projects to illustrate data integration and interoperability problems researchers continue to face and the strategies and tools that can be used to address them. We identify the digital health priorities we believe will assist jurisdictions, clinical researchers, technology, health, and aged care providers to better use, share and link data for smarter outcomes.

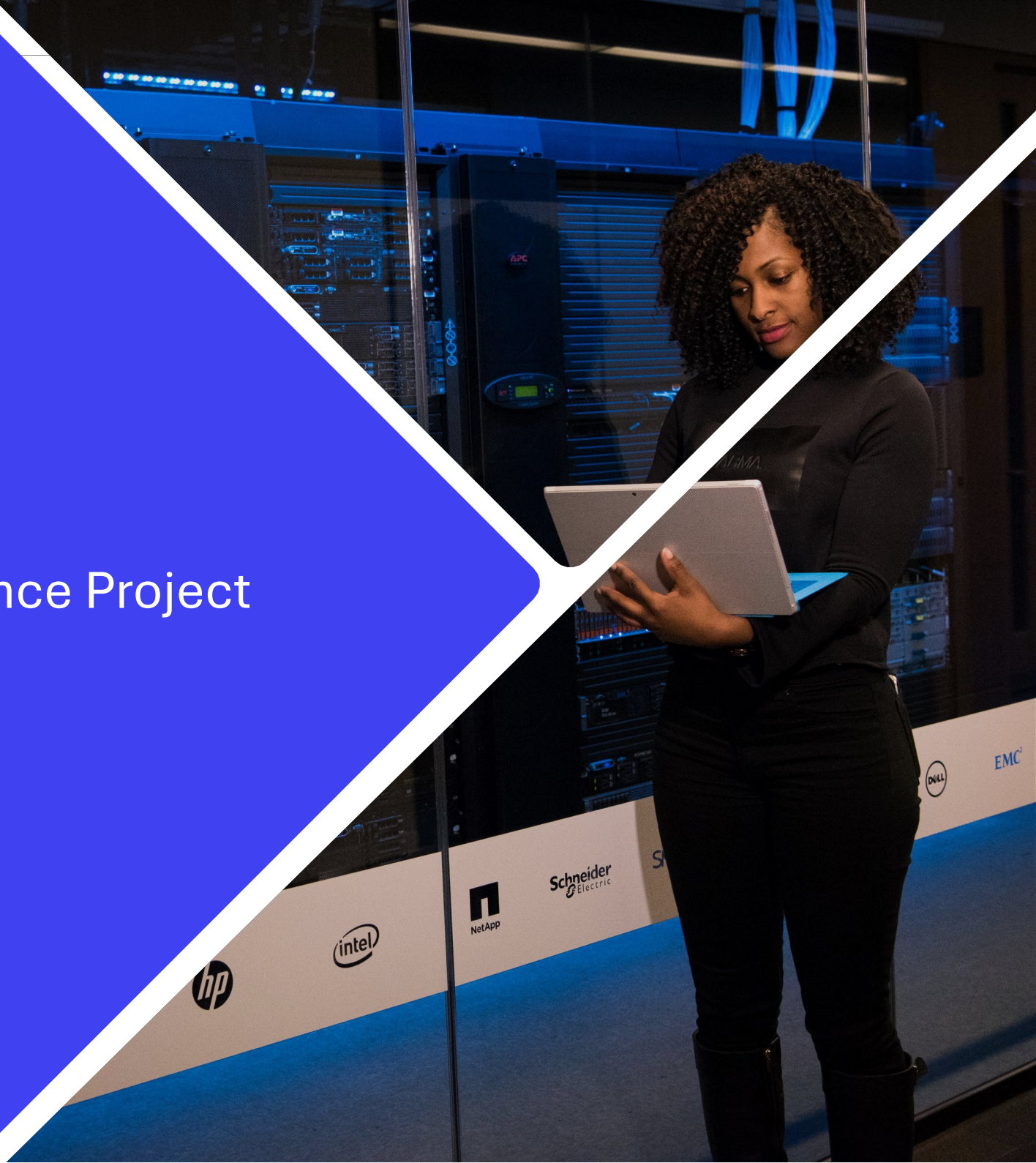


¹ Srinivasan, U., Ramachandran, D., Quilty, C., Rao, S., Nolan, M., and Jonas, D., 2018, Flying Blind: Australian Researchers and Digital Health, Volume 2: Health Data Series, Digital Health Cooperative Research Centre, Sydney. ISBN: 978 - 0 - 646 - 99212 - 9

² Ibid



DHCRC Data Governance Project Experiences



DHCRC Data Governance Project Experiences

In this section, we will review interoperability, data integration and data governance challenges faced by five DHCRC funded projects. Two of these projects share the common problem of secure linkage of large, disparate datasets, whilst three projects attempt to overcome inconsistencies in the content and logical structure of data collected to measure and support clinician decision making and patient outcomes. Whilst different solutions have been employed, these projects highlight the need to build data standards into underlying systems in which clinical data is collected to ensure Australian healthcare data is research ready and available to deliver timely point of care and population health outcomes.

We do not report here on the findings from these projects as they will be made available elsewhere.³



Clinical Data and Analytics Platform (DHCRC Project 0096)

Chief Investigator: Matt Bellgard

Queensland University of Technology

In April 2020, the DHCRC funded a collaborative project in partnership with Queensland University of Technology (QUT), Monash University, University of Sydney, Commonwealth Department of Health, Queensland Health, and the Ministry of Health NSW to demonstrate specific key features of the Clinical Data and Analytics Platform (CDAP) in response to the emerging COVID-19 epidemic. High-level project objectives involved testing data capture of clinical and patient reported data, validation of COVID-19 specific Bayesian Network decision support models and articulation of governance requirements for scaling the platform and others like it.

Multiple data custodians each requiring bespoke responses provided the main governance challenges faced by researchers. These challenges resulted in significant delays in accessing Australian public datasets. The strategies used to compensate for time delays, included acquisition of additional data datasets from interstate (Victorian Department of Health) and the UK NHS COVID-19 data (ISARIC-4C⁴); the use of a dynamic privacy preserving ingestion tool; and ongoing validation of the CDAP platform functionality using European linked datasets (IDDO⁵, LEOSS⁶).

³ For further information on each of these projects, please visit the DHCRC website <https://digitalhealthcrc.com/>.

⁴ <https://web.www.healthdatagateway.org/dataset/f06ec631-77d0-4b12-a21f-f11e7af49ba5>

⁵ <https://www.iddo.org/about-us/about-iddo>

⁶ <https://leoss.net/%20%20European%20COVID-19%20open%20data%20set>

DHCRC Data Governance Project Experiences

The project has contributed six validated Bayesian Network models to a public repository of such models; Python scripts, utility tools and instructions for generating causal Bayesian Networks compatible with international data sharing formats to two open research environments. A web-based, publicly available COVID Calculator (decision support tool) has been deployed for academic use. The researchers have also provided a report to the participating jurisdictions detailing improvements in their data governance processes and similar website guidance and reforms to the way the Public Health Act 2005 (PHA) is applied for granting access to health data for research purposes without consent.

Regrettably, the learning from this project is that multiple data custodians and inconsistent interpretation continues to stymie national projects and to create avoidable, costly bureaucratic hurdles. This was even more frustrating for researchers looking to contribute vital national population health information at the beginning of the pandemic. The previously stated solution they assert is to harmonise governance frameworks and their interpretation at a national level, incorporating clearly defined codes-of-practice for data sharing involving taxpayer funded clinical data. Fortunately, Queensland Health have recognised this issue and endorsed work through CDAP to deliver the necessary reforms demonstrating the value of the DHCRC investment.



Reviewing and managing chronic kidney disease to improve outcomes (DHCRC Project 0073)

Chief Investigator: Delia Hendrie (Suzanne Robinson)
Curtin University

In March 2020 the DHCRC funded a collaborative project in partnership with Curtin University Western Australia Department of Health, Western Australia Country Health Service (WACHS) and Western Australia Primary Health Alliance (WAPHA), to use linked data to determine the incidence, prevalence, progression, and economic burden of chronic kidney disease (CKD) in Western Australia. Analysis has focused on investigating factors associated with poor outcomes of CKD, with the aim of development models of care focused on improving health system usage and health practices that prioritise early intervention and improved patient outcomes.

This project knowingly took on the challenge of having to address multiple data sharing agreements in order to access pathology data. Unaccustomed to sharing their datasets, three pathology providers benefited from technical expertise within the DHCRC project participants that underpinned the separate negotiations for state-wide data access and sharing agreements with each of the major pathology labs.

DHCRC Data Governance Project Experiences

This project has been an exemplar case study for the application of Privacy Preserving Record Linkage (PPRL) tools and associated training. PPRL is a methodology for de-identifying personally identifiable information within a data set by creating one or more unique tokens that replace the identifiable data components, in turn enabling the secure matching of patient records held in other data sets. The particular PPRL model developed by the project researchers has been designed to be scalable and includes an encoding capability (as a standalone tool or as an API) for data custodians.⁷

A key feature of the project involves access to data from multiple systems to generate discipline benchmarks that allow for comparison of performance and access to underlying cases to support learning. The project will take learnings from seven PhD research projects that look at the lifecycle of data usage in supporting reflective practice including indicator development, data visualisation, data sense making and medico-legal considerations. Central to this project is how data can be accessed and compared within and across systems and services in Australia's private hospital sector.



Using practice analytics to understand variation and support reflective practice (DHCRC Project 0056)

Chief Investigator: Tim Shaw
University Sydney

In January 2020, the DHCRC funded a collaborative project in partnership with the University of Sydney, Monash University, Swinburne University of Technology, Cabrini Health, Adventist Healthcare, St John of God Health Care, Epworth Healthcare, Ramsay Hospital Research Foundation, Royal Australasian College of Physicians (RACP) and the Royal Australasian College of Surgeons (RACS). The aim of this project is to support clinicians, teams, and organisations to review their performance to support reflective practice, a requirement of Clinical Governance Standards and the Medical Board of Australia's Professional Performance Framework.

Underpinning the governance challenges for this project is that clinicians still have limited access to data to compare performance, especially in private settings. The most accessible data is contained in patient administration systems (PAS) systems, which lack clinical data points which support the measurement of many quality indicators. Other sources of clinical data access include registries which are often inaccessible and administratively challenging to access. Lack of clear governance and privacy standards associated with the use of these diverse data sources for performance feedback limit the ability to systematically develop resources and tools to support practice reflection and quality improvement across Australia's extensive private hospital sector.

⁷ Lim, D et al. Unlocking Potential within Health Systems Using Privacy-Preserving Record Linkage: Exploring Chronic Kidney Disease Outcomes through Linked Data Modelling. *Appl Clin Inform* 2022; 13: 901-909

DHCRC Data Governance Project Experiences

A number of jurisdictions and private organisations are developing systems that will support reflection and comparison within and between organisations however ability to do this in near real time and integrate this into clinical workflows to impact on practice is currently limited because of the paucity and lack of standardisation of the data recorded.



Predicting resident deterioration and acute care needs in aged care (DHCRC Project 0013)

Chief Investigator: Lawrence Cavedon

RMIT

In November 2019, one of the early DHCRC funded projects involved collaboration between RMIT University and Telstra Health to leverage residential aged care system data to develop algorithms to provide advanced indication of deteriorating condition. We draw here on work undertaken by the RMIT research team to highlight machine learning strategies that can be used to optimise secondary use of electronic health record (EHR) data which is not standardised and comprises structured and unstructured data formats. Their elegant and comprehensive paper⁸ provides useful understanding of the many data preparation and processing techniques required before machine learning algorithms can be employed.

Early on in this project the research team identified that important routine information about residents were contained in text-based notes which was hard to access. This required the project team to agree on data processing techniques, including addressing the challenge of how to de-identify the free-text progress notes.

Text-based data was converted into structured (numeric) features involving manual intervention to identify similar information recorded under different data categories. Once transformed, data could then be used with machine learning models. The transformed data however are application-specific and unlikely to scale to other applications.

The project undertook a comprehensive abstract analysis of research papers on EHR to understand which data types and data mining strategies are frequently used for EHR data, confirming the importance of unstructured data (clinical notes) to health applications. This paper⁹ includes a summary of the characteristics of data held in EHRs and the associated data mining strategies for addressing these.

⁸ Tabinda Sarwar, Sattar Seifollahi, Jeffrey Chan, Xiuzhen Zhang, Vural Aksakalli, Irene Hudson, Karin Verspoor, and Lawrence Cavedon. The Secondary Use of Electronic Health Records for Data Mining: Data Characteristics and Challenges. *ACM Comput. Surv.* 55, 2, Article 33 (January 2022), 40 pages. <https://doi.org/10.1145/3490234>

⁹ Ibid

DHCRC Data Governance Project Experiences



Aged Care Data Compare Project (ACDC) (DHCRC Project 0078)

Chief Investigator: Len Gray
University of Queensland

In June 2020, the DHCRC funded a collaborative project in partnership with University of Queensland, Bupa Health Foundation, and the Commonwealth Department of Health to produce and validate a prototype data hub to facilitate interoperability of aged care assessment data across residential aged care facilities (RACFs) that use different aged care IT solutions to enable benchmarking of provider quality indicators. Two sets of data standards were selected, HL7 FHIR standards that pertain to the technical task of sharing information and standards that pertain to information content - in this case functional and psychosocial information configured by interRAI.¹⁰

Because there is no standard dataset collected by all aged care providers in a comparable format, there is a plethora of descriptors used by software vendors and providers to characterise resident characteristics, care requirements and outcomes. This lack of standardisation of aged care data is a major barrier to comparing outcomes within and among organisations. To address this lack of a data standard¹¹, the project identified the interRAI LTCF system as a standard dataset that could be used to meet data content requirements, risk adjustment and calculation of quality indicators.

The project identified HL7 FHIR¹² content messaging solution to transmit this data to a central repository. Using these data content and data exchange standards respectively, the project has built a secure repository to enable storage and manipulation of data. A suite of presentation tools to enable performance comparison across providers for benchmarking is under development.

The project is moving into implementation mode but continues to experience challenges that are attributable to the lack of an aged care data governance framework and defined data standards, and to market uncertainty. A wide group of stakeholders must agree to participate and contribute. Some require external resources and there is anxiety about whether the solution will be compatible with future national regulatory and compliance approaches that are yet to be determined.

A scalable technology solution is about to be deployed and tested in the field under the next phase of the project, supported by a wide group of industry participants. The project, ACDC Plus, will offer a vendor-neutral SMART on FHIR application to extract data from one aged care application. It is hoped that such a solution, when coupled with an agreed approach to the structuring of clinical data can deliver quality indicators of care outcomes that are risk adjusted and which are calculated in near-real time, in turn helping to demonstrate to consumers and care providers alike what good quality care looks like.

¹⁰ <https://interrai.org/>

¹¹ Long Term Care Facility (LTCF)

¹² [HL7 FHIR](#) is the fastest growing interoperability standard globally for sharing healthcare data



Discussion



Discussion

DHCRC Project 0096 (CDAP) and DHCRC Project 0073 (CKD) both faced common challenges in accessing data held in data silos and governed by multiple data custodians imposing their own requirements for ethics and governance approval to enable data access. Both projects secured data sharing agreements. In the case of CDAP recommendations for improvements in jurisdictional governance have been provided; and CKD achieved state-wide agreements with each of the WA pathology labs, offering the potential for harmonisation at both state and national level. Key to these agreements was the offer of technical support through use of the privacy preserving tools, suggesting governance frameworks need to be supported by pragmatic tools. In addition, a change management process involving extensive consultancy and review of existing data governance frameworks was required that would not have been possible without the DHCRC program.

DHCRC Project 0056 (Practice Analytics) project is exploring a range of options to make data available for performance benchmarking for clinicians in private hospitals. Data governance solutions are likely to rely upon harmonisation of data capture and sharing codes of practice championed by clinical leaders.





Discussion

Both aged care projects, DHCRC Project 0013 (RMIT) and DHCRC Project 0078 (ACDC), have attempted solutions to inform care quality outcomes in the knowledge that the majority of routinely recorded data is not consistently standardised due to the use of customisable templates tailored to the (perceived) needs of individual facilities. Whilst both projects highlight the urgency for an aged care data strategy, with mandated datasets and formats for data capture, a comprehensive abstract analysis of research papers on EHR by RMIT confirmed the importance of accessing unstructured data (clinical notes) in aged care records.

Given increased use of machine learning techniques, such as natural language processing, evaluation of performance and the predictive power of structured and unstructured data is required to identify the applications where particular data types can safely and effectively be used. This will be particularly important in aged care where the temptation will be to mine unstructured data to deliver tactical intel for clinicians and carers in the short term without regard to the need to capture such information as part of the longitudinal care record.




A Call To Action





A Call To Action

The research projects showcased above set out to deliver digital health solutions to inform better care outcomes. Each have consistently encountered barriers due to the continued lack of harmonisation around data governance frameworks within private health and aged care and across jurisdictions relating to public healthcare data. These were barriers that had been identified in 2018 and which five years later have yet to be adequately addressed.

A national data and information governance framework for the sharing and joining of all health and aged care data needs to comprise:

-  Policies, guidelines, templates/tools, codes of practice and potentially national infrastructure to transparently assist in the decision-making process for the trustworthy use of clinical data.

-  Legislation that is harmonised with existing regulations and principles (such as the National Privacy Principles); and

-  Governance to ensure that there is buy-in and compliance by all stakeholders and the community.

Australia needs to act on two fronts to put into place a contemporary national data and information governance framework. We need immediate action to support today's researchers and action to deliver long term transformation.

A. Immediate, short-term solutions to deliver tangible support to harmonise data and information governance should include the following:

Support for the registering of data sharing agreements under the National Data Availability and Transparency Act 2022 and promotion of accreditation for data linkage entities



Drafting and testing of national principles for the consistent use and disclosure of personal data for research to underpin data protection legislation in every state and territory



Uniform respect and implementation by universities and jurisdictions of the National Mutual Agreements for ethics approvals.



These are deliverables that should include a national discussion, hosted by the Department of Health and Aged Care, involving Australia's national digital health and research agencies and representatives from each of Australia's universities and research institutes.

A Call To Action

B. We also need to define an achievable, more urgent, time frame for a sustainable and contemporary national framework for the governance of data and information, explicitly to enable:



Responsive research that can inform decisions on national population health issues for public good (such as the next epidemic);



Real time access to clinical data to better support evidence-based medicine, care delivery and coordination in the face of workforce challenges and resource constraints; and



Capacity to evaluate digital health solutions in diverse settings to lend confidence to the efficiency and trustworthiness of solutions relying upon artificial intelligence data models.

This transformative national approach to data and information governance needs to be led nationally and should include agreement on:



A clearly defined national risk-based framework for providing access to government data sets (e.g., AIHW) in de-identified form for trusted researchers without complex approval processes.



A formal program of research to feed into the three-year review of the Data Availability and Transparency Act 2022 to ensure timely amendment and extension, where the evidence supports the effectiveness of the Act.



National consent and authorisation policies that enable health and medical research AND clinical (EHR) data to flow continually and routinely across all points of the health system, from clinical points of service all the way to researchers, without compromising individual privacy.



The development and maintenance of a rich reusable national dataset (distributed) enabling researchers and developers of health and medical technology and pharmaceutical sectors to be massively more productive and drive innovation.

Australia can't afford to wait another five years to provide certainty for clinical researchers and innovators, those looking to run clinical trials or those responding to the next epidemic, nor can we continue to put a national data and information governance framework onto the back burner for future governments to address.

DHCRC snapshot

Our Purpose

The Digital Health Cooperative Research Centre (DHCRC) fosters innovation and facilitates the greater utilisation of data and digital technologies in healthcare to improve health outcomes, increase efficiencies in health and aged care delivery, and support the growth of a competitive Australian digital health technology industry. We do this by connecting universities, government, the digital technology industry and the healthcare sector to accelerate the implementation and translation of digital health technologies.



Our Focus Areas

Working with our Participant organisations and the broader digital health ecosystem, DHCRC works across three distinct, but inter-related, program areas:

Research and Development

Virtual care

including telehealth, hospital- in-the-home, remote monitoring and new models of care.

Effective use of data including clinical decision analytics, data standards, interoperability and data linkage.

Advancing a **sustainable and connected** health and aged care system including regional **and remote care**.

Education & Capacity Building

Supporting the next generation of digital health leaders and upskilling the current workforce.

The DHCRC has a highly successful education and training program that is delivering new and innovative learning opportunities.

Translation and Commercialisation

Developing the evidence base for digital technology to support the creation of new products and services and drive innovation.

Growing Australia's digital health technology sector, positioning Australia as a global leader in digital health innovation.

\$110M

Commonwealth and Participant funding from 2018-2026

60+

Industry, Government and Academic participant organisations across Australia

48

PhD and Masters students supported

60+

Research projects in delivery or completed



Australian Government
Department of Industry,
Science and Resources

AusIndustry
Cooperative Research
Centres Program

References

Lim, D et al. Unlocking Potential within Health Systems Using Privacy-Preserving Record Linkage: Exploring Chronic Kidney Disease Outcomes through Linked Data Modelling. *Appl Clin Inform* 2022; 13: 901-909

Md Momin Al Aziz, Md Nazmus Sadat, Dima Alhadidi, Shuang Wang, Xiaoqian Jiang, Cheryl L Brown, Noman Mohammed, Privacy-preserving techniques of genomic data—a survey, *Briefings in Bioinformatics*, Volume 20, Issue 3, May 2019, Pages 887–895, <https://doi.org/10.1093/bib/bbx139>

Srinivasan, U., Ramachandran, D., Quilty, C., Rao, S., Nolan, M., and Jonas, D., 2018, *Flying Blind: Australian Researchers and Digital Health*, Volume 2: Health Data Series, Digital Health Cooperative Research Centre, Sydney. ISBN: 978 - 0 - 646 - 99212 – 9

Tabinda Sarwar, Sattar Seifollahi, Jeffrey Chan, Xiuzhen Zhang, Vural Aksakalli, Irene Hudson, Karin Verspoor, and Lawrence Cavedon. The Secondary Use of Electronic Health Records for Data Mining: Data Characteristics and Challenges. *ACM Comput. Surv.* 55, 2, Article 33 (January 2022), 40 pages. <https://doi.org/10.1145/3490234>

Tayefi M, Ngo P, Chomutare T, et al. Challenges and opportunities beyond structured data in analysis of electronic health records. *WIREs Comput Stat.* 2021;13:e1549. <https://doi.org/10.1002/wics.1549>

Appendix

Box 1. Flying Blind report call to action

What is needed:

- Government policy regarding data release that reflects consumer sentiment because Consumers are willing to share their health data to support research.
- A well-documented governance framework that is transparent to researchers and enables them to access de-identified HMR datasets to address their specific research questions.
- A commitment to streamline secure data flow across state borders and jurisdictional boundaries for research.
- Policy incentives for data custodians to prepare data for research readiness.
- Standardised methodologies and technologies that support secure research environments that preserve privacy and confidentiality.
- Strategy for harnessing unstructured health data for health system research.
- Streamline processes to allow real-time de-identified service data to flow into research environment and enable research findings to flow back to pro-actively influence policy formulation and support evidence-based real-time service delivery.
- Vibrant HMR environment.

Actions required:

1. Develop and maintain a rich reusable national dataset (distributed) enabling researchers and developers of health and medical technology and pharmaceutical (HMR) sectors to be massively more productive and drive innovation.
2. Formulate policies that enable HMR data flows continually and routinely across all points of the health system, from clinical points of service all the way to researchers, without compromising individual privacy.
3. Clearly define risk-based frameworks providing access to government data sets (eg AIHW) in de-identified form for trusted researchers without complex approval processes.
4. Seek uniform respect and implementation by universities and jurisdictions to National Mutual Agreements for ethics approvals.
5. Design a national security and privacy framework for secure data management and for state-of-the art standardised methodologies to ensure data privacy and confidentiality.
6. Expedite implementation of the Data Sharing and Release Act for the health sector.